More Lectures on Geophysical Fluid Dynamics

Rick Salmon

Scripps Institution of Oceanography
University of California, San Diego

—

published online November 2020

Preface

This book is a sequel to my earlier book *Lectures on Geophysical Fluid Dynamics* published more than 20 years ago. Unlike *LGFD* the present book assumes a substantial background knowledge of geophysical fluid dynamics, such as that which could be obtained from *LGFD* or from any of the other introductory books. The emphasis here is on topics that have seen significant development in the past twenty years or were incompletely covered in the earlier book. The primary focus is on Lagrangian methods and variational principles.

This is a work of pedagogy in which I make no effort to cite original sources. Although some elements are new, most of the ideas presented here have been percolating in the field for a considerable time. I have sprinkled the text with some of the great names whose work has most most influenced me. The list will get longer as more chapters are added.

This is admittedly a covid project, undertaken with the advice and encouragement of my friend and colleague Nick Pizzo.

Rick Salmon
in the pandemic, October 2020

Contents

Chapter 1

Mean flow generated by a surface wave packet

We consider a surface gravity wave packet propagating to the right (in the direction of increasing x) with y increasing upward and $y = 0$ corresponding to the surface of the water when at rest (figure 1.1). The water velocity under the crests is rightward, in the direction of propagation. The water velocity under the troughs is leftward. The velocity field associated with the crests and troughs decays rapidly with depth. At a depth equal to the wavelength of the wave it has mostly disappeared. The crests and troughs move at the deep-water phase speed

$$c = \frac{\omega}{k} = \sqrt{\frac{g}{k}} \tag{1.1}$$

while the wave packet itself moves at the group velocity

$$c_g = \frac{d\omega}{dk} = \frac{1}{2}\sqrt{\frac{g}{k}} \tag{1.2}$$

Such a wave packet is easily observed in the wave tank at the Hydraulics Lab.

Less easy to observe is the field of mean velocity that is attached to the wave packet. This mean velocity is much smaller than the oscillatory fluid velocity within the wave packet itself, but the field of mean velocity is much broader. In the broader perspective of figure 1.2, the wave packet is but a speck near the surface of the water. The mean flow includes the Stokes drift in the wave packet itself, but also the much broader and deeper recirculating

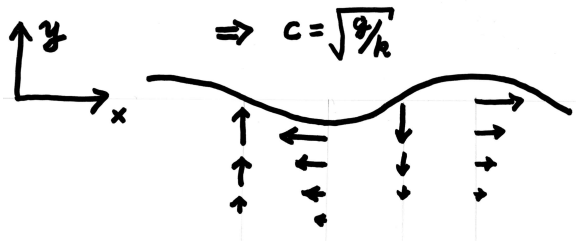

Figure 1.1: A close-up of the crests and troughs inside the wave packet.

flow outside the wave packet. To observe this mean flow, you need to seed the wave tank with neutrally buoyant beads.

The mean-flow pattern is attached to the wave packet and hence this *pattern*—not the fluid itself!—moves at the group velocity of the wave packet. But something strange happens if the waves in the packet experience breaking. If breaking occurs, removing (say) half the energy of the wave packet, then half the mean-flow pattern is jettisoned by the wave packet at the point of breaking and remains behind as a semi-permanent, stationary flow (figure 1.3). The rest of the mean-flow pattern continues to be carried along by the wave packet. This somewhat surprising behavior can also be observed in the wave tank.

The purpose of this chapter is to explain these observed features. We shall do so by solving the fluid equations in Lagrangian coordinates. In the Lagrangian description of the fluid, the independent variables are a set of fluid particle labels, (a, b), and the time τ. The dependent variables are the locations $x(a, b, \tau)$ and $y(a, b, \tau)$ of the fluid particle labeled (a, b) at the time τ. The pressure is $p(a, b, \tau)$. Each fluid particle keeps its same value of (a, b) as it moves around. Thus the value of (a, b) tells us which fluid particle we are considering. We use the special symbol τ for time to emphasize that $\partial/\partial\tau$ (with a and b held fixed) is the time derivative following a fluid particle. Thus

$$\frac{\partial}{\partial\tau} \equiv \frac{D}{Dt} \tag{1.3}$$

The labels can be arbitrarily assigned, but they must vary continuously throughout the flow, and the mapping from (a, b) to (x, y) must be one-to-one at all times. A common choice is to assign the (a, b) to equal (x, y) at some 'labeling time' in the past. We adopt this convention, and we further assume that, at the labeling time, the fluid was in a state of rest. The state

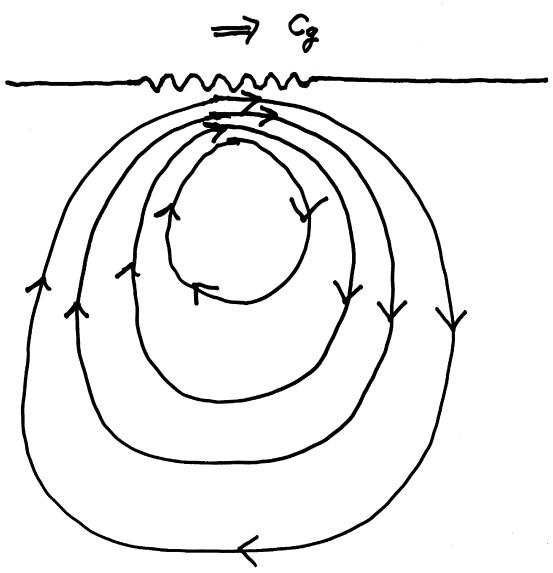

Figure 1.2: A broad recirculating mean flow is carried along by the wave packet at its group velocity. This figure is a poor imitation of a figure drawn by Michael McIntyre in 1981.

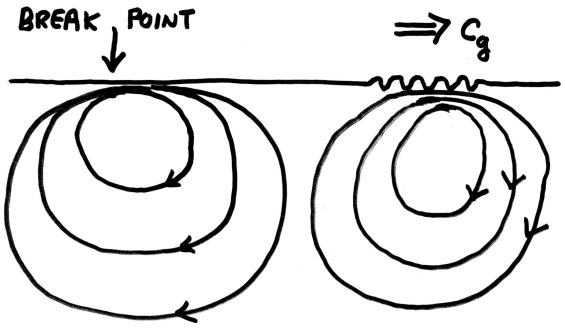

Figure 1.3: A portion of the mean-flow pattern is jettisoned by the wave packet at the point of breaking, and remains behind as a semi-permanent, stationary flow.

of rest is described by

$$x(a,b,\tau) = a \tag{1.4}$$
$$y(a,b,\tau) = b \tag{1.5}$$
$$p(a,b,\tau) = -gb \tag{1.6}$$

where g is the gravity constant. The fluid occupies the domain $b < 0$, and the free surface corresponds to $b = 0$. Since every fluid particle conserves its labels, and since the fluid particles on the free surface remain there, the free surface corresponds to a fixed boundary, $b = 0$, in Lagrangian coordinates. The free surface elevation at time τ is $y(a, 0, \tau)$. To express the free surface elevation in the conventional manner as $y = y(x, t)$, we would need to eliminate a between $x(a, 0, \tau)$ and $y(a, 0, \tau)$.

To get started we must write the equations of motion in Lagrangian coordinates. The equations are

$$x_{\tau\tau} = -\frac{\partial(p,y)}{\partial(a,b)} \tag{1.7}$$
$$y_{\tau\tau} = -\frac{\partial(x,p)}{\partial(a,b)} - g \tag{1.8}$$
$$\frac{\partial(x,y)}{\partial(a,b)} = 1 \tag{1.9}$$

where subscripts denote partial derivatives, and

$$\frac{\partial(F,G)}{\partial(a,b)} \equiv F_a G_b - F_b G_a \tag{1.10}$$

for any F and G. The fluid velocity is (x_τ, y_τ), and the acceleration is $(x_{\tau\tau}, y_{\tau\tau})$. Equation (1.9) states that the fluid is incompressible. Recalling that the Jacobians can be manipulated like fractions, we see that

$$\frac{\partial(p,y)}{\partial(a,b)} = \frac{\partial(a,b)}{\partial(x,y)}\frac{\partial(p,y)}{\partial(a,b)} = \frac{\partial(p,y)}{\partial(x,y)} = \frac{\partial p}{\partial x} \tag{1.11}$$

In Lagrangian coordinates, the acceleration terms are linear terms, and the pressure terms are nonlinear. This is just the opposite of the situation in Eulerian coordinates. The boundary condition is $p = 0$ at the free surface, $b = 0$. Note that this is the *only* boundary condition; the so-called kinematic boundary condition is automatically satisfied in Lagrangian coordinates. This is a primary advantage of the Lagrangian formulation.

The equations of motion are (1.7)-(1.9). Any solution of these equations is a physical possibility. However, we are interested in solutions with vanishing vorticity, because only such solutions can be excited from rest by pressure forces. The condition of vanishing vorticity is

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \frac{\partial(x_\tau, x)}{\partial(a,b)} + \frac{\partial(y_\tau, y)}{\partial(a,b)} = 0 \tag{1.12}$$

We solve the Lagrangian equations (1.7)-(1.9) with auxiliary condition (1.12) by pedestrian perturbation theory. We expand

$$x = a + x^{(1)}(a,b,\tau) + x^{(2)}(a,b,\tau) + \cdots \tag{1.13}$$

$$y = b + y^{(1)}(a,b,\tau) + y^{(2)}(a,b,\tau) + \cdots \tag{1.14}$$

$$p = -gb + p^{(1)}(a,b,\tau) + p^{(2)}(a,b,\tau) + \cdots \tag{1.15}$$

The leading terms represent the state of rest. At the next order, the wave packet will appear. At the order after that, we obtain the mean flow. The superscript-(1) variables are $\mathrm{O}(A)$, where A is the amplitude of the wave. The superscript-(2) variables are $\mathrm{O}(A^2)$. We substitute the expansions (1.13)-(1.15) into (1.7)-(1.9) and collect like powers of A. At the leading order, all the terms cancel because the state of rest satisfies the equations. At $\mathrm{O}(A)$ we obtain

$$x^{(1)}_{\tau\tau} = -g y^{(1)}_a - p^{(1)}_a \tag{1.16}$$

$$y^{(1)}_{\tau\tau} = +g x^{(1)}_a - p^{(1)}_b \tag{1.17}$$

$$x^{(1)}_a + y^{(1)}_b = 0 \tag{1.18}$$

The solution of interest is

$$x^{(1)} = A e^{kb} \sin(ka - \omega\tau) \tag{1.19}$$

$$y^{(1)} = -A e^{kb} \cos(ka - \omega\tau) \tag{1.20}$$

$$p^{(1)} = 0 \tag{1.21}$$

where A, k and ω are constants, and $\omega^2 = gk$. The result $p^{(1)} = 0$ seems bizarre, but remember that the full pressure at this order is $p = -gb + p^{(1)} = -gb$ and that b is a fluctuating function of (x, y, t).

At the next order, O(A^2), we obtain

$$x^{(2)}_{\tau\tau} = -gy^{(2)}_a - p^{(2)}_a \tag{1.22}$$

$$y^{(2)}_{\tau\tau} = +gx^{(2)}_a - p^{(2)}_b \tag{1.23}$$

$$x^{(2)}_a + y^{(2)}_b = A^2k^2e^{2kb} \tag{1.24}$$

At this order, the wave appears as a forcing term in the mass conservation equation (1.24). We also need to enforce the requirement (1.12) of vanishing vorticity. Substituting (1.13)-(1.15) into (1.12) we obtain

$$\left[\frac{\partial y^{(1)}_\tau}{\partial a} - \frac{\partial x^{(1)}_\tau}{\partial b}\right] + \left[\frac{\partial y^{(2)}_\tau}{\partial a} - \frac{\partial x^{(2)}_\tau}{\partial b} + \frac{\partial(x^{(1)}_\tau, x^{(1)})}{\partial(a,b)} + \frac{\partial(y^{(1)}_\tau, y^{(1)})}{\partial(a,b)}\right] = 0 \tag{1.25}$$

and substituting (1.19)-(1.21) into (1.25) we obtain

$$[\ 0\ ] + \left[\frac{\partial y^{(2)}_\tau}{\partial a} - \frac{\partial x^{(2)}_\tau}{\partial b} + 2\omega k^2A^2e^{2kb}\right] = 0 \tag{1.26}$$

Our task is to solve (1.22)-(1.24) and (1.26) for the superscript-(2) variables. A *particular* solution to (1.22)-(1.24) is

$$x^{(2)} = 0, \qquad y^{(2)} = \frac{1}{2}A^2ke^{2kb}, \qquad p^{(2)} = 0 \tag{1.27}$$

A *homogeneous* solution to (1.22)-(1.24) is

$$x^{(2)} = U(b)\tau, \qquad y^{(2)} = 0, \qquad p^{(2)} = 0 \tag{1.28}$$

where $U(b)$ is an arbitrary function. Our solution will be the solution (1.27) plus the solution (1.28) with the arbitrary function $U(b)$ chosen to satisfy (1.26), which implies $U'(b) = 2\omega k^2A^2e^{2kb}$. Thus we choose $U(b) = \omega kA^2e^{2kb} = cA^2k^2e^{2kb}$. Putting all this together, we have

$$x = a + Ae^{kb}\sin(ka - \omega\tau) + cA^2k^2e^{2kb}\tau + O(A^3) \tag{1.29}$$

$$y = b - Ae^{kb}\cos(ka - \omega\tau) + \frac{1}{2}A^2ke^{2kb} + O(A^3) \tag{1.30}$$

$$p = -gb + O(A^3) \tag{1.31}$$

where $c = \omega/k$ and $\omega^2 = gk$. Again, A and k are constants.

The solution (1.29)-(1.31) is the complete water-wave solution, correct to $O(A^2)$. The oscillatory components are familiar, as is the Stokes drift term, the last term shown in (1.29). The last term in (1.30) is less obvious. And where, you may ask, are the Stokes corrections to the basic wave?

The Stokes corrections are there, but they are hidden in the transformation back to Eulerian coordinates. Suppose, for example, we want the standard representation, $y = y(x,t)$, for the free surface elevation. The free surface corresponds to $b = 0$. At $b = 0$ we have

$$x = a + A\sin(ka - \omega\tau) + cA^2k^2\tau \tag{1.32}$$

$$y = -A\cos(ka - \omega\tau) + \frac{1}{2}A^2k \tag{1.33}$$

To obtain $y(x,t)$ we must eliminate a between (1.32) and (1.33). To get the waveform we may set τ=0. Then (1.32)-(1.33) become

$$x = a + A\sin(ka) \tag{1.34}$$

$$y = -A\cos(ka) + \frac{1}{2}A^2k \tag{1.35}$$

By (1.34) we have

$$a = x - A\sin(kx) + O(A^2) \tag{1.36}$$

Substituting (1.36) into (1.35) we obtain

$$\begin{aligned} y(x) &= -A\cos\left(kx - Ak\sin(kx)\right) + \frac{1}{2}A^2k \\ &\approx -A\cos(kx) + A\sin(kx)\left(-Ak\sin(kx)\right) + \frac{1}{2}A^2k \\ &= -A\cos(kx) - A^2k\sin^2(kx) + \frac{1}{2}A^2k \\ &= -A\cos(kx) + \frac{1}{2}A^2k\cos(2kx) \end{aligned} \tag{1.37}$$

which is the familiar, leading-order, Stokes correction to the wave form. To get the leading-order correction to the phase speed, it is necessary to compute the superscript-(3) terms in (1.13)-(1.15). We will not do this.

Up to now we have considered A to be a constant. Now we consider a wave packet in which k is constant but $A(a,\tau)$ varies slowly, on a time scale much larger than the wave period, and on a horizontal length scale much larger than the wavelength. We replace the constant A in (1.19)-(1.21) by the

slowly varying function $A(a,t)$. That is, we replace (1.19)-(1.21) by the linear superposition of basic waves that add up to make the wave packet. In the next order equations, (1.22)-(1.24) and (1.26), we replace A by $A(a,\tau)$. This replacement is legal, because the derivatives performed on the superscript-(1) solution to obtain (1.22)-(1.24) and (1.26) were derivatives with respect to the 'fast' variations of wave phase. The derivatives of $A(a,\tau)$ are negligible in comparison to these.

Since we are considering a wave packet, we can be even more specific. We know that the wave packet moves at the group velocity corresponding to the carrier wavenumber k. Thus assuming that $k > 0$ (wave propagating to the right) we take $A(a,\tau) = A(a - c_g\tau)$, where $A()$ is now an arbitrary but slowly varying function of its single argument, and c_g is given by (1.2). The superscript-(2) equations, (1.22)-(1.24) and (1.26), become

$$x^{(2)}_{\tau\tau} = -gy^{(2)}_a - p^{(2)}_a \tag{1.38}$$

$$y^{(2)}_{\tau\tau} = +gx^{(2)}_a - p^{(2)}_b \tag{1.39}$$

$$x^{(2)}_a + y^{(2)}_b = k^2 A(a - c_g\tau)^2 e^{2kb} \tag{1.40}$$

$$y^{(2)}_{\tau a} - x^{(2)}_{\tau b} = -2\omega k^2 A(a - c_g\tau)^2 e^{2kb} \tag{1.41}$$

Let $u^{(2)} = x^{(2)}_\tau$ and $v^{(2)} = y^{(2)}_\tau$. Then the time-derivative of (1.40) is

$$u^{(2)}_a + v^{(2)}_b = \frac{\partial}{\partial\tau}\left(k^2 A(a - c_g\tau)^2 e^{2kb}\right) \tag{1.42}$$

and eqn (1.41) becomes

$$v^{(2)}_a - u^{(2)}_b = -2\omega k^2 A(a - c_g\tau)^2 e^{2kb} \tag{1.43}$$

The right-hand side of (1.42) is much smaller than the right-hand side of (1.43), because $|A_\tau| << \omega A$. Thus the divergence of the mean velocity field is much smaller than its curl, and we may set

$$u^{(2)} = -\psi_b, \quad v^{(2)} = \psi_a \tag{1.44}$$

Eqn (1.43) becomes a Poisson equation for the mean streamfunction,

$$\psi_{aa} + \psi_{bb} = -2\omega k^2 A(a - c_g\tau)^2 e^{2kb} \tag{1.45}$$

The mean flow sees the free surface as a rigid lid, hence $\psi = 0$ at $b = 0$. This boundary condition is intuitively obvious, but it will be justified more formally below.

The solution of (1.45) is easiest if we consider the problem in the reference frame moving with the wave packet. This amounts to a Galilean transformation in label space. In the comoving frame the A^2-term in (1.45) becomes a steady source. Introducing the new independent variables

$$\alpha = a - c_g \tau \tag{1.46}$$

$$\beta = b \tag{1.47}$$

$$s = \tau \tag{1.48}$$

and applying the transformation equations

$$\frac{\partial}{\partial \tau} = \frac{\partial}{\partial s} - c_g \frac{\partial}{\partial \alpha} \tag{1.49}$$

$$\frac{\partial}{\partial a} = \frac{\partial}{\partial \alpha} \tag{1.50}$$

$$\frac{\partial}{\partial b} = \frac{\partial}{\partial \beta} \tag{1.51}$$

we find that (1.45) becomes

$$\psi_{\alpha\alpha} + \psi_{\beta\beta} = -2\omega k^2 A(\alpha)^2 e^{2k\beta} \equiv S(\alpha, \beta) \tag{1.52}$$

To justify the boundary condition $\psi = 0$, we take the time derivative of (1.38) at $b = 0$ (where $p^{(2)} = 0$). The result is

$$u^{(2)}_{\tau\tau} = -g v^{(2)}_a \qquad \text{at } b = 0 \tag{1.53}$$

In the comoving reference frame this becomes

$$c_g^2 u^{(2)}_{\alpha\alpha} = -g v^{(2)}_\alpha \qquad \text{at } b = 0 \tag{1.54}$$

Since the mean velocity vanishes very far from the wave packet, (1.54) implies

$$u^{(2)}_\alpha = -4k v^{(2)} \qquad \text{at } b = 0 \tag{1.55}$$

and hence $v^{(2)} \approx 0$ at $b = 0$ because $u^{(2)}$ varies on a scale much larger than k^{-1}.

Our solution for the mean flow generated by a wave packet boils down to the Poisson equation (1.52) with boundary condition $\psi = 0$. Although it is difficult to write down the analytical solution, the general character of

the solution is obvious from (1.52), which equates the vorticity of the mean flow to the source term $S(\alpha, \beta)$. This source term is negative because c_g is positive for a right-moving packet. The source term is also compact; it vanishes outside the wave packet, which extends only to a depth of order k^{-1}. Thus, stepping back, we see that the mean flow must be as depicted in figure 1.3: a broad anticyclonic flow around the wave packet which appears as a negative 'point-source' for mean vorticity. The 'point source' is non-vanishing only in a small region at, and just below, the free surface at $b = 0$. The rightward mean flow is confined to a very small depth (the depth of the wave packet itself) near $b = 0$. We recognize this shallow rightward mean flow as Stokes drift, but our calculation shows that there is more to the mean flow than just the Stokes drift.

Please note that there is considerable irony in our analysis of this flow in terms of vorticity. The vorticity of the fluid is precisely *zero*. Our mean flow does in fact have a non-vanishing vorticity, but its vorticity is cancelled by the vorticity contained in the wave packet itself. Thus we have solved a relatively complicated problem of irrotational flow by analyzing the vorticity equation for the flow.

But there is an additional benefit to our separation of the vorticity into mean and fluctuating parts. The benefit is this: The fluctuating vorticity has a very precarious existence. At any moment, wave-breaking (followed by turbulence and viscous dissipation) can destroy some or all of it. In contrast, the mean vorticity is much less affected by viscosity.

The effect of wave breaking is most easily understood by thinking of the fluid evolution as a movie. The breaking occurs in a single frame. In the frames before and after the breaking frame, the dynamics analyzed by us apply.

Suppose that the breaking removes half the wave energy. This means that the wave packet amplitude is reduced by a factor $1/\sqrt{2}$. The wavenumber k is unaffected. In the frames following the breaking frame, the wave packet continues to move at its same group velocity, but with only half its previous energy. But since (1.52) must apply to all the post-breaking frames, the mean flow dragged along by the wave packet can have only half the amplitude that it had before breaking. What happens to the additional mean-flow vorticity? It cannot have been dissipated by viscosity. The mean flow is too broad to be dissipated in a single movie frame. The answer is that this excess mean vorticity is simply jettisoned, left behind at the point of wave breaking. This explains the observations summarized at the beginning of this chapter. The

zero vorticity of the initial wave packet has been converted by jettisoning, into real, actual vorticity by the action of wave breaking.

Even if the wave packet doesn't break, it can cause permanent changes to the flow. Suppose you are a fluid particle or a neutrally buoyant glass bead very near the free surface. Before the wave packet arrives, you are at rest. As the wave packet passes, you move to the right in the Stokes drift of the mean flow. After the wave packet passes, you are at rest again, but at a location that is to the right of where you started. You have experienced a change in location. If you are a molecule of passive tracer, you have been transported by the wave packet. If there is a succession of wave packets, you just keep moving. The limiting case is a infinite sinusoidal wave in which you move steadily at the Stokes velocity.

Of course, if the wave breaks right on top of you, you acquire a semi-permanent velocity as a part of the jettisoned mean vorticity field. But let's continue to consider a non-breaking wave packet. Suppose, instead of a glass bead, you are a vorticity line or a part of a vortex tube that existed in the fluid before the arrival of the wave packet. As the wave packet passes you move with the attached mean flow, and then you stop. Thus, after passage of the wave packet, any already-existing field of vorticity has been re-arranged. There is a permanent change in the pre-existing flow even though the wave packet didn't break. This is the 'ratcheting effect' referred to in many of McIntyre's papers.

The analysis presented here falls within the domain of 'wave-mean' theory, the subject of the beautiful book by Oliver Bühler. This subject is made difficult by the fact that various investigators have approached it from quite dissimilar directions. Even if you learn it in one particular way, you may have trouble following someone else's application. Following an alternative approach, the right-hand side of (1.52) emerges as the curl of the pseudomomentum. The pseudomomentum is defined as

$$P = k\mathcal{A} = \frac{kE}{\omega} \tag{1.56}$$

where $\mathcal{A} = E/\omega$ is the wave action, and

$$E = gkA^2e^{2kb} \tag{1.57}$$

is the energy density of the wave packet. Thus

$$P = \omega kA^2e^{2kb} \tag{1.58}$$

and the curl of the pseudomomentum is

$$-\frac{\partial P}{\partial b} = -2\omega k^2 A^2 e^{2kb} \tag{1.59}$$

as claimed. This is a general result, true for all types of waves. We will meet it again, many times, in the following chapters.

Chapter 2

Mean flow generated by an internal wave packet

In this second chapter we compute the 'Bretherton flow' associated with an internal wave packet propagating through a stratified fluid. As in the previous chapter, we assume that the fluid was initially at rest. Now, however, we consider the full three-dimensional character of the flow. (Although surface gravity waves are often long-crested, internal waves frequently are not.) We do not consider coordinate system rotation, but that could easily be added. Also, we ignore top and bottom boundaries; our wave packet moves at its group velocity through a stratified fluid of infinite extent. However, the analysis of a vertical mode, with packet geometry in the horizontal, would proceed along similar lines.

We take the governing dynamics to be Boussinesq dynamics. In standard Eulerian form the Boussinesq equations are

$$\frac{Du}{Dt} = -p_x \tag{2.1}$$

$$\frac{Dv}{Dt} = -p_y \tag{2.2}$$

$$\frac{Dw}{Dt} = -p_z + \theta \tag{2.3}$$

$$u_x + v_y + w_z = 0 \tag{2.4}$$

$$\frac{D\theta}{Dt} = 0 \tag{2.5}$$

with θ the buoyancy, $D/Dt = \partial_t + u\partial_x + v\partial_y + w\partial_z$, and z positive up. In

Lagrangian coordinates these equations become

$$x_{\tau\tau} = -\frac{\partial(p,y,z)}{\partial(a,b,c)} \tag{2.6}$$

$$y_{\tau\tau} = -\frac{\partial(x,p,z)}{\partial(a,b,c)} \tag{2.7}$$

$$z_{\tau\tau} = -\frac{\partial(x,y,p)}{\partial(a,b,c)} + \theta(c) \tag{2.8}$$

$$\frac{\partial(x,y,z)}{\partial(a,b,c)} = 1 \tag{2.9}$$

We satisfy (2.5) by taking θ to be a prescribed function of the label c. The state of rest corresponds to

$$x(a,b,c,\tau) = a \tag{2.10}$$

$$y(a,b,c,\tau) = b \tag{2.11}$$

$$z(a,b,c,\tau) = c \tag{2.12}$$

$$p(a,b,c,\tau) = \int_0^c \theta(c')dc' \equiv p^{(0)}(c) \tag{2.13}$$

where the zero level is arbitrary. We solve (2.6)-(2.9) by a perturbation expansion about the state of rest, in which the waves are assumed to be weak, and the induced mean flow is assumed to be even weaker. Thus

$$x(a,b,c,\tau) = a + x^{(1)}(a,b,c,\tau) + x^{(2)}(a,b,c,\tau) + \cdots \tag{2.14}$$

$$y(a,b,c,\tau) = b + y^{(1)}(a,b,c,\tau) + y^{(2)}(a,b,c,\tau) + \cdots \tag{2.15}$$

$$z(a,b,c,\tau) = c + z^{(1)}(a,b,c,\tau) + z^{(2)}(a,b,c,\tau) + \cdots \tag{2.16}$$

$$p(a,b,c,\tau) = p^{(0)}(c) + p^{(1)}(a,b,c,\tau) + p^{(2)}(a,b,c,\tau) + \cdots \tag{2.17}$$

Before the wave packet arrives, the fluid is at rest. We introduce a constraint that enforces this. The appropriate constraint is that the potential vorticity has the value it had in the rest state. This is analogous to the vanishing of ordinary vorticity in the previous chapter. The Boussinesq equations conserve potential vorticity in the form

$$\frac{D}{Dt}(\boldsymbol{\omega} \cdot \nabla\theta) = 0 \tag{2.18}$$

where $\mathbf{u} = (u, v, w)$, and $\boldsymbol{\omega} = \nabla \times \mathbf{u}$ is the vorticity. Thus the appropriate constraint is

$$\boldsymbol{\omega} \cdot \nabla\theta = 0 \tag{2.19}$$

To translate (2.19) into Lagrangian coordinates, we note that

$$\boldsymbol{\omega} \cdot \nabla\theta = \epsilon_{ijk} \frac{\partial\theta}{\partial x_i} \frac{\partial u_r}{\partial x_j} \frac{\partial x_r}{\partial x_k} = \frac{\partial(u_r, x_r, \theta)}{\partial(x, y, z)} = \frac{\partial(u_r, x_r, \theta)}{\partial(a, b, c)} \tag{2.20}$$

where repeated indices are summed. Thus the appropriate form of the constraint is

$$\frac{\partial(x_\tau, x, \theta)}{\partial(a, b, c)} + \frac{\partial(y_\tau, y, \theta)}{\partial(a, b, c)} + \frac{\partial(z_\tau, z, \theta)}{\partial(a, b, c)} = 0 \tag{2.21}$$

However, since $\theta = \theta(c)$ and we assume $\theta'(c) \neq 0$, (2.21) takes the form

$$\frac{\partial(x_\tau, x)}{\partial(a, b)} + \frac{\partial(y_\tau, y)}{\partial(a, b)} + \frac{\partial(z_\tau, z)}{\partial(a, b)} = 0 \tag{2.22}$$

Compare (2.22) to (1.12). Our task is to solve (2.6)-(2.9) with the constraint (2.22). (Recognizing that the τ-derivative of (2.22) follows from (2.6)-(2.9), we emphasize that (2.22) is a constraint on the initial conditions.) At O(A) (2.6)-(2.9) become

$$x^{(1)}_{\tau\tau} = -p^{(1)}_a + \theta(c) z^{(1)}_a \tag{2.23}$$

$$y^{(1)}_{\tau\tau} = -p^{(1)}_b + \theta(c) z^{(1)}_b \tag{2.24}$$

$$z^{(1)}_{\tau\tau} = -p^{(1)}_c - \theta(c)(x^{(1)}_a + y^{(1)}_b) \tag{2.25}$$

$$x^{(1)}_a + y^{(1)}_b + z^{(1)}_c = 0 \tag{2.26}$$

These equations take a more familiar form if we define

$$\hat{p}^{(1)} \equiv p^{(1)} - \theta(c) z^{(1)} \tag{2.27}$$

Then (2.23)-(2.26) become

$$x^{(1)}_{\tau\tau} = -\hat{p}^{(1)}_a \tag{2.28}$$

$$y^{(1)}_{\tau\tau} = -\hat{p}^{(1)}_b \tag{2.29}$$

$$z^{(1)}_{\tau\tau} = -\hat{p}^{(1)}_c - N^2(c) z^{(1)} \tag{2.30}$$

$$x^{(1)}_a + y^{(1)}_b + z^{(1)}_c = 0 \tag{2.31}$$

where $N^2(c) \equiv \theta'(c)$. For simplicity we assume $N^2(c) = N_0^2$ (constant), corresponding to a uniform stratification in the state of rest. The solution of interest is

$$x^{(1)} = -\frac{k}{\omega^2}\, A(a,b,c,\tau)\sin\phi \tag{2.32}$$

$$y^{(1)} = -\frac{l}{\omega^2}\, A(a,b,c,\tau)\sin\phi \tag{2.33}$$

$$z^{(1)} = \frac{(k^2+l^2)}{\omega^2 m}\, A(a,b,c,\tau)\sin\phi \tag{2.34}$$

$$\hat{p}^{(1)} = A(a,b,c,\tau)\cos\phi \tag{2.35}$$

where

$$\phi = ka + lb + mc - \omega\tau \tag{2.36}$$

and

$$\omega^2 = \frac{N_0^2(k^2+l^2)}{(k^2+l^2+m^2)} \tag{2.37}$$

In this solution, k, l, m, ω are constants and $A(a,b,c,\tau)$ varies slowly compared to ϕ. We will also need an expression for the energy of the wave packet. The kinetic energy, averaged over a wave cycle, is

$$\frac{1}{2}\langle (x_\tau^{(1)})^2 + (y_\tau^{(1)})^2 + (z_\tau^{(1)})^2\rangle = \left(k^2+l^2+\frac{(k^2+l^2)^2}{m^2}\right)\frac{A^2}{4\omega^2} = \frac{K^4A^2}{4N_0^2m^2} \tag{2.38}$$

where $\mathbf{k} \equiv (k,l,m)$ and $K = |\mathbf{k}|$. Recalling that the kinetic energy equals the potential energy in internal waves, we find the energy density of the wave packet to be

$$E(a,b,c,\tau) = \frac{K^4}{2N_0^2m^2}A(a,b,c,\tau)^2 \tag{2.39}$$

It will also be handy to have the group velocity,

$$\mathbf{c}_g = \left(\frac{\partial\omega}{\partial k}, \frac{\partial\omega}{\partial l}, \frac{\partial\omega}{\partial m}\right) = \frac{N_0^2 m}{\omega K^4}\left(km, lm, -k^2-l^2\right) \tag{2.40}$$

Note that $\mathbf{k}\cdot\mathbf{c}_g = 0$ as expected for internal waves.

We note that solutions of the form (2.32)-(2.35) are not the only solutions to (2.28)-(2.31). In addition, we may have steady, two-dimensional, hydrostatic flows in horizontal planes. These horizontal flows come into play as we proceed.

At the next order, we obtain

$$\begin{aligned} x^{(2)}_{\tau\tau} &= -p^{(2)}_a + \theta(c) z^{(2)}_a - \frac{\partial(p^{(1)}, y^{(1)})}{\partial(a,b)} - \frac{\partial(p^{(1)}, z^{(1)})}{\partial(a,c)} \\ &= -\hat{p}^{(2)}_a + N_0^2 z^{(1)} z^{(1)}_a - \frac{\partial(\hat{p}^{(1)}, y^{(1)})}{\partial(a,b)} - \frac{\partial(\hat{p}^{(1)}, z^{(1)})}{\partial(a,c)} \end{aligned} \tag{2.41}$$

$$\begin{aligned} y^{(2)}_{\tau\tau} &= -p^{(2)}_b + \theta(c) z^{(2)}_b - \frac{\partial(p^{(1)}, z^{(1)})}{\partial(b,c)} - \frac{\partial(x^{(1)}, p^{(1)})}{\partial(a,b)} \\ &= -\hat{p}^{(2)}_b + N_0^2 z^{(1)} z^{(1)}_b - \frac{\partial(\hat{p}^{(1)}, z^{(1)})}{\partial(b,c)} - \frac{\partial(x^{(1)}, \hat{p}^{(1)})}{\partial(a,b)} \end{aligned} \tag{2.42}$$

$$\begin{aligned} z^{(2)}_{\tau\tau} &= -p^{(2)}_c - \theta(c)\left(x^{(2)}_a + y^{(2)}_b\right) - \frac{\partial(x^{(1)}, p^{(1)})}{\partial(a,c)} - \frac{\partial(y^{(1)}, p^{(1)})}{\partial(b,c)} \\ &= -\hat{p}^{(2)}_c - N_0^2 z^{(2)} + N_0^2 z^{(1)} z^{(1)}_c - \frac{\partial(x^{(1)}, \hat{p}^{(1)})}{\partial(a,c)} - \frac{\partial(y^{(1)}, \hat{p}^{(1)})}{\partial(b,c)} \end{aligned} \tag{2.43}$$

$$x^{(2)}_a + y^{(2)}_b + z^{(2)}_c = 0 \tag{2.44}$$

where

$$\hat{p}^{(2)} \equiv p^{(2)} - \theta(c) z^{(2)} \tag{2.45}$$

Note that the symbol $\hat{p}$ replaces p in the second lines of (2.41)-(2.43). Note also the correspondence between (2.45) and (2.27). In writing (2.41)-(2.43) we have used

$$\frac{\partial(x^{(1)}, y^{(1)})}{\partial(a,b)} = \frac{\partial(y^{(1)}, z^{(1)})}{\partial(a,c)} = 0 \quad \text{etc.} \tag{2.46}$$

because $x^{(1)}, y^{(1)}$ and $z^{(1)}$ differ only by a constant factor.

Compared to the second-order surface wave equations, these equations are a mess. And there is worse to come. If you attempt to evaluate the (1)-(1) terms in (2.41)-(2.43) by treating $A(a,b,c,\tau)$ as a constant, which is the approach followed in the previous chapter, then you find that all of the Jacobian terms in (2.41)-(2.43) actually vanish. Good riddance, you think. Things get simpler. Unfortunately, however, all the remaining (1)-(1) terms are negligible for a different reason. Take $z^{(1)} z^{(1)}_a$ for example. Since $z^{(1)} \propto \sin\phi$ and $z^{(1)}_a \propto \cos\phi$, their product is proportional to $\sin 2\phi$. To

the broad induced mean flow, this looks like a compact, rapidly fluctuating, source with an average value of zero. Squint your eyes and it disappears.

What has gone wrong? The answer is this: The (1)-(1) terms in (2.41)-(2.43) will contribute, but only if we take the slow variations of $A(a,b,c,\tau)$ into account. This is an enormous difference from the surface wave problem, best understood as follows. In the surface wave problem, the wave packet has a monopole component; the average of the source term does not vanish. The internal wave packet has no monopole component; averaging the source terms over the internal wave packet just gives zero. The mean flow responds, but its response will be a dipole. We must work harder.

Now take the term

$$\frac{\partial(x^{(1)},\hat{p}^{(1)})}{\partial(a,b)} \equiv [x^{(1)},\hat{p}^{(1)}] \propto [A\sin\phi, A\cos\phi] \tag{2.47}$$

Since

$$[A,A] = [\sin\phi,\cos\phi] = 0 \tag{2.48}$$

we have

$$\begin{aligned}[A\sin\phi, A\cos\phi] &= A[\sin\phi, A]\cos\phi + \sin\phi[A,\cos\phi]A \\ &= A\cos\phi[\phi,A]\cos\phi - \sin\phi[A,\phi]A\sin\phi \\ &= A[\phi,A] = A(kA_b - lA_a)\end{aligned} \tag{2.49}$$

Thus,

$$\frac{\partial(x^{(1)},\hat{p}^{(1)})}{\partial(a,b)} = -\frac{k}{\omega^2}(kA_b - lA_a)A \tag{2.50}$$

The other Jacobians can be evaluated in a manner analogous to (2.50). Then (2.41)-(2.44) become

$$x^{(2)}_{\tau\tau} = -\hat{p}^{(2)}_a - \frac{k}{2\omega^2 m}\left(km\partial_a + lm\partial_b - (k^2+l^2)\partial_c\right)A^2 \tag{2.51}$$

$$y^{(2)}_{\tau\tau} = -\hat{p}^{(2)}_b - \frac{l}{2\omega^2 m}\left(km\partial_a + lm\partial_b - (k^2+l^2)\partial_c\right)A^2 \tag{2.52}$$

$$z^{(2)}_{\tau\tau} = -\hat{p}^{(2)}_c - N_0^2 z^{(2)} - \frac{m}{2\omega^2 m}\left(km\partial_a + lm\partial_b - (k^2+l^2)\partial_c\right)A^2 \tag{2.53}$$

$$x^{(2)}_a + y^{(2)}_b + z^{(2)}_c = 0 \tag{2.54}$$

We have dropped terms such as $z^{(1)}z_a^{(1)}$ that phase-average to zero. Using (2.40), the momentum equations (2.51)-(2.53) can be written more succinctly as

$$\mathbf{u}_\tau^{(2)} = -\nabla\hat{p}^{(2)} - N_0^2 z^{(2)}\nabla c - \frac{\mathbf{k}K^4}{2m^2\omega N_0^2}\,\mathbf{c}_g\cdot\nabla A^2 \tag{2.55}$$

where $\mathbf{u}^{(2)} = (u^{(2)}, v^{(2)}, w^{(2)}) = (x_\tau^{(2)}, y_\tau^{(2)}, z_\tau^{(2)})$ and $\nabla = (\partial_a, \partial_b, \partial_c)$ is the gradient operator in label space. Note that ∇c is the vertical unit vector in label space. Similarly, the τ-derivative of (2.54) can be written

$$\nabla\cdot\mathbf{u}^{(2)} = 0 \tag{2.56}$$

The potential vorticity constraint is handled in a similar manner. To the first two non-vanishing orders, (2.22) becomes

$$y_{a\tau}^{(1)} - x_{b\tau}^{(1)} + y_{a\tau}^{(2)} - x_{b\tau}^{(2)} + \frac{\partial(x_\tau^{(1)}, x^{(1)})}{\partial(a,b)} + \frac{\partial(y_\tau^{(1)}, y^{(1)})}{\partial(a,b)} + \frac{\partial(z_\tau^{(1)}, z^{(1)})}{\partial(a,b)} = 0 \tag{2.57}$$

The first two terms vanish, and the Jacobian terms are handled in the same manner as above. We obtain

$$v_a^{(2)} - u_b^{(2)} = \frac{1}{\omega}(l\partial_a - k\partial_b)E(a,b,c,\tau) \tag{2.58}$$

after use of (2.39).

Our goal is to find the mean velocity field $\mathbf{u}^{(2)}$ induced by the wave packet. The left-hand side of (2.58) is the vertical component of its curl. The right-hand side of (2.58) is also the vertical component of a curl—the curl of the vector

$$\mathbf{P} \equiv \frac{\mathbf{k}}{\omega}E \tag{2.59}$$

Thus (2.58) can be rewritten as

$$\nabla c\cdot\left[\nabla\times\left(\mathbf{u}^{(2)} - \mathbf{P}\right)\right] = 0 \tag{2.60}$$

where $\mathbf{P}$ is the pseudomomentum of the wave packet.

Now, (2.60) is just a single equation, and we need to determine all *three* of the components of $\mathbf{u}^{(2)}$. It is obvious, therefore, that (2.60) cannot be enough. Surprisingly, however, it *is* enough. In a sense to be carefully qualified, (2.60) tells us everything that we really need to know about $\mathbf{u}^{(2)}$. But what does it actually tell us?

The horizontal components of $\mathbf{u}^{(2)}$ may be expressed quite generally as

$$u^{(2)} = -\psi_b + \lambda_a, \qquad v^{(2)} = \psi_a + \lambda_b \tag{2.61}$$

where $\psi(a, b, c, \tau)$ is the stream function and $\lambda(a, b, c, \tau)$ is the velocity potential in label space. Substituting (2.61) into (2.60) we obtain

$$\psi_{aa} + \psi_{bb} = \nabla c \cdot [\nabla \times \mathbf{P}] \tag{2.62}$$

Since the pseudomomentum is prescribed, this gives us ψ. If we could use another equation to determine λ, then we would know $u^{(2)}$ and $v^{(2)}$; and (2.56) in the form

$$\lambda_{aa} + \lambda_{bb} + w_c^{(2)} = 0 \tag{2.63}$$

would give us $w^{(2)}$. It appears that our work is far from done. But, in a sense to be carefully qualified, λ and hence $w^{(2)}$ turn out to be negligible. The simple equation (2.62) tells the whole story. Before explaining the qualifications, let us see what it says.

With no loss in generality, we may assume $l = 0$. This corresponds to rotating the horizontal axes until the a-axis points in the direction of phase propagation. Then (2.62) takes the form

$$\psi_{aa} + \psi_{bb} = -P_b \tag{2.64}$$

where P is the magnitude of $\mathbf{P}$. This is the same equation obtained in the previous chapter! The only differences are: (1) In this chapter a and b are horizontal coordinates, while in the previous chapter b was the vertical coordinate; and (2) In the previous chapter P decayed with depth in such a way that P_b was always positive, while here P_b takes both signs. The monopole has become a dipole. Figure 2.1 clarifies. If $A(a, b, c)$, the amplitude of the wave packet, is a single hump, then $P_b > 0$ (corresponding to a negative source in (2.64)) when $b < 0$; and $P_b < 0$ (corresponding to a positive source in (2.64)) when $b > 0$. The induced mean flow corresponding to ψ is a broad dipole, represented by the closed directional curves in figure 2.1.

This flow occurs at each level c as the wave packet passes the level. Before and after passage, the mean flow vanishes *unless* the internal waves inside the packet decide to break. If breakage occurs, then some of the mean flow is jettisoned as a semi-permanent dipole at the level at which breakage occurred. The remarks made in the previous chapter about re-arrangements of any pre-existing vorticity field apply here as well. In fact, virtually every remark we

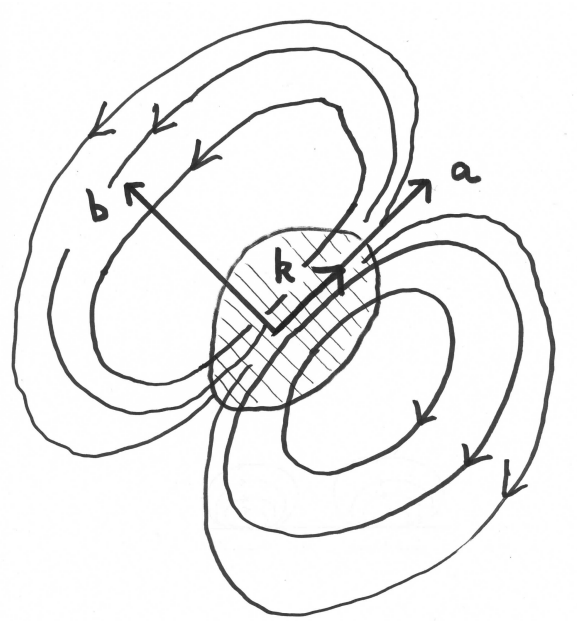

Figure 2.1: Wave packet and induced dipole on a surface of constant label c. The wave vector is $(k, 0, m)$, and the straight parallel lines are lines of constant phase (wave crests) within the packet. The packet moves upward, in the direction of increasing c, if $m < 0$, and downward if $m > 0$. The dipolar mean flow is present at level c only while the wave packet is present there *unless* the waves inside the packet break, losing some of their energy. If breakage occurs, some of the dipolar mean flow gets left behind at the level at which breakage occurred.

made about the surface wave packet is true here as well. Even the biggest difference between the two—monopole versus dipole—disappears if, in the previous chapter, the free surface boundary condition on the mean flow is replaced by a counter-rotating image vortex above the free surface. The two problems really seem to be the same problem!

But what about our assumption that the mean flow surrounding the internal wave packet takes the approximate form

$$u^{(2)} = -\psi_b, \qquad v^{(2)} = \psi_a, \qquad w^{(2)} = 0 \tag{2.65}$$

What justifies *that*? It turns out that this is not so much a matter of justification as it is a matter of *choice*. A skeptic would say that you have to solve the complete, second-order dynamical equations (2.55) and (2.56) to see everything that happens. Using (2.39) we rewrite these equations in the form

$$\bar{\mathbf{u}}^{(2)}_\tau = -\nabla \hat{p}^{(2)} - N_0^2 z^{(2)} \nabla c - \frac{\mathbf{k}}{\omega} \mathbf{c}_g \cdot \nabla E(a,b,c,\tau) \tag{2.66}$$

$$\nabla \cdot \mathbf{u}^{(2)} = 0 \tag{2.67}$$

Comparing (2.66)-(2.67) to (2.28)-(2.31), we see that they are the same except for the forcing terms contributed by the wave packet in (2.66). We are interested only in the mean flow response to these forcing terms. (Equations (2.66)-(2.67) also have homogeneous free-wave solutions analogous to (2.32)-(2.35), but it would be silly to consider them unless they were needed to satisfy boundary conditions.) But, again, what justifies our assertion that that this mean flow response takes the form (2.65) ?

In a sense, nothing at all. The free solutions of the mean flow equations comprise both non-zero-frequency modes satisfying the dispersion relation (2.37) and zero-frequency modes of the form (2.65). In general both are excited. The general procedure for determining the forced response has been beautifully explained by Lighthill. The modes that are most strongly excited by the wave packet are those whose space and time scales come closest to matching those of $A(a,b,c,\tau)$. But the scales of $A(a,b,c,\tau)$ are very slow. The wave packet, it is true, propagates at the group velocity corresponding to its wavenumber $\mathbf{k}$, and this is relatively fast. But the wave packet is broad; it contains many wavelengths; its envelope $A(a,b,c,\tau)$ therefore changes very slowly. It is this slow change that the mean flow sees as its source. Slow forcing begets slow response, and there is nothing slower than the zero-frequency

modes of the form (2.65). In asserting that (2.64) tells the whole story of the response, we are simply neglecting the response of the fast modes to the slow source.

Is this really okay? It depends on what you want to know. Intuition suggests that a moving wave packet might excite other waves. It might develop a wake. If the details of this wake are of interest to you, then it would be a terrible mistake to take (2.64) and (2.65) to be the solution. However, if you only want to know about the biggest part of the response, then you can embrace (2.64) and be thankful.

You may recall a method for deriving the quasigeostrophic equations which goes like this: Choose a particular state of rest. Compute the modes corresponding to a linearization about this state of rest. They comprise geostrophic modes and gravity modes. Returning to the fully nonlinear dynamics, expand the variables in these linear modes. Since the linear modes form a complete set, this involves no approximation. Now the approximation: Set all the gravity mode amplitudes to zero. The resulting equations describe geostrophic modes interacting with other geostrophic modes. The full dynamics has been *projected* onto the geostrophic modes. And another name for 'geostrophic mode' is 'potential vorticity.'

Our key equation (2.62) has a similar interpretation. It tells us how the gravity wave packet projects onto the zero-frequency mode. If we had worked this problem in rotating coordinates, we would have obtained, instead of (2.60),

$$f + \nabla c \cdot [\nabla \times (\bar{\mathbf{u}} - \mathbf{P})] = 0 \tag{2.68}$$

where f is the Coriolis parameter, and we would be saying that (2.68) describes the excitation of the *geostrophic mode* by a gravity wave packet.

In contrast to the surface wave case, the packet geometry is essential for the excitation of mean vorticity in the internal wave case. You may recall that a single, perfectly sinusoidal, internal gravity wave is an *exact* solution of the Boussinesq equations. A perfect sinusoid cannot therefore interact with itself to produce a mean flow. Our analysis is in agreement with this. The perfectly sinusoidal wave corresponds to A absolutely constant, and, therefore to the vanishing of the pseudomomentum curl in (2.68) and the forcing terms in (2.66). To get a response, we must consider a wave packet, rounding A at its edges, so to speak, giving the sinusoid some bandwith. In contrast, the surface wave comes automatically with 'rounded edges', because it has no choice but to decay with depth.

In hindsight, our calculation is not so much an expansion in powers of A as it is a *separation* of the flow into rapidly and slowly varying parts. This is something to keep in mind as we extend our methods to situations in which the waves and the mean flow interact in more complicated ways.

Although we have obtained the fundamental result (2.60) by enforcing the condition of vanishing potential vorticity, (2.60) could also be obtained by taking the vertical component of the curl of (2.66). One then uses

$$E_\tau + \mathbf{c}_g \cdot \nabla E = 0 \tag{2.69}$$

to obtain the τ-derivative of (2.60).[1] Since the potential vorticity vanishes before the arrival of the wave packet, (2.60) follows. However, this method of getting (2.60) requires more work and more ingenuity than our approach of concentrating on the potential vorticity by itself. But how would one know to do this? And what is so special about pseudomomentum, that seemingly strange visitor from another planet? The next chapter begins to address these questions.

[1] Eqn (2.69) applies because $\mathbf{k}$ and hence $\mathbf{c}_g$ are constant in our case. The general equation is $E_\tau + \nabla \cdot (\mathbf{c}_g E) = 0$.

Chapter 3

The variational approach

I have a friend who once said: "Hamiltonian philosophy is like avocado. You either like it or you don't."[1] In this chapter I will try to make you like it.

Variational principles offer two great advantages. First, they are wonderfully compact statements of the physics. Second, their symmetry properties correspond to conservation laws. The conservation of energy corresponds to the symmetry property that the Lagrangian is invariant to a translation in time. The conservation of momentum corresponds to the symmetry property that the Lagrangian is invariant to a translation in space. These two symmetry properties are what physicists call 'kinematical symmetries'; they apply to virtually every physical system. But fluid Lagrangians have an additional, 'particle relabeling' symmetry that corresponds to the conservation of potential vorticity. This symmetry property is what makes a fluid a fluid, in contrast (for example) to an elastic solid. In physical terms, the internal energy of the fluid cares only about how much the fluid has been squeezed and not, for example, about how much it has been twisted or stretched. The squeezing is measured by

$$\frac{\partial(x,y,z)}{\partial(a,b,c)} \tag{3.1}$$

and is not affected by a relabeling of the fluid particles that does not affect this Jacobian, just as you wouldn't live any differently if someone changed your social security number. The relabeling of fluid particles is closely analogous to the gauge transformation in electrodynamics. In the language of field theory, both are *local gauge symmetries*, a concept that underlies the

[1] Phil Morrison at Walsh Cottage, June, 1993.

four fundamental forces at work in the cosmos. This discussion can be made very broad, but we come quickly back to Earth.

The Lagrangian for Boussinesq dynamics is

$$L[x(a,b,c,\tau), y(a,b,c,\tau), z(a,b,c,\tau), p(a,b,c,\tau)] = \iiiint d\mathbf{a}d\tau \left(\tfrac{1}{2}(x_\tau)^2 + \tfrac{1}{2}(y_\tau)^2 + \tfrac{1}{2}(z_\tau)^2 + z\theta(c) + p\left[\frac{\partial(x,y,z)}{\partial(a,b,c)} - 1\right]\right) \tag{3.2}$$

where $d\mathbf{a} = da\,db\,dc$ and, as in Chapter 2, $\theta(c)$ is a prescribed function. The requirement that (3.2) be stationary with respect to the variations

$$\delta x(a,b,c,\tau), \quad \delta y(a,b,c,\tau), \quad \delta z(a,b,c,\tau), \quad \delta p(a,b,c,\tau) \tag{3.3}$$

yields the Boussinesq equations used in Chapter 2. Since, as in Chapter 2, we are not terribly concerned with boundary conditions, we let the integrals extend to infinity, and we require the variations to vanish there. (If, on the other hand, a free surface were present, then the δp-variation would yield both the incompressibility constraint and the boundary condition $p = 0$ at the free surface.) The pressure p is a Lagrange multiplier that enforces the incompressibility constraint. For more background on the variational principles of fluid mechanics, see *LGFD*. For a general introduction to variational principles, I recommend the classic book by Cornelius Lanczos.

Although it is conventional to take the variations (3.3) in sequence, we may combine them arbitrarily. That is, we can take any sort of variation that we like. The result will be a valid equation. A *symmetry property* is any sneaky type of variation we can find that will go undetected by some or all of the terms in the Lagrangian. Mathematicians will cringe at this definition, but its flexibility proves useful. As our primary example, we consider the particle relabeling symmetry of (3.2).

In varying the dependence of (x,y,z) on (a,b,c) we are actually varying the time dependent mapping between two spaces. Since the mapping is one-to-one, the 'forward mapping' from (a,b,c) to (x,y,z) is equivalent to an 'inverse mapping' from to (x,y,z) to (a,b,c). In other words, we could consider the variations $\delta a(x,y,z)$, etc., instead of $\delta x(a,b,c)$, etc. One way to do this, proceeding formally, would be to rewrite (3.2) in a form that exchanges dependent and independent variables, replacing

$$\iiiint da\;db\;dc\;d\tau \to \iiiint dx\;dy\;dz\;dt\;\frac{\partial(a,b,c)}{\partial(x,y,z)} \tag{3.4}$$

and so forth. It is probably better to think about particle relabeling like this: At every time you change the label on every fluid particle by an infinitesimal amount, from (a, b, c) to (a', b', c'), where

$$a' = a + \delta a(a, b, c, \tau), \quad b' = b + \delta b(a, b, c, \tau), \quad c' = c + \delta c(a, b, c, \tau) \tag{3.5}$$

and $\delta a(a, b, c, \tau)$ etc are continuous functions. This is a bit like changing your social security number. Would that make you behave any differently? It would, if your Lagrangian depended on your social security number, but that is unlikely to be the case.

Take a hard look at the last two terms in (3.2). Can we find a sneaky variation that will go undetected by these terms? We cannot vary the label c because the prescribed function $\theta(c)$ would detect it. Therefore, we must choose $\delta c = 0$. Similarly, the variations δa and δb must be such that

$$\frac{\partial(a', b', c)}{\partial(x, y, z)} = \frac{\partial(a, b, c)}{\partial(x, y, z)} \iff \frac{\partial(a', b', c)}{\partial(a, b, c)} = 1 \tag{3.6}$$

Otherwise these variations would be detected by the Jacobian term in (3.2). Since the variations are infinitesimal, (3.6) implies

$$\frac{\partial \delta a}{\partial a} + \frac{\partial \delta b}{\partial b} = 0 \tag{3.7}$$

and thus

$$\delta a = -\gamma_b, \quad \delta b = +\gamma_a \tag{3.8}$$

for some $\gamma(a, b, c, \tau)$. For such variations,

$$\delta L = \iiiint d\mathbf{a} d\tau \left(x_\tau \delta(x_\tau) + y_\tau \delta(y_\tau) + z_\tau \delta(z_\tau) \right) \tag{3.9}$$

According to the variational principle, δL must vanish. But the time derivatives in (3.2) are certainly affected by the variations, because holding (a', b', c') fixed means something different than holding (a, b, c) fixed. Let us compute

$$\delta x_\tau = x_\tau|_{\mathbf{a}'} - x_\tau|_{\mathbf{a}} \tag{3.10}$$

By the chain rule,

$$\begin{aligned} x_\tau|_{\mathbf{a}'} &= x_\tau|_{\mathbf{a}} + x_a a_\tau|_{\mathbf{a}'} + x_b b_\tau|_{\mathbf{a}'} + x_c c_\tau|_{\mathbf{a}'} \\ &= x_\tau|_{\mathbf{a}} - x_a (\delta a)_\tau|_{\mathbf{a}'} - x_b (\delta b)_\tau|_{\mathbf{a}'} \end{aligned} \tag{3.11}$$

since $\delta c = 0$. Thus

$$\begin{aligned}\delta x_\tau &= -x_a(\delta a)_\tau|_{\mathbf{a}'} - x_b(\delta b)_\tau|_{\mathbf{a}'} \\ &\approx -x_a(\delta a)_\tau|_{\mathbf{a}} - x_b(\delta b)_\tau|_{\mathbf{a}}\end{aligned} \tag{3.12}$$

where the terms neglected in the second step are of second order in the variations. Substituting this and the corresponding equations for y and z into (3.2) we obtain

$$\begin{aligned}\delta L &= \iiiint d\mathbf{a}d\tau\,[-x_\tau(x_a\delta a_\tau + x_b\delta b_\tau) - y_\tau(y_a\delta a_\tau + y_b\delta b_\tau) - z_\tau(z_a\delta a_\tau + z_b\delta b_\tau)] \\ &= \iiiint d\mathbf{a}d\tau\,[\delta a(x_\tau x_a + y_\tau y_a + z_\tau z_a)_\tau + \delta b(x_\tau x_b + y_\tau y_b + z_\tau z_b)_\tau] \\ &= \iiiint d\mathbf{a}d\tau\;\gamma\;[(x_\tau x_a + y_\tau y_a + z_\tau z_a)_{\tau b} - (x_\tau x_b + y_\tau y_b + z_\tau z_b)_{\tau a}] \\ &= \iiiint d\mathbf{a}d\tau\;\gamma\;[(x_{\tau b} x_a + y_{\tau b} y_a + z_{\tau b} z_a)_\tau - (x_{\tau a} x_b + y_{\tau a} y_b + z_{\tau a} z_b)_\tau] \\ &= -\iiiint d\mathbf{a}d\tau\;\gamma\;\frac{\partial}{\partial\tau}\left[\frac{\partial(x_\tau,x)}{\partial(a,b)} + \frac{\partial(y_\tau,y)}{\partial(a,b)} + \frac{\partial(z_\tau,z)}{\partial(a,b)}\right]\end{aligned} \tag{3.13}$$

where we have freely used integrations by parts and the fact that the variations vanish at infinity. Since γ is arbitrary and δL must vanish, it follows that

$$\frac{\partial}{\partial\tau}\left[\frac{\partial(x_\tau,x)}{\partial(a,b)} + \frac{\partial(y_\tau,y)}{\partial(a,b)} + \frac{\partial(z_\tau,z)}{\partial(a,b)}\right] = 0 \tag{3.14}$$

which is just the potential vorticity conservation law derived in Chapter 2. But, you say, if we already knew about potential vorticity conservation, why go through all this? Why re-derive it in this seemingly roundabout way?

The answer has to do with approximations. Francis Bretherton and Gerald Whitham pioneered a strategy of applying approximations directly to the Lagrangian. The great advantage of this approach is that conservation laws survive the approximations if the approximations do not violate the corresponding symmetry properties. Equally important, new conservation laws arising from symmetries *introduced* by the approximations are easily identified. Wave action conservation is an example of such a 'new' conservation law. Potential vorticity is always important, even when it vanishes. Approximations that respect the particle relabeling symmetry automatically conserve a *form* of potential vorticity, but it may not be easy to find the

exact form. The symmetry property is a reliable guide. It never fails to find the form.

In this chapter we recover the results of Chapter 2 by applying approximations to the Lagrangian (3.2). Subsequent chapters extend the method and make it easier to use. Our strategy is to substitute the expansions (2.14)-(2.17) directly into (3.2) and to simplify the result before taking variations. For example, we substitute

$$x_\tau^2 = \left(x_\tau^{(1)} + x_\tau^{(2)}\right)^2 = (x_\tau^{(1)})^2 + 2x_\tau^{(1)}x_\tau^{(2)} + (x_\tau^{(2)})^2 \tag{3.15}$$

and similarly, so that the Lagrangian depends on both the superscript (1) and (2) variables. Variations of the superscript (1) variables give the evolution equations for the superscript (1) variables. Variations of the superscript (2) variables give the evolution equations for the superscript (2) variables. However, since we will *prescribe* the superscript (1) variables to have the form (2.32)-(2.37) of the basic wave, we may omit the terms in the Lagrangian that involve *only* superscript (1) variables. That leaves out the first term in (3.15). The term following it drops out when we average over wave phase, which, again, is justified by the fact that the mean flow does not respond to a rapidly oscillating source. The rapid variations blur to zero. The sole remaining term in (3.15) is the last term. By similar reasoning we can replace the first *three* terms in the integrand of (3.2) by

$$(x_\tau^{(2)})^2 + (y_\tau^{(2)})^2 + (z_\tau^{(2)})^2 \tag{3.16}$$

As for

$$\theta(c)\left(c + z^{(1)} + z^{(2)} + \cdots\right) \tag{3.17}$$

the first term is not varied, the second term fluctuates rapidly, and we will not be varying $z^{(n)}$ for $n > 2$. Thus we may replace $\theta(c)z$ by $\theta(c)z^{(2)}$.

Now the more delicate part of the operation, the Jacobian term in (3.2). Written out in full it is

$$(p^{(0)} + p^{(1)} + \cdots)\left[\frac{\partial(a + x^{(1)} + \cdots, b + y^{(1)} + \cdots, c + z^{(1)} + \cdots)}{\partial(a,b,c)} - 1\right] \tag{3.18}$$

Since the terms in (3.16) are $\mathrm{O}(A^4)$, we must keep terms up to $\mathrm{O}(A^4)$ in(3.18). This is going to be an awful lot of terms. Let us consider them order by order. At O(1) we have

$$p^{(0)}\,[1 - 1] = 0 \tag{3.19}$$

Good!

At O(A) the only contributing terms are linear in the superscript (1) variables and these average to zero. Good again.

At O(A^2) there are two classes of terms of terms: terms with two superscript (1) variables and two 'rest variables' like a, b, c or $p^{(0)}(c)$; and terms with one superscript (2) variable and three rest variables. Since the Jacobian vanishes when two of its arguments are proportional to the same thing (such as $A \sin \phi$) the only terms in the former category are

$$p^{(1)} \left(x_a^{(1)} + y_b^{(1)} + z_c^{(1)} \right) \tag{3.20}$$

But these are terms that depend only on the prescribed superscript (1) variables and hence will not be varied. In the latter category, we have the term $p^{(0)}(c) z_c^{(2)}$ which integrates to zero. Terms that integrate to zero cannot contribute to the variations.

At O(A^3) we have only terms that are linear or cubic in the superscript (1) variables and these again average to zero. Thus, in the approximate Lagrangian, we have only the terms

$$(x_\tau^{(2)})^2 + (y_\tau^{(2)})^2 + (z_\tau^{(2)})^2 + \theta(c) z^{(2)} \tag{3.21}$$

and whatever comes from (3.18) at O(A^4).

At O(A^4) things finally get interesting. There are three classes of terms. The first class has four superscript (1) variables that will not be varied. The second class has two superscript (2) variables and two rest variables. This class will contribute the terms

$$p^{(2)} \left(x_a^{(2)} + y_b^{(2)} + z_c^{(2)} \right) \tag{3.22}$$

as well as the terms

$$p^{(0)} \frac{\partial(y^{(2)}, z^{(2)})}{\partial(b, c)} + p^{(0)} \frac{\partial(x^{(2)}, z^{(2)})}{\partial(a, c)} \tag{3.23}$$

which, after integrations by parts, take the simpler form

$$-\theta(c) z^{(2)} \left(x_a^{(2)} + y_b^{(2)} \right) \tag{3.24}$$

The third class of O(A^4) terms contains one superscript (2) variable and two superscript (1) variables. Introducing the compact notation

$$[A, B, C] \equiv \frac{\partial(A, B, C)}{\partial(a, b, c)} \tag{3.25}$$

these terms are

$$\begin{aligned} p^{(1)} \Big(& [a, y^{(1)}, z^{(2)}] + [a, y^{(2)}, z^{(1)}] \\ & + [x^{(1)}, b, z^{(2)}] + [x^{(2)}, b, z^{(1)}] + [x^{(1)}, y^{(2)}, c] + [x^{(2)}, y^{(1)}, c] \Big) \\ & = p^{(1)} \Big(\frac{\partial(y^{(1)}, z^{(2)})}{\partial(b,c)} + \frac{\partial(y^{(2)}, z^{(1)})}{\partial(b,c)} \\ & + \frac{\partial(x^{(1)}, z^{(2)})}{\partial(a,c)} + \frac{\partial(x^{(2)}, z^{(1)})}{\partial(a,c)} + \frac{\partial(x^{(1)}, y^{(2)})}{\partial(a,b)} + \frac{\partial(x^{(2)}, y^{(1)})}{\partial(a,b)} \Big) \end{aligned} \tag{3.26}$$

Putting all this together, we obtain the Lagrangian for the mean flow in the form

$$\begin{aligned} & L[\mathbf{x}^{(2)}, p^{(2)}] = \\ & \iiiint d\mathbf{a} d\tau \Big[\frac{1}{2} \mathbf{x}_\tau^{(2)} \cdot \mathbf{x}_\tau^{(2)} + z^{(2)} \theta(c) + p^{(2)} \nabla \cdot \mathbf{x}^{(2)} - \theta(c) z^{(2)} \left(x_a^{(2)} + y_b^{(2)} \right) \\ & \qquad - x^{(2)} \left[\frac{\partial(p^{(1)}, y^{(1)})}{\partial(a,b)} + \frac{\partial(p^{(1)}, z^{(1)})}{\partial(a,c)} \right] \\ & \qquad - y^{(2)} \left[\frac{\partial(p^{(1)}, x^{(1)})}{\partial(b,a)} + \frac{\partial(p^{(1)}, z^{(1)})}{\partial(b,c)} \right] \\ & \qquad - z^{(2)} \left[\frac{\partial(p^{(1)}, x^{(1)})}{\partial(c,a)} + \frac{\partial(p^{(1)}, y^{(1)})}{\partial(c,b)} \right] \Big] \end{aligned} \tag{3.27}$$

after integrations by parts. Using integration by parts again, the terms in the second line can be written in the simpler form

$$z^{(2)} \theta(c) + \hat{p}^{(2)} \nabla \cdot \mathbf{x}^{(2)} \tag{3.28}$$

where we define $\hat{p}^{(2)} = p^{(2)} - \theta(c) z^{(2)}$ as in Chapter 2. (Independent variations of $x^{(2)}, y^{(2)}, z^{(2)}, p^{(2)}$ are equivalent to independent variations of $x^{(2)}, y^{(2)}, z^{(2)}, \hat{p}^{(2)}$.) With this change of variable, (3.27) takes the form

$$\begin{aligned} L[\mathbf{x}^{(2)}, p^{(2)}] = \iiiint d\mathbf{a} d\tau \Big(& \frac{1}{2} \mathbf{x}_\tau^{(2)} \cdot \mathbf{x}_\tau^{(2)} + z^{(2)} \theta(c) + \hat{p}^{(2)} \nabla \cdot \mathbf{x}^{(2)} \\ & - x^{(2)} X(a,b,c,\tau) - y^{(2)} Y(a,b,c,\tau) - z^{(2)} Z(a,b,c,\tau) \Big) \end{aligned} \tag{3.29}$$

where X, Y, Z are defined to be the corresponding square-bracket terms in (3.27). Since we *prescribe* the linear-wave solution that determines X, Y, Z,

we may regard these functions as given functions that are not to be varied. However, instead of the specific choice made in Chapter 2, we could choose any other solution of the linear equations. If, instead of wave packets, you are interested in vertical modes or internal wave beams, here is your chance.

You can check that variations of (3.29) with respect to $\mathbf{x}^{(2)}$ and $p^{(2)}$ yield the equations (2.41)-(2.44) for the evolution of the mean flow. If we evaluate the square-bracket terms in (3.27) by using the wave packet solution (2.32)-(2.37) to the leading order equations, we recover all the results of Chapter 2.

As in Chapter 2, we can avoid detailed analysis of the the superscript (2) solutions by projecting the forces X, Y, Z onto the potential vorticity mode. Again, our reasoning is that a slow forcing produces a slow response, and nothing is slower than potential vorticity: In Lagrangian coordinates it just sits there. The point to be emphasized here is that a rather obvious symmetry property of (3.29) tells us just how to do this.

Look closely at the terms

$$z^{(2)}\theta(c) + \hat{p}^{(2)}\left(x_a^{(2)} + y_b^{(2)} + z_c^{(2)}\right) \tag{3.30}$$

in (3.29). What sneaky variations can we make that will go undetected by these terms? Because of the metrical nature of (3.30), as opposed to (3.1), it will be better to seek variations in the form $\delta x(a,b,c)$, etc., rather than $\delta a(x,y,z)$, etc. We cannot vary $z^{(2)}$ because that would be detected by the first term in (3.30). So we must vary only $x^{(2)}$ and $y^{(2)}$, and in such a way that

$$\delta\left(x_a^{(2)} + y_b^{(2)}\right) = 0 \tag{3.31}$$

This implies

$$\delta x^{(2)} = -\gamma_b, \quad \delta y^{(2)} = +\gamma_a \tag{3.32}$$

for some $\gamma(a,b,c,\tau)$. For such a variation,

$$\begin{aligned}\delta L &= \iiiint d\mathbf{a}d\tau\ \left(-x_{\tau\tau}^{(2)}\delta x^{(2)} - y_{\tau\tau}^{(2)}\delta y^{(2)} - \delta x^{(2)}X - \delta y^{(2)}Y\right) \\ &= \iiiint d\mathbf{a}d\tau\ \gamma\left(-u_{b\tau}^{(2)} + v_{a\tau}^{(2)} + X_b - Y_a\right)\end{aligned} \tag{3.33}$$

Since γ is arbitrary,

$$\left(v_a^{(2)} - u_b^{(2)}\right)_\tau = Y_a - X_b \tag{3.34}$$

which is a key result of Chapter 2.

Strictly speaking, (3.34) is not a conservation law; it does not fit the form of

$$\frac{\partial}{\partial \tau} \text{ something} = 0 \tag{3.35}$$

But in the form we have written (3.29), X, Y, Z could be *any* prescribed, slowly varying functions. They need not have the from of an internal wave packet, and in fact they could represent a slowly varying external source. Such a source could be putting potential vorticity into the flow, and therefore no conservation law of the form (3.35) should be expected. We can obtain a conservation law of the form (3.35) if we use (2.69) to pop a time derivative out of the right hand side of (3.34). By invoking (2.69) we confer a pedigree on X and Y. We say, "You are not just *any* X and Y; rather you are the X and Y that come from a wave packet obeying internal wave dynamics."

How would this distinction emerge from the variational principle? In a very elegant fashion, but only if we extend the variational principle in such a way that we can obtain the dynamics of *both* the superscript (1) variables and the superscript (2) variables from the same principle. In other words, we must restore some of the terms we discarded on the way to (3.29). This we will do in the following chapters.

What to do if you still don't like avocado. Don't stop eating. In the case, to which we aspire, of waves and mean flows that genuinely *interact*, the variational approach offers an efficiency and transparency that is far superior to brute force perturbation theory.

Most important, variational principles allow us to introduce the important concept of symmetry groups. In later chapters we will suggest an analogy between the labels (a, b, c) and the scalar and vector potentials in electrodynamics. Just as you cannot have a variational principle for electrodynamics without introducing the electrodynamic potentials, you cannot have a variational principle for fluid dynamics without introducing particle labels. And the particle labels, like the electrodynamic potentials, are not unique. You can find a gauge transformation of either that will go undetected in the physical realm, and these undetectable variations correspond to conservation laws. Just as the gauge transformation of the fluid labels corresponds to the conservation of potential vorticity, the gauge transformation of electrodynamics corresponds to the conservation of electric charge. Most important of all, you can choose the gauge to suit the particular problem you are trying to solve. And here we venture a prediction: Future fluid mechanicists will be talking

a lot about gauge theory and group theory. The ghost of Garrett Birkhoff will rise and walk.

Chapter 4

Whitham's method

It is time for an assessment. What have we actually accomplished? In the two examples so far considered, the wave packet was prescribed consistent with linear dynamics and then told not to change. The mean flow responded, but in a highly subservient manner: It was powerless to alter the wave packet that had generated it. Our prescription of the wave packet itself was a very particular one. The amplitude A was allowed to vary slowly, but the wave vector $\mathbf{k}$ was held absolutely constant. Thus the wave packet in Chapters 1-3 corresponds to a quantum particle with a definite momentum but a poorly determined location.

All of these shortcomings must be fixed, and in Chapter 3 we hinted that variational principles would be our key to victory. But Chapter 3 offered a variational redux of Chapter 2 that hardly made things simpler. In fact, readers of Chapter 3 could be forgiven for thinking that the variational approach was actually more complicated than the more straightforward perturbation theory of Chapter 2.

In this chapter, we demonstrate the power and flexibility of the variational approach by deriving wave/mean equations in which the waves and the mean flow interact in an almost unrestricted manner. The straightforward derivation of these equations, by applying perturbation expansions directly to the fluid equations, is complicated. We obtain the same results more simply by using variational methods that are specially designed for our purpose.

We begin with the observation that fluid mechanists have widely differing attitudes to variational principles. In one corner are the physicists, for whom there is a single Ur-principle that is likely to be the particle mechanics one introduced in Chapter 3. All other variational principles must be 'derived'

from this Ur-principle by transformations and approximations.

In the other corner are the mathematicians. Their philosophy is the more pragmatic: First you find a set of equations that you like, and *then* you look to see if they are derivable from a variational principle. If so, the equations may be easier to analyze. Gerald Whitham inclined to the latter viewpoint. In the 1960's he developed a powerful variational method for studying linear and nonlinear waves and extended it to the case of wave-current interactions. Our methods are but a slight extension of those proposed by him.

Once again we consider internal gravity waves propagating through a stratified fluid. The dispersion relation is

$$\omega = \left[\frac{N_0^2(k^2+l^2)+f_0^2m^2}{(k^2+l^2+m^2)}\right]^{1/2} + \mathbf{U}\cdot\mathbf{k} \;\equiv\; \omega_r(k,l,m) + \mathbf{U}\cdot\mathbf{k} \tag{4.1}$$

where the wave vector $\mathbf{k}(x,y,z,t) = (k,l,m)$ and the mean velocity $\mathbf{U}(x,y,z,t)$ are slowly varying functions of space and time. ω_r is the *relative* or *intrinsic* frequency—the frequency measured in a reference frame moving with the mean velocity—and $\mathbf{U}\cdot\mathbf{k}$ is the Doppler shift. We are in rotating coordinates, but both the Coriolis parameter f_0 and the Vaisala frequency N_0 will be assumed constant.

Our general strategy has three main steps: First, find a variational principle for the waves by themselves. Whitham tells us how to do that. Second, find a variational principle for the mean flow (however defined) by itself. That will be supplied. Third, couple the waves to the mean flow using the Doppler term.

For the waves by themselves the general Lagrangian proposed by Whitham is

$$L_w[\phi,\mathcal{A}] = \iiiint dtd\mathbf{x}\;\left(\omega - \omega_r(k,l,m) - \mathbf{U}\cdot\mathbf{k}\right)\mathcal{A} \tag{4.2}$$

where

$$\omega = -\phi_t, \quad \mathbf{k} = \nabla\phi \tag{4.3}$$

are merely abbreviations for the derivatives of the phase, ϕ,

$$\mathcal{A} = \frac{E}{\omega_r} \tag{4.4}$$

is the wave action, and E is the energy density of the wave. At this stage, we consider the mean flow $\mathbf{U}(x,y,z,t)$ to be prescribed. We propose that (4.2) is *always* correct. To consider a particular type of wave, it is merely a matter

of inserting the appropriate expression for $\omega_r(\mathbf{k})$. To consider internal waves, we adopt the expression in (4.1). What justifies this? We postpone that question until later. For the moment we are content to show that variations of (4.2) yield good equations.

Whatever the choice for ω_r, $\delta L_w/\delta\mathcal{A} = 0$ implies the dispersion relation (4.1). For variations of ϕ we obtain

$$\begin{aligned}\delta L_w[\phi,\mathcal{A}] &= \iiiint dtd\mathbf{x}\ \left(-(\delta\phi)_t - \frac{\partial\omega_r}{\partial\mathbf{k}}\cdot\nabla(\delta\phi) - \mathbf{U}\cdot\nabla(\delta\phi)\right)\mathcal{A} \\ &= \iiiint dtd\mathbf{x}\ \left(\mathcal{A}_t + \nabla\cdot[(\mathbf{c}_g+\mathbf{U})\ \mathcal{A}]\right)\delta\phi \end{aligned} \tag{4.5}$$

where $\mathbf{c}_g(\mathbf{k}) = \partial\omega_r/\partial\mathbf{k}$ is the relative group velocity. Thus we have the action conservation equation,

$$\mathcal{A}_t + \nabla\cdot[(\mathbf{c}_g+\mathbf{U})\ \mathcal{A}] = 0 \tag{4.6}$$

As in the previous chapter, we ignore 'boundary contributions' that might arise from integrations by parts. One way to justify this is to say that we are considering an unbounded fluid that is quiescent at infinity. A better attitude is this: We assume that the variations themselves vanish at infinity, thereby recovering equations that hold only in the interior of the fluid, and are therefore incomplete in the sense that they omit the boundary conditions. This attitude is better because boundary conditions, which are always important, amount to arbitrary specifications of the way in which the fluid connects to its surroundings. Sometimes the boundary conditions can be incorporated into the variational principle, but that is seldom worthwhile.

Now for the second step: A variational principle for the mean flow. Asked about mean flows, nine people out of ten vote quasigeostrophic. The majority rules. The Lagrangian for quasigeostrophic dynamics is:

$$L_{qg}[\alpha,\beta,\psi] = \iiiint dt\mathbf{x}\ \left(-\alpha\beta_t + \psi\frac{\partial(\alpha,\beta)}{\partial(x,y)} + \frac{1}{2}\nabla_2\psi\cdot\nabla_2\psi + \frac{1}{2}\frac{f_0^2}{N_0^2}\psi_z^2\right) \tag{4.7}$$

The variables α, β and ψ depend on (x,y,z,t), but $\nabla_2 \equiv (\partial_x,\partial_y)$ is the two-dimensional gradient operator. $\delta L_{qg} = 0$ implies

$$\delta\alpha: \quad \beta_t + [\psi,\beta] = 0 \tag{4.8}$$

$$\delta\beta: \quad \alpha_t + [\psi,\alpha] = 0 \tag{4.9}$$

$$\delta\psi: \quad [\alpha,\beta] = \nabla_2^2\psi + \frac{f_0^2}{N_0^2}\psi_{zz} \equiv q \tag{4.10}$$

where

$$[A, B] \equiv \frac{\partial(A, B)}{\partial(x, y)} \tag{4.11}$$

We see that α and β are potential vorticity labels that are conserved following the quasigeostrophic flow. By taking the time-derivative of (4.10) and using the Jacobi identity,

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \tag{4.12}$$

we obtain the potential vorticity equation,

$$q_t + [\psi, q] = 0 \tag{4.13}$$

We could add the term f_0 to the right hand side of (4.10) by inserting a term $-f_0\psi$ in the integrand of (4.7), but, because f_0 is constant, this is an unnecessary complication.

Now for the final step. We couple L_w to L_{qg} by replacing $\mathbf{U}$ with the geostrophic velocity $\mathbf{u}_G = (u_G, v_G) \equiv (-\psi_y, \psi_x)$ in (4.2), and by assuming that the Lagrangian for the entire system is the sum, $L_w + L_{qg}$, of (4.2) and (4.7). The first part of this—the replacement—seems obvious, but the second part—the summation—involves a subtlety that must be explained. First we see what it gives. Instead of (4.10), $\delta(L_w + L_{qg})/\delta\psi = 0$ implies

$$\delta\psi : \quad [\alpha, \beta] = q - \nabla_2 \times (\mathbf{k}\mathcal{A}) \tag{4.14}$$

where $\nabla_2 \times (A, B) \equiv B_x - A_y$. As in previous chapters, we define the pseudomomentum

$$\mathbf{p} = \mathbf{k}\mathcal{A} \tag{4.15}$$

The same manipulation that led to (4.13) now leads to

$$Q_t + [\psi, Q] = 0 \tag{4.16}$$

where

$$Q \equiv q - \nabla_2 \times \mathbf{p} \tag{4.17}$$

The wave action equation (4.6) is unchanged, but now $\mathbf{U} = \mathbf{u}_G$

The most interesting effect of the coupling and summation is the generalization of (4.13) to (4.16)-(4.17). By these equations, the quantity $q - \nabla_2 \times \mathbf{p}$ is conserved following the *mean motion* of fluid particles. Consider waves

propagating into a region of fluid that is initially at rest. Before the arrival of the waves, $q = \mathbf{p} = 0$, and hence

$$q - \nabla_2 \times \mathbf{p} = 0 \tag{4.18}$$

By (4.16), (4.18) applies at all times. Hence

$$q = \nabla_2 \times \mathbf{p} \tag{4.19}$$

even when waves are present. Equation (4.19) is a concise general description of Bretherton flow. If wave breaking destroys the pseudomomentum $\mathbf{p}$ before the quasigeostrophic flow has time to react, then real, actual, potential vorticity q is created and remains behind after the remaining wave energy propagates away. We have seen all this before!

But there is more. By applying $\nabla = (\partial_x, \partial_y, \partial_z)$ to the dispersion relation (4.1), we obtain the refraction equation

$$\frac{\partial \mathbf{k}}{\partial t} + ((\mathbf{c}_g + \mathbf{u}_G) \cdot \nabla)\, \mathbf{k} = -k \nabla u_G - l \nabla v_G \tag{4.20}$$

where again $\mathbf{k} = \nabla\phi$. If the waves do not break, then the action $\mathcal{A} = E/\omega_r(\mathbf{k})$ is conserved. However, the refractive change in $\mathbf{k}$ predicted by (4.20) causes a change in $\omega_r(\mathbf{k})$ that can be determined from (4.1). If $\omega_r(\mathbf{k})$ increases, then the wave energy E must also increase to keep their ratio constant. For internal waves, $\omega_r(\mathbf{k})$ is always between f_0 and N_0. Suppose that $\omega_r \approx f_0$ initially. Then refractive changes in $\mathbf{k}$ can only increase ω_r, thereby increasing the wave energy E. This wave energy can only have come from the mean flow. We therefore conclude that inertial waves remove energy from the quasigeostrophic mean flow.

Such conclusions must be tempered by the realization that, unless the waves break, the evolution of the entire system is time reversible in the sense that any mechanical system can be made to 'run backwards' if, at any time, we reverse the velocity of *all* its constituents. Thus there will be solutions of our equations in which inertial waves transfer energy to the mean flow. But such behavior is *improbable* in the sense that it corresponds to an inertial spectral peak becoming sharper rather than spreading out. The generic behavior is a loss of mean energy to inertial waves.

Now we rewrite our complete Lagrangian as

$$L[\phi, A, \alpha, \beta, \psi] = L_w + L_{qg} = \iiiint dt d\mathbf{x}\ (-\phi_t \mathcal{A} - \alpha\beta_t) - \int dt\ H \tag{4.21}$$

where

$$H[\phi, A, \alpha, \beta, \psi]$$
$$= \iiint d\mathbf{x} \left(\omega_r \mathcal{A} - \psi \frac{\partial(\alpha,\beta)}{\partial(x,y)} - \frac{1}{2}\nabla_2\psi\cdot\nabla_2\psi - \frac{1}{2}\frac{f_0^2}{N_0^2}\psi_z^2 + \nabla_2\psi\times\mathbf{k}\mathcal{A}\right) \tag{4.22}$$

is the Hamiltonian, and $(\phi, \mathcal{A})$ and (α, β) form canonical pairs. Using (4.3), (4.14), and integrations by parts to evaluate (4.22) we find that

$$H = \iiint d\mathbf{x} \left(E + \frac{1}{2}\nabla_2\psi\cdot\nabla_2\psi + \frac{1}{2}\frac{f_0^2}{N_0^2}\psi_z^2\right) \tag{4.23}$$

The energy (4.23) is conserved. The equations corresponding to $\delta L = 0$ are

$$\delta\mathcal{A}: \quad \omega \equiv -\phi_t = \omega_r(k,l,m) + \mathbf{u}_G\cdot\mathbf{k} \tag{4.24}$$
$$\delta\phi: \quad \mathcal{A}_t + \nabla\cdot\left[(\mathbf{c}_g + \mathbf{u}_G)\,\mathcal{A}\right] = 0 \tag{4.25}$$
$$\delta\psi: \quad Q \equiv [\alpha,\beta] = \psi_{xx} + \psi_{yy} + \frac{f_0^2}{N_0^2}\psi_{zz} - \nabla_2\times(\mathbf{k}\mathcal{A}) \tag{4.26}$$
$$\delta\alpha: \quad \beta_t + [\psi,\beta] = 0 \tag{4.27}$$
$$\delta\beta: \quad \alpha_t + [\psi,\alpha] = 0 \tag{4.28}$$

Again, (4.26)-(4.28) imply

$$Q_t + [\psi, Q] = 0 \tag{4.29}$$

We are finally in a position to explain why it was correct to obtain the complete Lagrangian by *summing* L_w and L_{qg}. Such an explanation is needed because any Lagrangian can be multiplied by a constant factor (such as -1) without changing the equations that result from it. However, it would have been be a severe mistake to (say) reverse the sign of either L_w or L_{qg} by itself and then expect $\delta(L_w + L_{qg}) = 0$ to give us good equations. In particular, reversing the sign of only L_w, or only L_{qg}, would yield a system that conserved the *difference* between wave energy and quasigeostrophic energy. Our full Lagrangian was assembled with malice aforethought, by choosing the relative sign so that the energy came out right. This consideration must be added to our simple recipe, which now consists of four steps: Choose your dispersion relation; choose your mean flow; use the Doppler term to couple

them together; and, finally, add the two Lagrangians with weights carefully chosen to give the right energy.

What about the momentum? Just as energy conservation results from time-translation symmetry, momentum conservation results from space-translation symmetry. There is an easy method for extracting momentum conservation from our Lagrangian. It is the energy momentum formalism, which tells us that

$$\frac{\partial T^{ij}}{\partial x^j} = 0 \tag{4.30}$$

where $(x^0, x^1, x^2, x^3,) \equiv (t, x, y, z)$, and

$$T^{ij} \equiv \frac{\partial \phi^r}{\partial x^i} \frac{\partial \mathcal{L}}{\partial(\partial \phi^r / \partial x^j)} - \mathcal{L}\delta_{ij}, \tag{4.31}$$

Here $\mathcal{L}$ is the integrand of (4.21), and $(\phi^1, \phi^2, \phi^3, \phi^4, \phi^5) \equiv (\phi, \mathcal{A}, \alpha, \beta, \psi)$. Repeated indices are summed. The conservation law (4.30) assumes that $\mathcal{L}$ depends only on the ϕ^r and their first derivatives, and contains no explicit dependence on the x^i.

We focus on x-direction momentum. According to the energy momentum formalism, it obeys the conservation law

$$\frac{\partial T^{10}}{\partial t} + \frac{\partial T^{11}}{\partial x} + \frac{\partial T^{12}}{\partial y} + \frac{\partial T^{13}}{\partial z} = 0 \tag{4.32}$$

where

$$T^{10} = \phi_x \frac{\partial \mathcal{L}}{\partial \phi_t} + \beta_x \frac{\partial \mathcal{L}}{\partial \beta_t} = -k\mathcal{A} - \alpha\beta_x \tag{4.33}$$

$$T^{12} = \phi_x \frac{\partial \mathcal{L}}{\partial \phi_y} + \alpha_x \frac{\partial \mathcal{L}}{\partial \alpha_y} + \beta_x \frac{\partial \mathcal{L}}{\partial \beta_y} = -k\mathcal{A}(c_{gy} + v_g) - u_g v_g \tag{4.34}$$

$$T^{13} = \phi_x \frac{\partial \mathcal{L}}{\partial \phi_z} + \psi_x \frac{\partial \mathcal{L}}{\partial \psi_z} = -k\mathcal{A}c_{gz} + v_g \frac{f_0^2}{N_0^2}\psi_z \tag{4.35}$$

and $\mathbf{c}_g = (c_{gx}, c_{gy}, c_{gz}) = (\partial_k \omega_r, \partial_l \omega_r, \partial_m \omega_r)$ is the relative group velocity. We shall not require T^{11}. Thus (4.32) takes the form

$$\frac{\partial}{\partial t}(\alpha\beta_x + k\mathcal{A}) - \frac{\partial T^{11}}{\partial x} + \frac{\partial}{\partial y}(u_g v_g + k\mathcal{A}(c_{gy} + v_g)) + \frac{\partial}{\partial z}\left(-v_g \frac{f_0^2}{N_0^2}\psi_z + k\mathcal{A}c_{gz}\right) = 0 \tag{4.36}$$

If the flow is periodic in the x-direction, then (4.36) implies

$$\frac{\partial}{\partial t}\langle\alpha\beta_x+k\mathcal{A}\rangle+\frac{\partial}{\partial y}\langle u_g v_g+k\mathcal{A}(c_{gy}+v_g)\rangle+\frac{\partial}{\partial z}\langle -v_g\frac{f_0^2}{N_0^2}\psi_z+k\mathcal{A}c_{gz}\rangle=0 \quad (4.37)$$

where $\langle\cdot\rangle$ denotes the x-average. By assuming periodicity in x, we are associating x with longitude; then $\langle\cdot\rangle$ corresponds to the 'zonal average'.

To make sense of the first term in (4.37), we integrate (4.26), rewritten as

$$\nabla_2\times(\alpha\nabla\beta)=\nabla_2\times(\mathbf{u}_g-\mathbf{k}\mathcal{A})+\frac{f_0^2}{N_0^2}\psi_{zz} \quad (4.38)$$

over the 'polar cap' north of latitude y. By Stokes's theorem,

$$\langle\alpha\beta_x+k\mathcal{A}\rangle=\langle u_g\rangle+\int dy\ \langle\frac{f_0^2}{N_0^2}\psi_{zz}\rangle \quad (4.39)$$

where the y-integral is over the polar cap. (Technically, our equations apply to a periodic channel.) Substituting (4.39) into (4.37) we obtain the equation

$$\begin{aligned}&\frac{\partial}{\partial t}\langle u_g\rangle(y,z)+\frac{\partial}{\partial t}\int dy\ \langle\frac{f_0^2}{N_0^2}\psi_{zz}\rangle\\&\quad+\frac{\partial}{\partial y}\langle u_g v_g+k\mathcal{A}(c_{gy}+v_g)\rangle+\frac{\partial}{\partial z}\langle-\psi_x\frac{f_0^2}{N_0^2}\psi_z+k\mathcal{A}c_{gz}\rangle=0\end{aligned} \quad (4.40)$$

for the zonally averaged eastward flow. This is a favorite equation of meteorologists and of oceanographers studying the Antarctic Circumpolar Current. The last two terms in (4.40) represent the divergence of the Eliassen-Palm flux. The term proportional to $\psi_x\psi_z$ can be interpreted as vertical form drag. The y-integral represents the average layer thickness in the polar cap. In steady state, its time derivative vanishes. More generally, one speaks about 'residual velocity.' For further discussion, see other books. The point here is that conservation laws like (4.40) flow *automatically* from the Lagrangian and its symmetry properties. They do not require ingenuity to discover.

Energy and momentum conservation correspond to a symmetry of space-time that is shared by all physical systems: the property that physical law is the same everywhere and at all times. Physicists regard space-time symmetry as *kinematic*. However, a Lagrangian may also possess *dynamic* symmetries corresponding to conservation laws that are particular to that system. Our Lagrangian, (4.21), has two dynamic symmetries. One of these is the

particle-relabeling symmetry introduced in Chapter 3. In present context, it might better be called the potential-vorticity-relabeling symmetry. It corresponds to the fact that variations $\delta\alpha,\ \delta\beta$ that leave $[\alpha,\beta]$ undisturbed go undetected by (4.22). This leads to the conservation law (4.29). The other dynamic symmetry of (4.21) is 'phase-translation symmetry'. It corresponds to the action conservation law (4.25), but its physical content remains to be explained.

Our recipe adapts easily to other situations. To change the nature of the waves being considered, choose a different $\omega_r(\mathbf{k})$. If (4.2) and (4.4) are always correct, you do not even need to know *how* the wave energy E depends on wave amplitude. But *are* these equations always correct? And, most importantly, *why* is wave action conserved?

To address these questions we consider the Lagrangian corresponding to linear internal waves in the absence of a mean flow. The convenient variables are the particle displacements (ξ,η,ζ) from the state of rest. In terms of these variables the linear internal wave equations take the form

$$\xi_{tt} - f_0\eta_t = -p_x \tag{4.41}$$

$$\eta_{tt} + f_0\xi_t = -p_y \tag{4.42}$$

$$\zeta_{tt} = -p_z - N_0^2\zeta \tag{4.43}$$

$$\xi_x + \eta_y + \zeta_z = 0 \tag{4.44}$$

For simplicity we first consider the case $f_0 = 0$ in which the motion can be confined (without loss of generality) to the xz-plane. In this case $\eta \equiv 0$ and the continuity equation is solved by setting

$$\xi = -\gamma_z, \quad \zeta = \gamma_x \tag{4.45}$$

The governing equation is a vorticity equation in the xz-plane, for which the Lagrangian is

$$L[\gamma] = \iiint dxdzdt \left(\frac{1}{2}\gamma_{xt}^2 + \frac{1}{2}\gamma_{zt}^2 - \frac{1}{2}N_0^2\gamma_x^2\right) \tag{4.46}$$

The simplifying feature is that (4.46) depends on the single variable $\gamma(x,z,t)$. The resulting dynamics is

$$\delta\gamma: \quad (\gamma_{xx} + \gamma_{zz})_{tt} + N_0^2\gamma_{xx} = 0 \tag{4.47}$$

which admits solutions of the form

$$\gamma = A\cos(kx + mz - \omega t) \tag{4.48}$$

where A, k, and m are arbitrary constants, and

$$\omega^2 = \frac{N_0^2 k^2}{k^2 + m^2} \tag{4.49}$$

Whitham's method is to seek solutions in the form

$$\gamma = A(x,z,t)\cos(\phi(x,z,t)) \tag{4.50}$$

in which $A(x,z,t)$, $k(x,z,t) \equiv \phi_x$, $m(x,z,t) \equiv \phi_x$, and $\omega(x,z,t) \equiv -\phi_t$ all vary slowly compared to $\phi(x,z,t)$. This assumption justifies the approximations

$$\gamma_x \approx -A\phi_x \sin\phi, \quad \gamma_{xt} \approx -A\phi_x\phi_t\cos\phi, \quad \text{etc} \tag{4.51}$$

which, when substituted into (4.46), yield

$$L[A,\phi] = \iiint dxdzdt \left(\frac{1}{2}A^2(\phi_x^2+\phi_z^2)\phi_t^2\cos^2\phi - \frac{1}{2}N_0^2A^2\phi_x^2\sin^2\phi\right) \tag{4.52}$$

Whitham's final step is to average over the fast oscillations of the trigonometric terms, effectively replacing the $\cos^2\phi$ and $\sin^2\phi$ by $1/2$. The resulting *averaged Lagrangian*

$$L[A,\phi] = \iiint dxdzdt\ \frac{1}{4}A^2\left((\phi_x^2+\phi_z^2)\phi_t^2 - N_0^2\phi_x^2\right) \tag{4.53}$$

depends only on A and the first derivatives of ϕ. Thus (4.53) posesses a ϕ-translation symmetry that (4.52) does not. It is this symmetry property that corresponds to action conservation.

The Lagrangian (4.53) fits the form

$$L[A,\phi] \equiv \iiint dxdzdt\ \mathcal{L}(A,\phi_t,\phi_x,\phi_z) = \iiint dxdzdt\ A^2\mathcal{F}(\phi_t,\phi_x,\phi_z) \tag{4.54}$$

The variations yield

$$\delta A: \quad \mathcal{F}(\phi_t,\phi_x,\phi_z) = 0 \tag{4.55}$$

and

$$\delta\phi: \quad \frac{\partial}{\partial t}\left(A^2\frac{\partial\mathcal{F}}{\partial\phi_t}\right)+\frac{\partial}{\partial x}\left(A^2\frac{\partial\mathcal{F}}{\partial\phi_x}\right)+\frac{\partial}{\partial z}\left(A^2\frac{\partial\mathcal{F}}{\partial\phi_z}\right)=0 \tag{4.56}$$

Since (4.55) relates the wavenumbers ϕ_x and ϕ_z to the frequency $-\phi_t$, it can only be the dispersion relation. By (4.30)-(4.31), the energy density corresponding to (4.54) is

$$E=\phi_t\frac{\partial\mathcal{L}}{\partial\phi_t}-\mathcal{L}=A^2\left(\phi_t\frac{\partial\mathcal{F}}{\partial\phi_t}-\mathcal{F}\right)=A^2\phi_t\frac{\partial\mathcal{F}}{\partial\phi_t} \tag{4.57}$$

Thus the conserved quantity in (4.56) is E/ω, as expected. For the particular case (4.53),

$$E=\frac{1}{4}A^2\left((\phi_x^2+\phi_z^2)\phi_t^2+N_0^2\phi_x^2\right) \tag{4.58}$$

In asserting (4.2) we are assuming that $\mathcal{F}$ can always be written in the form

$$\mathcal{F}(\phi_t,\phi_x,\phi_z)=-\phi_t-\omega(\phi_x,\phi_z) \tag{4.59}$$

but the Lagrangian (4.53) does not fit this form. However,

$$L[\mathcal{A},\phi]=\iiint dxdzdt\;\mathcal{A}\left(-\phi_t-\sqrt{\frac{N_0^2\phi_x^2}{\phi_x^2+\phi_z^2}}\right) \tag{4.60}$$

is equivalent to (4.53) because they both yield the same dispersion relation and action equations.

At the heart of Whitham's method lie the steps between (4.46) and (4.53). They seem reasonable, based upon the separation in scales. Do they require a more rigorous justification? Whitham certainly thought so, for his book includes a demonstration that the results obtained by his variational method are equivalent to those obtained by more traditional asymptotic expansions applied directly to the equations governing the waves. He concludes: "The direct use of [asymptotic expansions] in the equations leads to the required results but without the generality and insight of the variational approach" (p. 402). The focus of Whitham's work was on *nonlinear* waves. Our use of his methods for the linear-waves portion of our wave/mean Lagrangian scarcely does him justice.

The Lagrangian corresponding to the full tree-dimensional equations (4.41)-(4.44) with $f_0 \neq 0$ is

$$L[\xi,\eta,\zeta,p] = \iiint d\mathbf{x}dt \, \left(\frac{1}{2}\xi_t^2+\frac{1}{2}\eta_t^2+\frac{1}{2}\zeta_t^2+f_0\xi\eta_t-\frac{1}{2}N_0^2\zeta^2+p(\xi_x+\eta_y+\zeta_z)\right) \tag{4.61}$$

For a single wavetrain we can rotate the z-axis so that nothing depends on y. Then (4.45) applies, and the dynamics (4.41)-(4.44) reduces to a vorticity equation in the xz-plane,

$$(\gamma_{xx}+\gamma_{zz})_{tt} + N_0^2\gamma_{xx} + f_0\eta_{zt} = 0 \tag{4.62}$$

coupled to the momentum equation in the y-direction,

$$\eta_{tt} - f_0\gamma_{zt} = 0 \tag{4.63}$$

which are together equivalent to the Lagrangian

$$L[\gamma,\eta] = \iiint dxdzdt \, \left(\frac{1}{2}\gamma_{xt}^2 + \frac{1}{2}\gamma_{zt}^2 + \frac{1}{2}\eta_t^2 - f_0\gamma_z\eta_t - \frac{1}{2}N_0^2\gamma_x^2\right) \tag{4.64}$$

Plane wave solutions of (4.62) and (4.63) take the form

$$\gamma = A\cos(kx+mz-\omega t), \quad \eta = A\frac{f_0 m}{\omega}\cos(kx+mz-\omega t) \tag{4.65}$$

Since our convention is that the (relative) frequency is always positive—so that the wave propagates in the direction of (k,m)—we see that γ and η are either in phase or 180 degrees out of phase depending on the sign of m. Substituting the analogues

$$\gamma = A(x,z,t)\cos\phi(x,z,t), \quad \eta = A(x,z,t)\frac{f_0\phi_z}{\omega(\phi_x,\phi_z)}\cos\phi(x,z,t) \tag{4.66}$$

of (4.50) back into (4.64) and averaging, we obtain

$$L[A,\phi] = \iiint dxdzdt \, \frac{1}{4}A^2\left(\phi_x^2\phi_t^2 + \phi_z^2\phi_t^2 - f_0^2\phi_z^2 - N_0^2\phi_x^2\right) \tag{4.67}$$

which fits the form (4.54) and can be handled in the same way. The key difference between the prescription (4.50) and the prescription (4.66) is that the latter includes a specification of modal *structure*—the relative amplitudes and relative phases of γ and η. Like the dispersion relation $\omega_r(\phi_x,\phi_y,\phi_z)$, this structure must be taken from linear theory.

Chapter 5

Wave-mean theory

We are gradually working our way toward a formulation of wave-mean interaction that is closely analogous to classical electrodynamics. The two ingredients of classical electrodynamics are electromagnetic waves and charged particles. The former are analogous to inertia-gravity waves, and the latter are analogous to fluid particles possessing potential vorticity. Potential vorticity is analogous to electric charge. Potential vorticity, we will argue, is the true 'slow variable' of wave-mean theory. Apart from the rapid vibrations caused by passing waves, potential vorticity moves at the relatively slow speed of the fluid itself. The influence of the waves moves at the usually much greater group velocity. The characterization of the flow associated with the potential vorticity as a *mean* flow is unfortunate, because it implies the need for an averaging that seems only to be an unnecessary complication.

The equations of classical electrodynamics are very difficult to solve in their general from. Textbooks and research papers focus on two special cases. In the first of these, the electromagnetic waves are prescribed, and one computes the motion of the charged particles from the Lorentz equations. This type of problem is analogous to the problems we have solved in the first three chapters. There the wave packet was prescribed, and we calculated the *virtual* vorticity associated with the wave packet. If the wave packet experienced dissipation, then some of this virtual vorticity was converted into real, actual vorticity.

The distinction between virtual and actual vorticity has no analogue in *classical* electrodynamics, because classical electromagnetic waves, unlike inertia gravity waves, do not interact. But electromagnetic waves *do* interact, albeit weakly, in *quantum* electrodynamics, producing virtual charges analo-

gous to our virtual vorticity. In this analogy, the Bretherton dipole depicted in Chapter 2 consists of a photon (the wave packet) and an electron/positron pair (the counter-rotating virtual vortices in figure 2.1.)

In the other class of simplified electrodynamics problem, one prescribes the motion of the electric charges, and one computes the radiated electromagnetic waves using Maxwell's equations. Such problems are analogous to fluids problems in which the motion of the vortices is prescribed and one computes the radiated inertia-gravity waves. Think of counter-rotating point vortices or the Kirchoff vortex patch. Such fluids problems are more complicated than their electrodynamic analogues, again because the inertia-gravity waves, unlike electromagnetic waves, interact with one another, but they have been successfully attacked by Lighthill, Crow, Ford, and others using the method of matched asymptotic expansions.

The most difficult type of electrodynamics problem is one in which neither the waves nor the motion of the charges (potential vorticity) is prescribed. That is the type of problem we consider in this chapter.

To simplify things somewhat, we consider the fluid motion to be confined to two space dimensions, x and y. All reference to z will be dropped. The quasigeostrophic flow degenerates into ordinary, two-dimensional, incompressible Euler flow—the equations of two-dimensional turbulence. The relative frequency $\omega_r(k,l)$ can be anything you like, but now it has only two arguments. Notation simplifies considerably. The need to distinguish between ∇ and ∇_2 disappears. (It arose from the fact that the quasigeostrophic velocity field is a two-dimensional vector field that nevertheless depends on (x,y,z).) In this chapter $\nabla = (\partial_x, \partial_y)$. Vorticity replaces potential vorticity.

The two-dimensional analogue of (4.22)-(4.23) is

$$L[\phi, A, \alpha, \beta, \psi] = \iiint dxdydt\ (-\phi_t \mathcal{A} - \alpha\beta_t) - \int dt\ H \tag{5.1}$$

where

$$H[\phi, A, \alpha, \beta, \psi] = \iint dxdy\ \left(\omega_r(\phi_x, \phi_y)\mathcal{A} - [\alpha, \beta]\psi - \frac{1}{2}\nabla\psi \cdot \nabla\psi + [\psi, \phi]\mathcal{A}\right) \tag{5.2}$$

and $[A, B] \equiv A_x B_y - A_y B_x$. The variations yield

$$\delta\mathcal{A}: \quad -\phi_t = \omega_r(\phi_x, \phi_y) + [\psi, \phi] \tag{5.3}$$

$$\delta\phi: \quad \mathcal{A}_t + \nabla \cdot (\mathbf{c}_g\, \mathcal{A}) + [\psi, \mathcal{A}] = 0 \tag{5.4}$$

$$\delta\beta: \quad \alpha_t + [\psi, \alpha] = 0 \tag{5.5}$$

$$\delta\alpha: \quad \beta_t + [\psi, \beta] = 0 \tag{5.6}$$

$$\delta\psi: \quad \nabla^2\psi - [\mathcal{A}, \phi] = [\alpha, \beta] \tag{5.7}$$

As in Chapter 4 the last three equations combine to give

$$Q_t + [\psi, Q] = 0, \quad Q = \nabla^2\psi - [\mathcal{A}, \phi] \tag{5.8}$$

where the last term in (5.8) is minus the curl of the pseudomomentum. Using the same Jacobi identity we used to obtain (5.8), we operate on (5.3) and (5.4) to get the equation for the evolution of pseudomomentum,

$$[\mathcal{A}, \phi]_t + [\psi, [\mathcal{A}, \phi]] = -[\nabla \cdot (\mathbf{c}_g\, \mathcal{A}), \phi] - [\mathcal{A}, \omega_r(\phi_x, \phi_y)] \tag{5.9}$$

Adding (5.8) and (5.9) we obtain

$$\nabla^2\psi_t + [\psi, \nabla^2\psi] = -[\nabla \cdot (\mathbf{c}_g\, \mathcal{A}), \phi] - [\mathcal{A}, \omega_r(\phi_x, \phi_y)] \tag{5.10}$$

which shows the critical influence of the choice of $\omega_r(\phi_x, \phi_y)$ in determining how the waves affect the mean flow.[1]

We begin our long trek in the direction of electrodynamics by observing that in electrodynamics charged particles such as the electron really are *point* particles. Since vorticity is analogous to electric charge, the electrodynamic analogy requires the vorticity to be concentrated in point vortices. Thus we assume

$$[\alpha, \beta] = \sum_i \Gamma_i \delta\left(\mathbf{x} - \mathbf{x}_i(t)\right) \tag{5.11}$$

where $\mathbf{x}_i(t)$ is the location at time t of a point vortex with strength Γ_i. The subscripts i replace the continuous vorticity labels α and β. The ansatz

[1] A word about units. Every term in the integrand of (5.2) should have the same units. If $\omega_r\mathcal{A}$ includes the depth integration, as it probably would in the case of surface waves, then it would be appropriate to multiply the $\nabla\psi \cdot \nabla\psi$ term (for example) by a constant depth H_0 over which the mean kinetic energy has also been integrated. Alternatively, such a factor could be absorbed into the definitions of ψ, α, and β. By not including such a notation-cluttering factor, I am implicitly adopting that approach.

(5.11) conserves vorticity because

$$\iint dxdy\ [\alpha,\beta] = \iint d\alpha d\beta = \sum_i \Gamma_i \tag{5.12}$$

are constant. The Hamiltonian (5.2) becomes

$$H[\phi, A, \mathbf{x}_i, \psi] = -\sum_i \Gamma_i \psi(\mathbf{x}_i(t)) + \iint dxdy\ \left(\omega_r(\phi_x,\phi_y)\mathcal{A} - \frac{1}{2}\nabla\psi\cdot\nabla\psi + [\psi,\phi]\mathcal{A}\right) \tag{5.13}$$

To fully convert from (α,β) to $\mathbf{x}_i$ we must transform the term

$$\iiint dxdydt\ \alpha\beta_t \tag{5.14}$$

in (5.1). It becomes

$$\begin{aligned}&\iiint dxdydt\ \alpha\frac{\partial(x,y,\beta)}{\partial(x,y,t)} = \iiint d\alpha d\beta d\tau\ \alpha\frac{\partial(x,y,\beta)}{\partial(\alpha,\beta,\tau)}\\ &= \iiint d\alpha d\beta d\tau\ -x\frac{\partial(\alpha,y,\beta)}{\partial(\alpha,\beta,\tau)} = \iiint d\alpha d\beta d\tau\ x\frac{\partial y}{\partial\tau} = \int dt \sum_i \Gamma_i x_i \frac{dy_i}{dt}\end{aligned} \tag{5.15}$$

Thus the complete Lagrangian takes the form

$$\begin{aligned}L[\phi, A, \mathbf{x}_i, \psi] = &\iiint dxdydt\ \left(-\phi_t\mathcal{A} - \omega_r(\phi_x,\phi_y)\mathcal{A} + \frac{1}{2}\nabla\psi\cdot\nabla\psi - [\psi,\phi]\mathcal{A}\right)\\ &+ \int dt\left(-\sum_i \Gamma_i x_i\frac{dy_i}{dt} + \sum_i \Gamma_i\psi(\mathbf{x}_i(t))\right)\end{aligned} \tag{5.16}$$

Suppose no waves are present; suppose $\mathcal{A}\equiv 0$. Then the Lagrangian (5.16) simplifies to

$$\begin{aligned}L[\mathbf{x}_i,\psi] = &-\int dt\sum_i \Gamma_i x_i\frac{dy_i}{dt}\\ &+\iiint dxdydt\ \left(\frac{1}{2}\nabla\psi\cdot\nabla\psi + \sum_i \Gamma_i\psi(\mathbf{x})\delta\left(\mathbf{x}-\mathbf{x}_i(t)\right)\right)\end{aligned} \tag{5.17}$$

The last term in (5.17) is an alternative way of writing the last term in (5.16). The equations resulting from (5.17) are

$$\delta \mathbf{x}_i : \quad \frac{d\mathbf{x}_i}{dt} = (-\psi_y(\mathbf{x}_i, t), \psi_x(\mathbf{x}_i, t)) \tag{5.18}$$

and

$$\delta\psi : \quad \nabla^2\psi = \sum_i \Gamma_i \delta\left(\mathbf{x} - \mathbf{x}_i(t)\right) \tag{5.19}$$

We solve (5.19) for

$$\psi(\mathbf{x}, t) = \frac{1}{2\pi} \sum_i \Gamma_i \ln|\mathbf{x} - \mathbf{x}_i(t)| \tag{5.20}$$

and substitute the result (5.20) back into (5.17). The second line of (5.17) simplifies as follows:

$$\begin{aligned}
&\iiint dxdydt \ -\frac{1}{2}\psi\nabla^2\psi + \int dt \sum_i \Gamma_i \psi(\mathbf{x}_i) \\
&= \int dt \ \left(-\frac{1}{2} + 1\right) \sum_i \Gamma_i \psi(\mathbf{x}_i) = \int dt \ \frac{1}{4\pi} \sum_{i \neq j} \Gamma_i \Gamma_j \ln|\mathbf{x}_i - \mathbf{x}_j| \\
&= \int dt \ \frac{1}{2\pi} \sum_{i > j} \Gamma_i \Gamma_j \ln|\mathbf{x}_i - \mathbf{x}_j|
\end{aligned} \tag{5.21}$$

and the result is the Lagrangian,

$$L[\mathbf{x}_i] = \int dt \left(-\sum_i \Gamma_i x_i \frac{dy_i}{dt} + \ \frac{1}{2\pi} \sum_{i > j} \Gamma_i \Gamma_j \ln|\mathbf{x}_i - \mathbf{x}_j| \right) \tag{5.22}$$

for point vortex dynamics,

$$\frac{d\mathbf{x}_i}{dt} = \frac{1}{2\pi} \sum_{j \neq i} \Gamma_j \frac{(y_j - y_i, x_i - x_j)}{|\mathbf{x}_i - \mathbf{x}_j|} \tag{5.23}$$

Three important points. First, it is generally quite *wrong* to substitute an equation resulting from the variational principle back into the Lagrangian. (See what happens if you substitute the dispersion relation (5.3) back into (5.1)-(5.2).) However, it *is* legitimate to use the equation obtained by varying

a particular field to eliminate that same field from the Lagrangian. To prove this to yourself, consider the analogous problem of finding the points at which the gradient of a multivariate function vanishes. Thus it is legal to use (5.20) to eliminate ψ from (5.17).

Second, in (5.21)-(5.23) we have thrown away the interaction of point vortices with themselves. This is an arbitrary modification required to eliminate infinite terms. It has *no* fundamental justification. The frequently heard argument that the self-interaction vanishes by symmetry, because a lone vortex cannot decide which way to move, is specious: A vortex *patch* in the presence of other vortices, however distant, would experience 'tidal forces' that would contribute to its motion. The neglect of point-vortex self-interaction is analogous to the neglect of the electron's self-energy in electrodynamics. It has no justification beyond the need to avoid infinities.

Third, point vortex dynamics exhibits action-at-a-distance, a feature seemingly at odds with electrodynamics. This action-at-a-distance is solely the result of our decision that the mean flow be incompressible. Some such restriction is always necessary to avoid the embarrassing circumstance that the mean flow develops waves of its own. Action-at-a-distance disappears when we develop the electrodynamic analogy more generally, without attempting to combine it with wave-mean theory. See the next chapter for that.

Now we return to (5.16) and consider what happens when waves are present ($\mathcal{A} \neq 0$). Instead of (5.19) we obtain

$$\delta\psi : \qquad \nabla^2\psi = \sum_i \Gamma_i \delta\left(\mathbf{x} - \mathbf{x}_i(t)\right) + [\mathcal{A}, \phi] \tag{5.24}$$

The solution is

$$\psi(\mathbf{x}, t) = \frac{1}{2\pi} \sum_i \Gamma_i \ln\left|(\mathbf{x} - \mathbf{x}_i(t)\right| + \psi_w(\mathbf{x}, t) \tag{5.25}$$

where (suppressing the time-dependence)

$$\psi_w(x, y) = \iint dx'dy'\ \rho(x', y') \frac{1}{2\pi} \ln|\mathbf{x} - \mathbf{x}'| \tag{5.26}$$

and

$$\rho = [\mathcal{A}, \phi] = \nabla\mathcal{A} \times \mathbf{k} \tag{5.27}$$

In classical electrodynamics the electrons are particles, but the waves are waves. In quantum electrodynamics the waves too become particles: photons.

The fluid analogues of photons are wave packets, small regions of the flow in which the action $\mathcal{A}$ is nonzero. We want to calculate the stream function field generated by a single wave packet. It is given by (5.26) with

$$\rho = \nabla\mathcal{A} \times \mathbf{k}_p(t) \tag{5.28}$$

where $\mathbf{k}_p(t)$ is the wave vector associated with the wave packet. We assume that $\mathbf{k}_p$ depends only on time; its variation within the wave packet is assumed negligible. The integration in (5.26) is over the area of the wave packet, the region of the flow in which $\mathcal{A} \neq 0$. Since we are only considering a single wave packet, we assume for convenience that it is located near $\mathbf{x} = 0$. This is still a messy problem. We make it easy by saying that we only want to know the stream function very far from the wave packet. For $r \equiv |\mathbf{x}| >> |\mathbf{x}'|$,

$$\ln|\mathbf{x} - \mathbf{x}'| \approx \ln r - \frac{\mathbf{x}\cdot\mathbf{x}'}{r^2} \tag{5.29}$$

and (5.26) becomes

$$\psi_w(\mathbf{x}) \approx \frac{\ln r}{2\pi}\iint dx'dy'\ \rho(x',y') - \frac{1}{2\pi}\frac{\mathbf{x}}{r^2}\cdot\iint dx'dy'\ \mathbf{x}'\rho(x',y') \tag{5.30}$$

The first term in (5.30) vanishes, because $\mathcal{A} = 0$ at the boundary of the wave packet. In the second term,

$$\iint dxdy\ \mathbf{x}\rho(x,y) = \iint dxdy\ \mathbf{x}\,(\mathcal{A}_x l_p - \mathcal{A}_y k_p) = \mathcal{A}_p(-l_p, k_p) \tag{5.31}$$

after integration by parts, where

$$\mathcal{A}_p = \iint dxdy\mathcal{A} \tag{5.32}$$

is the total action inside the wave packet. Thus the streamfunction response to a single wave packet at $\mathbf{x}_p(t)$ is

$$\psi(\mathbf{x}) = \frac{1}{2\pi}\frac{(\mathbf{x}-\mathbf{x}_p)\times\mathbf{k}_p}{|\mathbf{x}-\mathbf{x}_p|^2}\mathcal{A}_p \tag{5.33}$$

With a little fooling around, you can see that the stream function given by (5.33) has the same pattern as the Bretherton dipole depicted in Chapter 2.

The streamfunction response to many point vortices and many wave packets is clearly

$$\psi(\mathbf{x},t) = \sum_i \Gamma_i \psi_m(\mathbf{x},\mathbf{x}_i) + \sum_p \mathcal{A}_p \psi_d(\mathbf{x},\mathbf{x}_p,\mathbf{k}_p) \tag{5.34}$$

where

$$\psi_m(\mathbf{x},\mathbf{x}_i) \equiv \frac{1}{2\pi} \ln|\mathbf{x}-\mathbf{x}_i(t)| \tag{5.35}$$

is the response to a monopole at $\mathbf{x}_i$, and

$$\psi_d(\mathbf{x},\mathbf{x}_p,\mathbf{k}_p) \equiv \frac{1}{2\pi} \frac{(\mathbf{x}-\mathbf{x}_p)\times\mathbf{k}_p}{|\mathbf{x}-\mathbf{x}_p|^2} \tag{5.36}$$

is the response to a Bretherton dipole with wavevector $\mathbf{k}_p$ at $\mathbf{x}_p$. The constants Γ_i and $\mathcal{A}_p$ measure the strength of the monopole and the dipole, respectively. $\mathcal{A}_p$ is always positive but Γ_i can have either sign. Unless dissipation occurs $\mathcal{A}_p$ and Γ_i remain constant.

Our dipole calculation assumed that $\mathbf{x}$ is very far from the wave packet at $\mathbf{x}_p$. Is that a serious restriction? Yes and no. First of all, it is not fundamentally different from our treatment of the monopoles. A realistic vortex *patch* looks like a point vortex only if you are very far away from it. Close up, you see a monopole only if the patch really is a *point*, i.e. only if point vortices *actually exist*. Likewise our dipole formula holds right up to the wave packet only if 'point dipoles' *actually exist*. So we have a choice. Either we assume the existence of these idealized entities—point vortices and point dipoles—or we assume that the more realistic structures—vortex patches and wave packets of finite size—are widely separated in space. From a distance each vortex or wave packet sees the others as a dilute gas of monopoles and dipoles (figure 5.1).

In the more realistic Bretherton dipole depicted in Chapter 2, the dipole pattern covered the wave packet itself, but there we were careful to say that the wave packet consisted of a single hump. If the wave packet had instead consisted of two humps, then the stream function pattern would have been a quadrupole. This corresponds to keeping more terms in (5.29).

One feature of our treatment remains an enormous restriction: We refuse to allow each monopole or Bretherton dipole to interact with itself. More precisely, self-interaction such as wave breaking can occur, but it must be imposed by us. *We* decide if and when the wave packets collapse and release

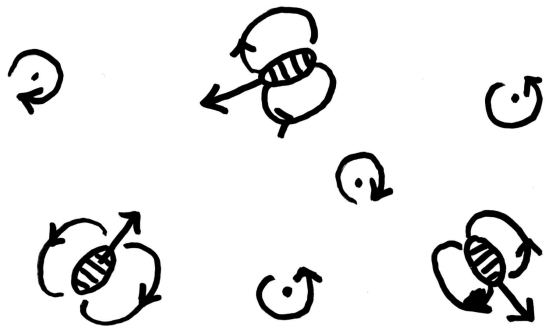

Figure 5.1: A dilute gas of Bretherton dipoles and point vortices.

their virtual vorticity. This is not a defect of our model. Rather, it is an indication of how idealized it is. Models, to be useful, can never include everything. For instance, people criticize Lagrangian methods for their neglect of dissipation. Certainly dissipation is important, but it can always be put in 'by hand.' The advantage of Lagrangian methods is that they do the conservative part of the dynamics correctly. They conserve the right things.

Our next step is to substitute (5.34) back into the Lagrangian (5.16), removing its dependence on ψ. However, since our aim is to produce a Lagrangian that depends only the point vortex locations $\mathbf{x}_i(t)$, the wave packet locations $\mathbf{x}_p(t)$, and their wave vectors $\mathbf{k}_p(t)$, we must transform *all* of the terms in (5.16). If we integrate the first term in (5.16) over the p-th wave packet, we obtain

$$\begin{aligned} -\iiint d\mathbf{x}dt\ \phi_t\mathcal{A} &= \iiint d\mathbf{x}dt\ \phi\mathcal{A}_t = -\iiint d\mathbf{x}dt\ \phi\ \frac{d\mathbf{x}_p}{dt}\cdot\nabla\mathcal{A} \\ = -\int dt\ \frac{d\mathbf{x}_p}{dt}\cdot\iint d\mathbf{x}\ \phi\nabla\mathcal{A} &= \int dt\ \frac{d\mathbf{x}_p}{dt}\cdot\iint d\mathbf{x}\ \mathcal{A}\nabla\phi = \int dt\ \frac{d\mathbf{x}_p}{dt}\cdot\mathbf{k}_p\mathcal{A}_p \end{aligned} \tag{5.37}$$

where we have used integrations by parts and the relation

$$\left(\frac{\partial}{\partial t} + \frac{d\mathbf{x}_p}{dt}\cdot\nabla\right)\mathcal{A}(\mathbf{x},t) = 0 \tag{5.38}$$

which follows from the definition of $\mathbf{x}_p(t)$: $d\mathbf{x}_p(t)/dt$ is the velocity of the wave envelope. The second term in (5.16) becomes

$$-\int dt\ \omega_r(\mathbf{k}_p)\mathcal{A}_p \tag{5.39}$$

The three terms in (5.16) containing ψ combine as

$$\begin{aligned}
&\iiint dtd\mathbf{x}\ \left(\frac{1}{2}\nabla\psi\cdot\nabla\psi - [\psi,\phi]\mathcal{A}\right) + \int dt \sum_i \Gamma_i\psi(\mathbf{x}_i(t),t) \\
&= \iiint dtd\mathbf{x}\left(-\frac{1}{2}\ \psi\nabla^2\psi - [\psi,\phi]\mathcal{A} + \ \sum_i \Gamma_i\psi\delta(\mathbf{x}-\mathbf{x}_i)\right) \\
&= \iiint dtd\mathbf{x}\ \left(-\frac{1}{2}\psi\nabla^2\psi + [\mathcal{A},\phi]\psi + \psi\left(\nabla^2\psi - [\mathcal{A},\phi]\right)\right) \\
&= \frac{1}{2}\iiint dtd\mathbf{x}\ \psi\nabla^2\psi
\end{aligned} \tag{5.40}$$

where we have used (5.24). The last integral in (5.40) is

$$\begin{aligned}
\iint d\mathbf{x}\ \psi\nabla^2\psi = \iint d\mathbf{x}\ \Bigg[&\left(\sum_i \Gamma_i\psi_m(\mathbf{x},\mathbf{x}_i) + \sum_p \mathcal{A}_p\psi_d(\mathbf{x},\mathbf{x}_p,\mathbf{k}_p)\right)\times \\
&\nabla^2\left(\sum_j \Gamma_j\psi_m(\mathbf{x},\mathbf{x}_j) + \sum_q \mathcal{A}_q\psi_d(\mathbf{x},\mathbf{x}_q,\mathbf{k}_q)\right)\Bigg]
\end{aligned} \tag{5.41}$$

We simplify (5.41) by neglecting the dipole-dipole interactions, which are expected to be weak: The velocity field associated with the monopoles falls off like $1/r$, whereas the velocity field associated with Bretherton dipoles falls off like $1/r^2$. Dropping these terms from (5.41) gives us

$$\begin{aligned}
&\iint d\mathbf{x}\ \psi\nabla^2\psi \approx \iint d\mathbf{x}\sum_i \Gamma_i\nabla^2\psi_m(\mathbf{x},\mathbf{x}_i)\left(\sum_j \Gamma_j\psi_m(\mathbf{x},\mathbf{x}_j) + 2\sum_p \mathcal{A}_p\psi_d(\mathbf{x},\mathbf{x}_p,\mathbf{k}_p)\right) \\
&= \iint d\mathbf{x}\sum_i \Gamma_i\delta(\mathbf{x}-\mathbf{x}_i)\left(\sum_j \Gamma_j\psi_m(\mathbf{x},\mathbf{x}_j) + 2\sum_p \mathcal{A}_p\psi_d(\mathbf{x},\mathbf{x}_p,\mathbf{k}_p)\right) \\
&= \sum_i \Gamma_i\left(\sum_j \Gamma_j\psi_m(\mathbf{x_i},\mathbf{x}_j) + 2\sum_p \mathcal{A}_p\psi_d(\mathbf{x_i},\mathbf{x}_p,\mathbf{k}_p)\right)
\end{aligned} \tag{5.42}$$

Putting all this together, we obtain the Lagrangian

$$L[\mathbf{x}_i,\mathbf{x}_p,\mathbf{k}_p] = \int dt\left(\sum_p \mathcal{A}_p\mathbf{k}_p\cdot\dot{\mathbf{x}}_p - \sum_i \Gamma_i x_i\dot{y}_i - H[\mathbf{x}_i,\mathbf{x}_p,\mathbf{k}_p]\right) \tag{5.43}$$

where

$$H[\mathbf{x}_i, \mathbf{x}_p, \mathbf{k}_p] = \sum_p \mathcal{A}_p\, \omega_r(\mathbf{k}_p) - \frac{1}{2\pi} \sum_i \sum_{j>i} \Gamma_i \Gamma_j \ln |\mathbf{x}_i - \mathbf{x}_j| \\ - \frac{1}{2\pi} \sum_i \sum_p \Gamma_i \mathcal{A}_p \frac{(\mathbf{x}_i - \mathbf{x}_p) \times \mathbf{k}_p}{|\mathbf{x}_i - \mathbf{x}_p|^2} \tag{5.44}$$

is the Hamiltonian. For every wave packet there are two canonical pairs, (x_p, k_p) and (l_p, y_p), and for every point vortex there is one canonical pair, (x_i, y_i). Again, Γ_i and $\mathcal{A}_p$ are constants. The Hamiltonian (5.44) contains $\Gamma\Gamma$ terms and $\Gamma\mathcal{A}$ terms. If we had not dropped the dipole/dipole interactions it would also contain $\mathcal{A}\mathcal{A}$ terms.

The equations corresponding to (5.43)-(5.44) are

$$\delta \mathbf{k}_p : \quad \dot{\mathbf{x}}_p = \frac{1}{\mathcal{A}_p} \frac{\partial H}{\partial \mathbf{k}_p} = \mathbf{c}_g(\mathbf{k}_p) + \mathbf{U}_m(\mathbf{x}_p) \tag{5.45}$$

$$\delta \mathbf{x}_p : \quad \dot{\mathbf{k}}_p = -\frac{1}{\mathcal{A}_p} \frac{\partial H}{\partial \mathbf{x}_p} = -k_p \nabla U_m(\mathbf{x}_p) - l_p \nabla V_m(\mathbf{x}_p) \tag{5.46}$$

$$\delta \mathbf{x}_i : \quad \dot{\mathbf{x}}_i = \frac{1}{\Gamma_i} \left(\frac{\partial H}{\partial y_i}, -\frac{\partial H}{\partial x_i} \right) = \mathbf{U}_m(\mathbf{x}_i) + \mathbf{U}_d(\mathbf{x}_i) \tag{5.47}$$

where

$$\mathbf{U}_m(\mathbf{x}) = (U_m(\mathbf{x}), V_m(\mathbf{x})) = \sum_i \Gamma_i \left(-\frac{\partial \psi_m}{\partial y}(\mathbf{x}, \mathbf{x}_i),\ \frac{\partial \psi_m}{\partial x}(\mathbf{x}, \mathbf{x}_i) \right) \\ = \frac{1}{2\pi} \sum_i \Gamma_i \frac{(y_i - y, x - x_i)}{|\mathbf{x}_i - \mathbf{x}|^2} \tag{5.48}$$

is the velocity field induced by the point vortices, and

$$\mathbf{U}_d(\mathbf{x}) = \sum_p \mathcal{A}_p \left(-\frac{\partial \psi_d}{\partial y}(\mathbf{x}, \mathbf{x}_p, \mathbf{k}_p),\ \frac{\partial \psi_d}{\partial x}(\mathbf{x}, \mathbf{x}_p, \mathbf{k}_p) \right) \tag{5.49}$$

is the velocity field induced by the wave packets. The total velocity is $\mathbf{U}(\mathbf{x}) = \mathbf{U}_m(\mathbf{x}) + \mathbf{U}_d(\mathbf{x})$. In our approximation, the wave packets talk to point vortices but not to one another, while the point vortices talk to both point vortices and wave packets. We can add the missing physics if necessary; it would, for example, add the term $\mathbf{U}_d(\mathbf{x}_p)$ to (5.45).

The dynamics (5.45)-(5.47) conserves the energy (5.44) and the momentum

$$\mathbf{M} = \sum_p \mathcal{A}_p \mathbf{k}_p + \sum_i \Gamma_i (y_i, -x_i) \tag{5.50}$$

Conservation of $\mathbf{M}$ is proved by considering variations of the form

$$\delta \mathbf{x}_i = \delta \mathbf{x}_p = \boldsymbol{\epsilon}(t) \tag{5.51}$$

where $\boldsymbol{\epsilon}(t)$ is an arbitrary infinitesimal vector. If we think of the interactions between the Bretherton dipoles and the point vortices as the sum of pair interactions between each dipole/vortex pair, then pairwise conservation of (5.50) tells us that the refraction of wave packet p (i.e. the change in $\mathbf{k}_p$) caused by vortex i is accompanied by a change in the position of vortex i. Bühler and McIntyre refer to this as 'remote recoil.'

Conservation of (5.50) also governs wave breaking in the following sense. If the p-th wave packet is completely destroyed by wave breaking, then $\mathcal{A}_p$ is suddenly replaced by two counter-rotating vortices with a dipole moment equal to ΓD where D is the separation between counter-rotating vortices of strength $\pm\Gamma$. See Bühler's book, p. 282 (first edition).

The dispersion relation $\omega_r(\mathbf{k})$ remains arbitrary. It can be anything you want. If your interest is in surface gravity waves, you would choose

$$\omega_r(\mathbf{k}) = \sqrt{g|\mathbf{k}|} \tag{5.52}$$

Whatever your choice, it is tempting to suppose that wave-vector stretching is *generic*. Wave vectors tend to get longer for the same reason that fluid particles tend to move apart. How does this affect the energy? It all depends on your choice of $\omega_r(\mathbf{k})$. If $\omega_r(\mathbf{k})$ increases with $|\mathbf{k}|$, then wave-vector stretching increases the first term in (5.44), drawing energy out of the other two terms. If, as in our approximate dynamics, the wave-vector stretching is produced by the vortical flow, this energy is drawn directly from the vortical flow. We have already made an argument like this for near-inertial waves in three dimensions. Now we see how the argument applies more generally. (We are not yet ready to discuss Rossby waves, for which $\omega_r(\mathbf{k})$ *decreases* with $|\mathbf{k}|$.)

In a famous 1949 paper, Onsager considered the equilibrium statistical mechanics of a system of point vortices. Our system reduces to Onsager's system when no waves are present ($\mathcal{A}_p \equiv 0$). Our phase space is larger than the one considered by Onsager because it contains dimensions corresponding

to the wave packet locations $\mathbf{x}_p$ and their wave vectors $\mathbf{k}_p$. However, the difference is not merely a matter of extra dimensions. In Onsager's problem the volume of the phase space is *finite*, because the point vortices are confined to a box. In our problem the phase space has infinite volume because $-\infty < \mathbf{k}_p < \infty$. We expect an ultraviolet catastrophe in which energy spreads to ever larger $|\mathbf{k}_p|$ by the process of wave vector stretching. If the wave vector stretching increases the first term in (5.44), as would be the case for gravity waves, this increase must be compensated by a decrease in the other two terms. Could this decrease be associated with pattern formation? Remember: In our problem the momentum invariant (5.50) is nontrivial, and must be treated on a par with the energy.

To illustrate the use of (5.43), we consider the system comprising a single wave packet with action $\mathcal{A}_p$ and wavevector (k_p, l_p) located at (x_p, y_p); a point vortex of strength $-\Gamma$ located at x_1, y_1; and a second point vortex of strength $+\Gamma$ located at x_2, y_2, as shown in figure 5.2. Initially,

$$y_p = l_p = 0, \quad x_2 = x_1, \quad y_2 = -y_1 \tag{5.53}$$

and, by symmetry, these conditions hold at all later times. (Warning: It is illegal to substitute the symmetry conditions (5.53) into the Lagrangian *before* taking the variations.) The Lagrangian is

$$\begin{aligned} &L[x_p, y_p, k_p, l_p, x_1, y_1, x_2, y_2] = \\ &\int dt \, \Big[\mathcal{A}_p \left(k_p \dot{x}_p + l_p \dot{y}_p - \omega_r(k_p, l_p) \right) + \Gamma \left(x_1 \dot{y}_1 - x_2 \dot{y}_2 \right) - \frac{\Gamma^2}{2\pi} \ln |\mathbf{x}_1 - \mathbf{x}_2| \\ &\qquad + \frac{\mathcal{A}_p}{2\pi} \left(-\Gamma \frac{(\mathbf{x}_1 - \mathbf{x}_p) \times \mathbf{k}_p}{|\mathbf{x}_1 - \mathbf{x}_p|^2} + \Gamma \frac{(\mathbf{x}_2 - \mathbf{x}_p) \times \mathbf{k}_p}{|\mathbf{x}_2 - \mathbf{x}_p|^2} \right) \Big] \end{aligned} \tag{5.54}$$

and the equations are

$$\delta k_p : \quad \dot{x}_p = c_g(k_p, 0) - \frac{\Gamma y_1}{\pi d^2} \tag{5.55}$$

$$\delta x_p : \quad \dot{k}_p = \frac{2\Gamma k_p}{\pi} \left(\frac{(x_1 - x_p) y_1}{d^4} \right) \tag{5.56}$$

$$\delta x_1 : \quad \dot{y}_1 = \frac{\mathcal{A}_p k_p}{\pi} \frac{(x_1 - x_p) y_1}{d^4} \tag{5.57}$$

$$\delta y_1 : \quad \dot{x}_1 = -\frac{\Gamma}{4\pi y_1} + \frac{\mathcal{A}_p k_p}{2\pi} \frac{(x_1 - x_p)^2 - y_1^2}{d^4} \tag{5.58}$$

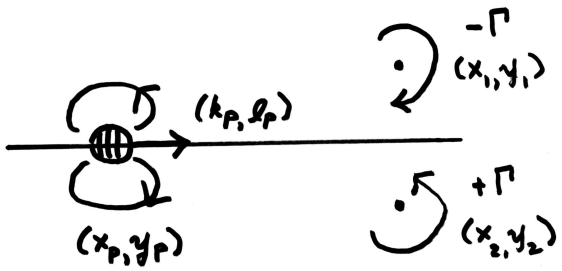

Figure 5.2: A wave packet with wavevector (k_p, l_p) at (x_p, y_p) propagates toward a pair of counter-rotating vortices at (x_1, y_1) and (x_2, y_2).

where c_g is the x-component of the group velocity, and

$$d^2 \equiv (x_1 - x_p)^2 + y_1^2 \tag{5.59}$$

is the squared distance between the wave packet and either vortex. On the left hand sides of (5.55)-(5.58), we have used the symmetry conditions (5.53). Because of these symmetry conditions, we do not need the evolution equations for y_p, l_p, x_2, and y_2.

The equations (5.55)-(5.58) conserve energy in the form

$$E = \omega_r(k_p)\mathcal{A}_p + \frac{\Gamma^2}{2\pi}\ln y_1 - \frac{\mathcal{A}_p\Gamma}{\pi}\frac{y_1 k_p}{d^2} \tag{5.60}$$

and momentum in the form

$$M = \mathcal{A}_p k_p - 2\Gamma y_1 \tag{5.61}$$

If we define

$$X(t) \equiv x_p(t) - x_1(t) \tag{5.62}$$

we can write (5.55)-(5.58) more simply as three equations

$$\dot{X} = c_g(k_p) + \frac{\Gamma}{4\pi d^2 y_1}(X^2 - 3y_1^2) - \frac{\mathcal{A}_p k_p}{2\pi d^4}(X^2 - y_1^2) \tag{5.63}$$

$$\dot{k}_p = -\frac{2\Gamma k_p}{\pi d^4} X y_1 \tag{5.64}$$

$$\dot{y}_1 = -\frac{\mathcal{A}_p k_p}{\pi d^4} X y_1 \tag{5.65}$$

in the three unknowns k_p, y_1 and X, where now $d^2 = X^2 + y_1^2$.

We have little need for these equations. The two conserved quantities, (5.60) and (5.61), make this an integrable system. Eliminating y_1 between (5.60) and (5.61), we obtain an expression for the energy in terms of X and k_p. The motion is confined to curves of constant $E(X, k_p)$. We determine the solution by plotting $E(X, k_p)$ or, even more conveniently, by plotting

$$E(k_p, d^2) = \omega_r(k_p)\mathcal{A}_p + \frac{\Gamma^2}{2\pi}\ln(\mathcal{A}_p k_p - M) - \frac{\mathcal{A}_p k_p}{2\pi}\frac{(\mathcal{A}_p k_p - M)}{d^2} \tag{5.66}$$

in which we have dropped additive constants. Only the last term in (5.66) involves d^2.

Consider a gravity wave packet, initially at $X = -\infty$ with $k_p > 0$, approaching the vortex pair from the left. While the wave packet is still far away from the vortex pair (d^2 very large) the last term in (5.66) is negligible. According to (5.64), k_p increases with time on $X < 0$. This increase in k_p occurs because the velocity field associated with the vortices is squeezing the wave packet in the x-direction. Since $c_g(k_p) > 0$ the wave energy $\omega_r \mathcal{A}_p$ and the vortex-interaction energy—the middle term in (5.66)—both increase with k_p. The increase in the latter corresponds to the two vortices being pushed apart by the velocity field associated with the dipole. The increase in these two terms must be balanced by the last term in (5.66), which represents the energy stored in the superposed velocity fields of the wave packets and vortices. These superposed fields tend to cancel as the wave packet approaches the vortex pair. k_p reaches its maximum value at $X = 0$ where

$$d^2 = y_1^2 = \left(\frac{\mathcal{A}_p k_p - M}{2\Gamma}\right)^2 \tag{5.67}$$

Substituting (5.67) into (5.66), we obtain an equation for this maximum value of k_p. After passing $X = 0$, the solution 'unwinds', and k_p returns to its original value as $X \to \infty$. Such solutions, easy because of the assumed symmetry, build physical intuition, but they tell us little about the behavior of a large number of wave packets and vortices.

Chapter 6

The analogy with electrodynamics

Chapter 5 broached an analogy between fluid mechanics and classical electrodynamics. Vorticity, or potential vorticity, is analogous to electric charge, and electromagnetic waves are analogous to whatever waves are present in the fluid: acoustic waves, surface waves, inertial waves, whatever. The wave-mean theory pursued in Chapter 5 clouded this analogy. It introduced action-at-a-distance, a decidedly non-electrodynamical characteristic, by assuming that the mean flow was either quasigeostrophic or incompressible Euler. In this chapter we pursue the electrodynamical analogy in its purer form, without attempting to divide the flow into waves and a mean flow. This leads to an important question, which we take up at the end of the chapter: Is the concept of a *mean* flow really necessary, or even helpful? Vorticity, we will suggest, is the true slow variable of fluid mechanics, and the averaging used to define other slow variables might be an unnecessary complication.

The analogy is between electrodynamics in two space dimensions on the one hand, and *quasi-two-dimensional* fluid dynamics on the other. What do we mean by that? We mean that the fluid must be stratified, and its isopycnal surfaces must not overturn. It is helpful if the fluid motion is also hydrostatic, but we can relax that requirement perturbatively. We consider shallow-water dynamics, which fulfills both of these restrictions. The generalization to three-dimensional, stratified, Boussinesq dynamics has been given elsewhere.

We begin with a quick review of electrodynamics. First, there are the

Maxwell equations:

$$\nabla \cdot \mathbf{B} = 0 \tag{6.1}$$

$$\nabla \cdot \mathbf{E} = \sum_i q_i \delta(\mathbf{x} - \mathbf{x}_i(t)) = q \tag{6.2}$$

$$\mathbf{B}_t + \nabla \times \mathbf{E} = 0 \tag{6.3}$$

$$c^2 \nabla \times \mathbf{B} - \mathbf{E}_t = \sum_i q_i \dot{\mathbf{x}}_i \delta(\mathbf{x} - \mathbf{x}_i(t)) = \mathbf{j} \tag{6.4}$$

where $\mathbf{B}(x, y, z, t)$ is the magnetic field; $\mathbf{E}(x, y, z, t)$ is electric field; q_i is the charge on the particle located at $\mathbf{x}_i(t)$; q is the charge density; $\mathbf{j}$ is the current density; and c is the speed of light. The charged particles obey Lorentz's equations,

$$m_i \ddot{\mathbf{x}}_i = q_i \left[\mathbf{E}(\mathbf{x}_i, t) + \dot{\mathbf{x}}_i \times \mathbf{B}(\mathbf{x}_i, t)\right] \tag{6.5}$$

where m_i is the mass of the charged particle. The acceleration term in (6.5) is non-relativistic. This violates Lorentz invariance, but we could easily replace it by the correct, relativistic expression. Shortly we will neglect it entirely.

The potential representation

$$\mathbf{E} = -\nabla\psi - \mathbf{A}_t \tag{6.6}$$

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{6.7}$$

automatically satisfies the two homogeneous Maxwell equations, (6.1) and (6.3). More importantly, it allows us to formulate a variational principle for electrodynamics. The Lagrangian for classical electrodynamics is $L = L_1 + L_2$ where

$$\begin{aligned} L_1[\psi, \mathbf{A}] &= \frac{1}{2} \iiiint dt d\mathbf{x} \left(\mathbf{E} \cdot \mathbf{E} - c^2 \mathbf{B} \cdot \mathbf{B}\right) \\ &= \frac{1}{2} \iiiint dt d\mathbf{x} \left[(\nabla\psi + \mathbf{A}_t) \cdot (\nabla\psi - + \mathbf{A}_t) - c^2 (\nabla \times \mathbf{A}) \cdot (\nabla \times \mathbf{A})\right] \end{aligned} \tag{6.8}$$

is the Lagrangian for the fields by themselves, and

$$\begin{aligned} L_2[\psi, \mathbf{A}, \mathbf{x}_i] = &\frac{1}{2} \sum_i m_i \dot{\mathbf{x}}_i \cdot \dot{\mathbf{x}}_i \\ &\iiiint dt d\mathbf{x} \sum_i q_i \left[-\psi(\mathbf{x}, t) + \mathbf{A}(\mathbf{x}, t) \cdot \dot{\mathbf{x}}_i\right] \delta(\mathbf{x} - \mathbf{x}_i(t)) \end{aligned} \tag{6.9}$$

couples the fields and the charges. In (6.9) we have written the kinetic energy of the particles in its non-relativistic form, but, again, this could easily be generalized.

The potential representation (6.6)-(6.7) is not unique. The gauge transformation

$$\mathbf{A} \to \mathbf{A} + \nabla\lambda \tag{6.10}$$

$$\psi \to \psi - \lambda_t \tag{6.11}$$

where $\lambda(x, y, z, t)$ is an arbitrary function, alters the potentials $\mathbf{A}$ and ψ, but not the physical fields $\mathbf{E}$ and $\mathbf{B}$. Similarly, the Lagrangian is invariant to variations of the form $\delta\mathbf{A} = \nabla(\delta\lambda)$ and $\delta\psi = -(\delta\lambda)_t$, where $\delta\lambda(x, y, z, t)$ is an arbitrary infinitesimal function. This symmetry property corresponds to a conservation law, the conservation of electric charge.

We specialize to the two-dimensional case, in which

$$\mathbf{A}(x, y, t) = (A, B, 0) \tag{6.12}$$

$$\mathbf{E}(x, y, t) = (E_1, E_2, 0) = (-\psi_x - A_t, -\psi_y - B_t, 0) \tag{6.13}$$

$$\mathbf{B}(x, y, t) = (0, 0, B_3) = (0, 0, B_x - A_y) \tag{6.14}$$

$$\mathbf{x}_i(t) = (x_i(t), y_i(t), 0) \tag{6.15}$$

(The vertical component, B_3, of $\mathbf{B}$ must not be confused with the y-component, B, of $\mathbf{A}$.) In two dimensions, (6.8) and (6.9) become

$$L_1[\psi, A, B] = \frac{1}{2}\iiint dtdxdy \left[(\psi_x + A_t)^2 + (\psi_y + B_t)^2 - c^2(B_x - A_y)^2\right] \tag{6.16}$$

and

$$L_2[\psi, A, B, x_i, y_i] = \frac{1}{2}\sum_i m_i \left(\dot{x}_i^2 + \dot{y}_i^2\right)$$
$$+ \iiint dtdxdy \sum_i q_i \left[-\psi(x, y, t) + A(x, y, t)\dot{x}_i + B(x, y, t)\dot{y}_i\right] \delta(\mathbf{x} - \mathbf{x}_i(t)) \tag{6.17}$$

The last term in (6.17) may also be written

$$\int dt \sum_i q_i \left[-\psi(x_i(t), y_i(t), t) + A(x_i(t), y_i(t), t)\dot{x}_i + B(x_i(t), y_i(t), t)\dot{y}_i\right] \tag{6.18}$$

The variational principle $\delta(L_1 + L_2) = 0$ yields the following equations:

$$\delta A: \quad (A_t + \psi_x)_t + c^2(B_x - A_y)_y = \sum_i q_i \dot{x}_i \delta(\mathbf{x} - \mathbf{x}_i(t))$$

$$\iff -E_{1,t} + c^2 B_{3,y} = j_1(\mathbf{x}, t) \tag{6.19}$$

$$\delta B: \quad (B_t + \psi_y)_t - c^2(B_x - A_y)_x = \sum_i q_i \dot{y}_i \delta(\mathbf{x} - \mathbf{x}_i(t))$$

$$\iff -E_{2,t} - c^2 B_{3,x} = j_2(\mathbf{x}, t) \tag{6.20}$$

$$\delta\psi: \quad (-\psi_x - A_t)_x + (-\psi_y - B_t)_y = \sum_i q_i \delta(\mathbf{x} - \mathbf{x}_i(t))$$

$$\iff E_{1,x} + E_{2,y} = q(\mathbf{x}, t) \tag{6.21}$$

$$\delta x_i: \quad m_i \ddot{x}_i = q_i[(-\psi_x - A_t) + (B_x - A_y)\dot{y}_i]$$

$$\iff m_i \ddot{x}_i = q_i[E_1 + B_3 \dot{y}_i] \tag{6.22}$$

$$\delta y_i: \quad m_i \ddot{y}_i = q_i[(-\psi_y - B_t) - (B_x - A_y)\dot{x}_i]$$

$$\iff m_i \ddot{y}_i = q_i[E_2 - B_3 \dot{x}_i] \tag{6.23}$$

The first three of these can also be written in the forms

$$A_{tt} - c^2\nabla^2 A + (\psi_t + c^2 A_x + c^2 B_y)_x = j_1 \tag{6.24}$$

$$B_{tt} - c^2\nabla^2 B + (\psi_t + c^2 A_x + c^2 B_y)_y = j_2 \tag{6.25}$$

$$-\nabla^2\psi - (A_x + B_y)_t = q \tag{6.26}$$

in which we have rearranged the terms on the left-hand sides. We simplify (6.24)-(6.26) by adopting the *Lorenz gauge*, that is, by arbitrarily requiring the potentials to satisfy

$$\psi_t + c^2(A_x + B_y) = 0 \tag{6.27}$$

Then (6.24)-(6.26) take the simpler forms

$$A_{tt} - c^2\nabla^2 A = j_1 \tag{6.28}$$

$$B_{tt} - c^2\nabla^2 B = j_2 \tag{6.29}$$

$$\psi_{tt} - c^2\nabla^2\psi = c^2 q \tag{6.30}$$

What permits us to impose (6.27)? Suppose that ψ, A, and B do not satisfy (6.27); suppose

$$\psi_t + c^2(A_x + B_y) = R(x, y, t) \tag{6.31}$$

where $R \neq 0$. The transformation (6.10)-(6.11) allows us to satisfy (6.27) if we can find a $\lambda(x, y, t)$ that satisfies

$$\lambda_{tt} - c^2 \nabla^2 \lambda = -R \tag{6.32}$$

This, we know, is always possible.

The new form (6.28)-(6.30) of Maxwell's equations is an improvement over (6.24)-(6.26) in that only a single potential appears in each equation, and the wave operator $\partial_{tt} - c^2 \nabla^2$ (the d'Alembertian) reminds us that all information travels at the speed of light. There is no action-at-a-distance. To have any hope of actually solving the equations, it is better to consider (6.28)-(6.30) instead of (6.24)-(6.26). Thus the gauge freedom of the potentials makes the equations easier to solve.

Despite the advantages of the Lorenz gauge, people sometimes prefer the *Coulomb gauge*, which imposes the condition

$$A_x + B_y = 0 \tag{6.33}$$

instead of (6.27). The gauge transformation (6.10)-(6.11) allows us to satisfy (6.33) if we can find a λ that satisfies

$$\nabla^2 \lambda = R \tag{6.34}$$

at every time. This equation—Poisson's equation—always has a solution, and therefore the restriction (6.33) is allowed. If (6.33) holds, then (6.26) takes the simpler form

$$-\nabla^2 \psi = q \tag{6.35}$$

but (6.24)-(6.25) remain complicated. Nevertheless, the Coulomb gauge is sometimes advantageous.

The Coulomb gauge raises an interesting paradox that was not apparently fully resolved until the paper of Brill and Goodman in 1967: Eqn (6.35) *looks like* action-at-a-distance, even though the physics must be the same as (6.28)-(6.30). For the resolution of this paradox, see any advanced book on electrodynamics, especially Jackson. (Hint: (6.34) looks like action-at-a-distance too.) The Coulomb gauge turns out to be a good choice for wave/mean theory, but it does *not* by itself impose action-at-a-distance.

Now we turn to the shallow water equations. We obtain shallow water dynamics from the Lagrangian for electrodynamics in four steps. First, we associate the potentials (ψ, A, B) with new physical variables:

$$\hat{h} \equiv \frac{h}{h_0} = B_x - A_y \tag{6.36}$$

$$\hat{h}u = -\psi_y - B_t \tag{6.37}$$

$$\hat{h}v = \psi_x + A_t \tag{6.38}$$

where $h(x, y, t)$ is the fluid depth; h_0 is the constant mean depth; $\mathbf{u}(x, y, t) = (u, v)$ is the fluid velocity in the (x, y) direction; and $c = \sqrt{gh_0}$. Equations (6.36)-(6.38) replace (6.13)-(6.14). We see that the fluid depth is analogous to the vertical component of the magnetic field. The horizontal momenta are analogous to the two horizontal components of the electric field, but note the sign flip between (6.37) and (6.38).

Second, we replace q_i with $-\Gamma_i$ and regard $\mathbf{x}_i(t) = (x_i(t), y_i(t))$ as the location of a point vortex with potential vorticity Γ_i. The sign flip between charge and vorticity is inconsequential, and is merely a matter of convention. (It is an accident of history that electrons were assigned a negative charge.)

Third, we drop the first term in (6.17). In other words, we assume that $m_i = 0$. This amounts to replacing (6.17) by

$$L_2^{sw} = \iiint dtdxdy \sum_i \Gamma_i \left[\psi(x, y, t) - A(x, y, t)\dot{x}_i - B(x, y, t)\dot{y}_i\right] \delta(\mathbf{x} - \mathbf{x}_i(t)) \tag{6.39}$$

where the superscript denotes 'shallow water.' This has the effect of replacing the Lorentz force law by

$$0 = \mathbf{E}(\mathbf{x}_i, t) + \dot{\mathbf{x}}_i \times \mathbf{B}(\mathbf{x}_i, t) \tag{6.40}$$

By the new interpretations (6.36)-(6.38), (6.40) is equivalent to

$$0 = \hat{h}(-v, u) + \hat{h}(\dot{y}_i, -\dot{x}_i) \tag{6.41}$$

from which it follows that $\dot{\mathbf{x}}_i = \mathbf{u}(\mathbf{x}_i, t)$; the point vortices move at the fluid velocity. Because Γ_i has divided out of (6.40), the attraction or repulsion between charges—vortices—disappears.

Fourth, we replace (6.16) by

$$L_1^{sw}[\psi, A, B] = \frac{1}{2} \iiint dtdxdy \left(\frac{(\psi_x + A_t)^2}{(B_x - A_y)} + \frac{(\psi_y + B_t)^2}{(B_x - A_y)} - c^2(B_x - A_y)^2 \right) \tag{6.42}$$

in which the first two terms acquire denominators. If these denominators were absent, L_1^{sw} would, like L_1, be quadratic in the potentials, and the equations resulting from $\delta L_1^{sw} = 0$ would be linear. The fact that L_1 is quadratic means that electromagnetic waves are linear waves that do not interact. The presence of denominators in (6.42) means that the shallow water gravity waves *do* interact, even in the absence of vorticity.

From the standpoint of electrodynamics, our assumption that $m_i = 0$ seems reasonable. Charged particles of negligible or zero mass are conceivable. On the other hand, the appearance of the denominators in (6.42) seems to lie wholly outside the realm of electrodynamics. In a sense, it does not. In *quantum* electrodynamics, electromagnetic waves do interact, even in the absence of electric charge, by the creation of 'virtual charge.' The virtual charge corresponds to Bretherton flow.

The variational principle $\delta(L_1^{sw} + L_2^{sw}) = 0$ implies:

$$\delta A: \quad \left(\frac{\psi_x + A_t}{B_x - A_y}\right)_t + \left(c^2(B_x - A_y) + \frac{1}{2}\frac{(\psi_x + A_t)^2}{(B_x - A_y)^2} + \frac{1}{2}\frac{(\psi_y + B_t)^2}{(B_x - A_y)^2}\right)_y = -\sum_i \Gamma_i \dot{x}_i \delta(\mathbf{x} - \mathbf{x}_i(t)) \tag{6.43}$$

$$\delta B: \quad -\left(\frac{\psi_y + B_t}{B_x - A_y}\right)_t + \left(c^2(B_x - A_y) + \frac{1}{2}\frac{(\psi_x + A_t)^2}{(B_x - A_y)^2} + \frac{1}{2}\frac{(\psi_y + B_t)^2}{(B_x - A_y)^2}\right)_x = \sum_i \Gamma_i \dot{y}_i \delta(\mathbf{x} - \mathbf{x}_i(t)) \tag{6.44}$$

$$\delta\psi: \quad \left(\frac{\psi_x + A_t}{B_x - A_y}\right)_x + \left(\frac{\psi_y + B_t}{B_x - A_y}\right)_y = \sum_i \Gamma_i \delta(\mathbf{x} - \mathbf{x}_i(t)) \equiv Q \tag{6.45}$$

$$\delta x_i: \quad (-\psi_x - A_t) + (B_x - A_y)\dot{y}_i = 0 \tag{6.46}$$

$$\delta y_i: \quad (-\psi_y - B_t) - (B_x - A_y)\dot{x}_i = 0 \tag{6.47}$$

When translated into the fluid variables (6.36)-(6.38), (6.43)-(6.47) become the shallow water equations:

$$v_t + \left(gh + \frac{1}{2}u^2 + \frac{1}{2}v^2\right)_y = -Qu \tag{6.48}$$

$$u_t + \left(gh + \frac{1}{2}u^2 + \frac{1}{2}v^2\right)_x = Qv \tag{6.49}$$

$$v_x - u_y = Q \tag{6.50}$$

$$\dot{\mathbf{x}}_i = \mathbf{u}(\mathbf{x}_i, t) \tag{6.51}$$

where Q is the vorticity density. The continuity equation

$$\hat{h}_t + \nabla \cdot (\hat{h}\mathbf{u}) = 0 \tag{6.52}$$

is automatically satisfied by (6.36)-(6.38). That is, the potential representation (6.36)-(6.38) automatically satisfies (6.52) in the same way that (6.6)-(6.7) automatically satisfies (6.3). (The other homogeneous Maxwell equation (6.1) is trivially satisfied in two dimensions.) This correspondence is the means by which the analogy was discovered: If we rewrite (6.52) in the form of a vanishing three-dimensional spacetime divergence,

$$(\partial_t, \partial_x, \partial_y) \cdot (\hat{h}, \hat{h}u, \hat{h}v) = 0 \tag{6.53}$$

then it is obvious that (6.53) implies

$$(\hat{h}, \hat{h}u, \hat{h}v) = (\partial_t, \partial_x, \partial_y) \times (-\psi, A, B) \tag{6.54}$$

for some ψ, A, B. Eqn (6.54) is equivalent to (6.36)-(6.38), and at this point the analogy reveals itself.

The automatic satisfaction of mass conservation by the potential representation (6.36)-(6.38) is analogous to the automatic satisfaction of mass conservation by the labeling of fluid particles in earlier chapters. Neither the potential representation nor the particle labeling is unique; both admit a *gauge freedom* that corresponds to potential vorticity conservation. This gauge freedom helps us to solve the fluid equations. In other words, a physically undetectable change in the description of the system—gauge transformation of the potentials or relabeling of the fluid particles—can lead to physically interesting solutions. The following chapter offers a stunning example.

Shallow-water dynamics with point vortices is of limited interest. Fortunately, this can all be extended to three-dimensional, rotating, stratified, hydrostatic Boussinesq dynamics with continuously distributed vorticity. These generalizations cover territory of considerable interest to oceanographers and meteorologists. But what about the *most general* case of a three-dimensional, non-hydrostatic, perfect fluid, stratified or not? Is the general case analogous to electrodynamics?

Beautiful variational principles cover the general three-dimensional case, but none of them resemble electrodynamics, and it seems unlikely that the analogy between electrodynamics and fluid dynamics could be extended to

the unrestricted three-dimensional case. To see why, consider three-dimensional, constant-density flow governed by

$$\mathbf{v}_t + \mathbf{v}\cdot\nabla\mathbf{v} = -\nabla p \tag{6.55}$$

$$\nabla\cdot\mathbf{v} = 0 \tag{6.56}$$

For such flow the simplest vorticity structures are *loops*, vortex tubes of infinitesimal thickness. On the other hand, charged particles like the electron really are *point* particles. Apart perhaps from Hill's vortex, there is nothing resembling point vortices in three dimensional fluids.

In a sense, two-dimensional fluids and three-dimensional fluids live in separate worlds. They obey similar momentum equations, but they behave quite differently. In two dimensions vorticity looks like particles (although the 'particles' could be regarded as the projections of three-dimensional vortex lines). In three dimensions vorticity looks like 'strings' that either form a loop or terminate at boundaries. In three dimensional electrodynamics you still have point particles; strings might exist, but only in spaces of much higher dimension, such as, maybe, 26.

The limit $B_x - A_y \to 1$, in which the denominators in (6.42) disappear, corresponds to a flat fluid surface. Thus infinitesimal fluid waves behave like electromagnetic waves; in the absence of vorticity they do not interact. For small amplitude shallow water waves, we set $A = A'$ and $B = x + B'$, and regard A' and B' as small. Then

$$\frac{h}{h_0} = 1 + B'_x - A'_y \tag{6.57}$$

is nearly constant. Expanding

$$\frac{1}{(B_x - A_y)} = 1 - (B'_x - A'_y) + (B'_x - A'_y)^2 + \cdots \tag{6.58}$$

we obtain the leading order nonlinear approximation to (6.42) in the form

$$L_1^{sw} = \frac{1}{2}\iiint dtdxdy\left(\left[(\psi_x + A'_t)^2 + (\psi_y + B'_t)^2\right](1 - B'_x + A'_y) - c^2(B'_x - A'_y)^2\right) \tag{6.59}$$

The cubic terms in (6.59) give rise to wave-wave interactions.

In 1936 Heisenberg and Euler computed the analogous, leading order, quantum correction to the Lagrangian for classical electrodynamics. It is

$$L_1 = \frac{1}{2}\iiint dtd\mathbf{x}\left(\mathbf{E}\cdot\mathbf{E} - c^2\mathbf{B}\cdot\mathbf{B}\right) + \epsilon\left((\mathbf{E}\cdot\mathbf{E} - c^2\mathbf{B}\cdot\mathbf{B})^2 + 7(\mathbf{E}\cdot c\mathbf{B})^2\right) \tag{6.60}$$

where ϵ is a small parameter, inversely proportional to the fourth power of the mass of the electron (the lightest charged particle). The Lagrangian (6.60) is what physicists call an *effective Lagrangian*. Although you use it as you would use the classical Lagrangian (6.8), it incorporates quantum effects at lowest order. In two dimensions (6.60) takes the form

$$L_1 = \frac{1}{2}\iiint dtdxdy\left[(E_1^2 + E_2^2 - c^2B_3^2) + \epsilon(E_1^2 + E_2^2 - c^2B_3^2)^2\right] \tag{6.61}$$

We compare the electrodynamic Lagrangian (6.61) to the shallow water Lagrangian (6.59) by writing the latter in terms of $\mathbf{E}$ and $\mathbf{B}$:

$$L_1^{sw} = \frac{1}{2}\iiint dtdxdy\left((E_1^2 + E_2^2)(1 - B_3') - c^2B_3'^2\right) \tag{6.62}$$

We see that the quantum correction to the Lagrangian for classical electrodynamics is quartic, whereas the nonlinear correction to the Lagrangian for linear shallow water dynamics is cubic. The analogy between the two corrections therefore appears to be weak. However, the analogy improves if we carry things another step, comparing the shallow water correction to the *field* quantization of (6.61), in which the electromagnetic fields themselves become particles. In the particle realm, the quartic correction in (6.61) corresponds to the interaction between 4 photons shown in figure 6.1. This diagram results from the combination of 4 simpler diagrams of the form shown in figure 6.2.

Figure 6.2 is the basic vertex of quantum electrodynamics. In figure 6.2 a photon *temporarily* converts itself into an electron and a positron. The cubic correction in (6.59) corresponds to a basic vertex of this type. The photon corresponds to a gravity wave packet, and the electron/positron pair corresponds to the pair of counter-rotating vortices in Bretherton flow. Compare figure 2.1.

Figure 6.2 cannot represent a complete, beginning-to-end, process in quantum electrodynamics, because the permanent conversion of a photon

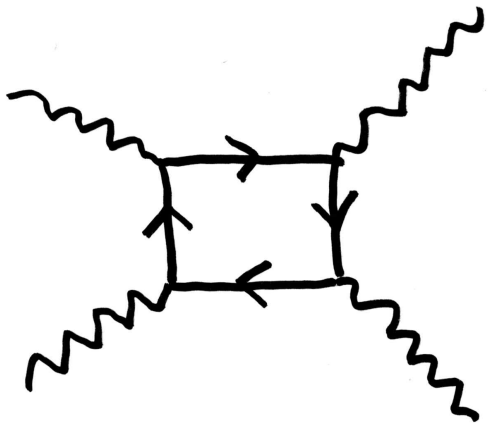

Figure 6.1: The simplest beginning-to-end interaction between photons in a vacuum. Two photons go in, and two photons go out.

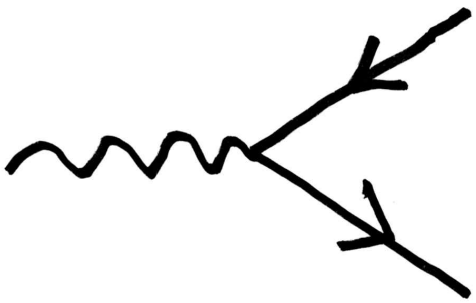

Figure 6.2: The basic vertex of quantum electrodynamics. A photon produces a virtual electron/positron pair. Energy-momentum conservation prevents the pair from separating from the photon. OR: A gravity wave packet produces counter-rotating virtual vortices (Bretherton flow). Potential vorticity conservation prevents the virtual vortices from separating from the wave packet. However, if wave breaking occurs, the virtual vortices become real, actual vorticity.

into an electron/positron pair would badly violate energy-momentum conservation: Because electrons and positrons have mass, it would be possible to find a reference frame in which the final momentum vanishes (particles moving in opposite directions at the same speed). But in this same reference frame the initial momentum could not vanish, because photons can never be brought to rest.

Likewise, potential vorticity conservation forbids the virtual vorticity in Bretherton flow from becoming real, actual vorticity *unless* dissipation (wave breaking) acts. Dissipation is a real feature of fluid dynamics, but it has no analogue in quantum electrodynamics. Thus the basic vertex of figure 6.2 describes the creation of vorticity by wave breaking in a Bretherton dipole. In contrast, figure 6.1 is the simplest possible beginning-to-end interaction between photons. Two photons go in, two photons go out, and the leptons rattling around inside the diagram are mere aids to the computation.

Does this somewhat fanciful analogy add anything to geophysical fluid dynamics? I think it does. Just as electrons and photons represent the fundamental particles of quantum electrodynamics, wave packets and point vortices constitute the fundamental particles of gfd. It is pointless to seek anything more basic. The charged particles, the vortices, are slow in the sense that they move at the typically slow speed of the fluid. The wave packets move at the typically much larger group speed. Wave/mean theory is meaningful when these two speeds are well separated. To be sure, the vortices also experience a fast *Zitterbewegung* [trembling motion] caused by passing waves, but this trembling motion does not add appreciably to their displacement. Vorticity, it seems, is the true slow variable of fluid mechanics. Instead of wave/mean theory, we should speak of wave/vorticity theory.

Vorticity is defined without reference to averaging, and this begs the question: Does averaging play a useful role in fluid dynamics? The concept of averaging entered our field in a big way with the work of Osborne Reynolds. Not everyone embraced it initially. (See for example George Stokes's uncomprehending review of a Reynolds paper in the book *A Voyage Through Turbulence*, p. 26.) Although the importance of averaging to turbulence theory has now long been taken for granted, the mathematical framework erected by Taylor, Batchelor, and Chandrasekhar has not fulfilled its high initial promise. The closure problem of turbulence remains largely unsolved.

We end this chapter by showing how the Coulomb gauge facilitates the calculation of Bretherton flow. If the potentials satisfy $A_x + B_y = 0$, then $A = \gamma_y$ and $B = -\gamma_x$ for some $\gamma(x, y, t)$. The representation (6.36)-(6.38)

becomes

$$\hat{h} = -\nabla^2\gamma \tag{6.63}$$

$$\hat{h}u = -\psi_y + \gamma_{tx} \tag{6.64}$$

$$\hat{h}v = \psi_x + \gamma_{ty} \tag{6.65}$$

Thus ψ is the stream function and γ_t is the velocity potential for the momentum $\hat{h}(u,v)$. In the limit of weak nonlinearity, the waves-only part of the Lagrangian becomes

$$\begin{aligned} L_1^{sw}[\psi,\gamma] =& \frac{1}{2}\iiint dtdxdy\left(\nabla\psi\cdot\nabla\psi + \nabla\gamma_t\cdot\nabla\gamma_t - c^2(\nabla^2\gamma)^2\right) \\ &+ \frac{1}{2}\iiint dtdxdy\left(\nabla\psi\cdot\nabla\psi + 2[\psi,\gamma_t] + \nabla\gamma_t\cdot\nabla\gamma_t\right)\nabla^2\gamma \end{aligned} \tag{6.66}$$

where the first line contains all the quadratic terms and the second line contains all the cubic terms. (6.66) is the Lagrangian for shallow water dynamics in which the vorticity vanishes (fluid initially at rest) and the gravity waves are assumed to be small in amplitude. Linear waves, which are governed by the first line of (6.66), obey

$$\delta\psi: \quad \nabla^2\psi = 0 \tag{6.67}$$

and

$$\delta\gamma: \quad \nabla^2\gamma_{tt} = c^2\nabla^2\nabla^2\gamma \tag{6.68}$$

and are thus confined to the variable γ. Setting $\gamma = \bar{\gamma} + \gamma'$ and $\psi = \bar{\psi} + 0$, we rewrite (6.66) as

$$\begin{aligned} L_1[\psi,\gamma] =& \frac{1}{2}\iiint dtdxdy\left(\nabla\gamma'_t\cdot\nabla\gamma'_t - c^2(\nabla^2\gamma')^2\right) \\ &\frac{1}{2}\iiint dtdxdy\left(\nabla\bar{\psi}\cdot\nabla\bar{\psi} + \nabla\bar{\gamma}_t\cdot\nabla\bar{\gamma}_t - c^2(\nabla^2\bar{\gamma})^2\right) \\ &+ \frac{1}{2}\iiint dtdxdy\left(\nabla\bar{\psi}\cdot\nabla\bar{\psi} + 2[\bar{\psi},\gamma_t] + \nabla\gamma_t\cdot\nabla\gamma_t\right)\nabla^2\gamma \end{aligned} \tag{6.69}$$

Under the usual scaling assumptions that $\gamma' = O(a)$, $\bar{\psi} = O(a^2)$, and $\bar{\gamma} = O(a^2)$, the first line in (6.69) is $O(a^2)$, the second line is $O(a^4)$, and we should keep only $O(a^4)$ terms in the third line. To prevent the wave from exciting

other waves (i.e. from generating a wake), we set $\bar{\gamma} = 0$. As emphasized in earlier chapters, this is a separate assumption that holds only if the envelope of γ' is very smooth. Under these assumptions (6.69) simplifies to

$$\begin{aligned} L_1[\psi,\gamma] =& \frac{1}{2}\iiint dtdxdy \left(\nabla\gamma'_t\cdot\nabla\gamma'_t - c^2(\nabla^2\gamma')^2\right) \\ & \frac{1}{2}\iiint dtdxdy \left(\nabla\bar{\psi}\cdot\nabla\bar{\psi} + 2[\bar{\psi},\gamma'_t]\nabla^2\gamma'\right) \end{aligned} \tag{6.70}$$

The rest of the calculation is standard: We set $\gamma' = A\cos\phi$ and apply Whitham's method to get the equation for $\bar{\psi}$, the Bretherton flow.

Chapter 7

The gauge freedom of fluid mechanics

Lagrangian coordinates have a bad reputation. The case against them goes like this: In seeking solutions to the Lagrangian equations, one sets oneself the task of solving for the trajectory of every fluid particle. However, Eulerian solutions are known in which the velocity field is a simple one, but the particle trajectories are chaotic and therefore impossible to determine analytically. Since the Lagrangian method requires you to find the fluid particle trajectories, it would have failed to find these solutions. Ergo, the Eulerian method succeeds in cases for which the Lagrangian method would have failed. Lagrangian fluid mechanics is a step too far.

This is a bad rap, and it is easy to rebut. A simple argument shows that there are solutions to the Lagrangian equations that the Eulerian method would fail to find. You probably know one already. It is the Gerstner wave, which is among the earliest solutions to the fluid equations (1802). Fluid mechanicists view the Gerstner solution as *parametric*, but it is unabashedly Lagrangian.

This chapter focuses on two-dimensional incompressible flow governed by

the Eulerian equations in the form

$$\left(\frac{\partial}{\partial t}+u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}\right)u=-\frac{\partial p}{\partial x} \tag{7.1}$$

$$\left(\frac{\partial}{\partial t}+u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}\right)v=-\frac{\partial p}{\partial y}-g \tag{7.2}$$

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \tag{7.3}$$

and by the Lagrangian equations in the form

$$\frac{\partial^2 x}{\partial \tau^2}=-\frac{\partial(p,y)}{\partial(a,b)} \tag{7.4}$$

$$\frac{\partial^2 y}{\partial \tau^2}=-\frac{\partial(x,p)}{\partial(a,b)}-g \tag{7.5}$$

$$\frac{\partial(x,y)}{\partial(a,b)}=1 \tag{7.6}$$

You first saw (7.4)-(7.6) in Chapter 1. We hold two interpretations in mind. Either x and y are horizontal coordinates and $g=0$, or y is vertical and g is the gravity constant. In the latter case we assume that a free surface is present.

In the Eulerian approach we solve (7.1)-(7.3) for $\mathbf{u}(\mathbf{x},t)$ and $p(\mathbf{x},t)$. In the Lagrangian approach we solve (7.4)-(7.6) for $\mathbf{x}(\mathbf{a},\tau)$ and $p(\mathbf{a},\tau)$. Suppose we have done the latter. If we know $\mathbf{x}(\mathbf{a},\tau)$, then we know $\mathbf{u}(\mathbf{a},\tau)=\partial\mathbf{x}/\partial\tau$. To express our Lagrangian solution in the Eulerian form $\mathbf{u}(\mathbf{x},t)$, we must invert $\mathbf{x}(\mathbf{a},\tau)$ to get $\mathbf{a}(\mathbf{x},t)$ and substitute the latter into $\mathbf{u}(\mathbf{a},\tau)$.

In principle it is always possible to find the inverse of $\mathbf{x}(\mathbf{a},\tau)$, because the mapping from particle labels to locations is one-to-one. But what is possible in principle can be impossible in practice. If the mapping $\mathbf{a}\to\mathbf{x}$ involves transcendental functions, as in the case of the Gerstner wave, then it may be practically impossible to obtain $\mathbf{a}\leftarrow\mathbf{x}$ as needed to obtain an analytical expression for $\mathbf{x}\to\mathbf{u}$. If there is no analytical expression for $\mathbf{x}\to\mathbf{u}$, then it would have been impossible to obtain $\mathbf{u}(\mathbf{x},t)$ by solving the Eulerian equations. In summary, just as there are Eulerian solutions that cannot be written in Lagrangian form, so too there are Lagrangian solutions that cannot be written in Eulerian form.

But there is more. The Lagrangian mass-conservation equation (7.6) requires that the ratio of area in physical space to area in label space be

unity. This requirement is unnecessarily strong. Fluid incompressibility is ensured if the fluid particles corresponding to a *particular* area in label space always occupy the same *amount* of area in physical space, but the ratio can depend on location in label space. In fact, although we shall not consider such, it could even depend on time. The Lagrangian formulation only requires that you have a method for keeping track of fluid particles and that every fluid particle has one, and only one, location at any particular time.

To illustrate what we mean, suppose that we define a new set of fluid particle labels, α and β, by the equations

$$\alpha = F(a,b), \quad \beta = G(a,b) \tag{7.7}$$

where $F(a,b)$ and $G(a,b)$ are arbitrary, time-independent functions. The only rigid requirement is that

$$J \equiv \frac{\partial(a,b)}{\partial(\alpha,\beta)} \neq 0 \tag{7.8}$$

By prohibiting J from changing sign, we ensure that the particle relabeling is one-to-one. The function J, which is determined by F and G, can be considered an arbitrary sign-definite function of either (a,b) or (α,β).

Next we transform the independent variables in (7.4)-(7.6) from (a,b,τ) coordinates to (α,β,τ) coordinates. The result is

$$J(\alpha,\beta)\,\frac{\partial^2 x}{\partial\tau^2} = -\frac{\partial(p,y)}{\partial(\alpha,\beta)} \tag{7.9}$$

$$J(\alpha,\beta)\,\frac{\partial^2 y}{\partial\tau^2} = -\frac{\partial(x,p)}{\partial(\alpha,\beta)} - g\,J(\alpha,\beta) \tag{7.10}$$

$$\frac{\partial(x,y)}{\partial(\alpha,\beta)} = J(\alpha,\beta) \tag{7.11}$$

These are three equations in the three independent variables $x(\alpha,\beta,\tau)$, $y(\alpha,\beta,\tau)$, and $p(\alpha,\beta,\tau)$. Although (7.9)-(7.11) appear more complicated than (7.4)-(7.6), they are actually easier to solve, because the arbitrary function $J(\alpha,\beta)$ can be chosen to facilitate the solution. Again, the only rigid requirement is that $J(\alpha,\beta)$ not change sign.

Our first example is the Gerstner wave. To get the Gerstner wave, suppose that $J(\alpha,\beta)$ depends only on β; suppose

$$J = f'(\beta) \tag{7.12}$$

where $f(\beta)$ is a function to be determined. The labels α and β can always be re-defined to satisfy (7.12). We are merely adopting J as one of our two labels. If (7.12) holds then (7.9)-(7.11) become

$$\delta x: \qquad f'(\beta)\frac{\partial^2 x}{\partial \tau^2} = -\frac{\partial(p,y)}{\partial(\alpha,\beta)} \tag{7.13}$$

$$\delta y: \qquad f'(\beta)\frac{\partial^2 y}{\partial \tau^2} = -\frac{\partial(x,p)}{\partial(\alpha,\beta)} - f'(\beta)g \tag{7.14}$$

$$\delta p: \qquad \frac{\partial(x,y)}{\partial(\alpha,\beta)} = f'(\beta) \tag{7.15}$$

We find that

$$x = \alpha + A(\beta)\sin(k\alpha - \omega\tau) \tag{7.16}$$

$$y = \beta - A(\beta)\cos(k\alpha - \omega\tau) \tag{7.17}$$

$$p = -gf(\beta) \tag{7.18}$$

satisfy (7.13)-(7.14) for any $A(\beta)$, provided that $\omega^2 = gk$. The free surface corresponds to $f(\beta) = 0$. Eqn (7.18) looks like hydrostatic balance, but remember that β can be considered a function of (x, y, t).

Substituting (7.16) and (7.17) into (7.15) we obtain

$$\frac{\partial(x,y)}{\partial(\alpha,\beta)} = 1 + [kA(\beta) - A'(\beta)]\ \cos(k\alpha - \omega\tau) - kA(\beta)A'(\beta) = f'(\beta) \tag{7.19}$$

To eliminate the α, τ-dependence on the left hand side, we require $kA(\beta) - A'(\beta) = 0$, which in turn implies

$$A(\beta) = A_0 e^{k\beta} \tag{7.20}$$

where A_0 is a constant. Then (7.19) becomes

$$\frac{\partial(x,y)}{\partial(\alpha,\beta)} = 1 - k^2 A_0^2 e^{2k\beta} = f'(\beta) \tag{7.21}$$

To keep (7.21) positive we set $A_0 = k^{-1}$ and confine the fluid to the region of negative β. Then (7.21) implies

$$f(\beta) = B + \beta - \frac{1}{2k}e^{2k\beta} \tag{7.22}$$

where B is an arbitrary constant. Putting all this together, we obtain the fluid motion in the form

$$x = \alpha + \frac{e^{k\beta}}{k}\sin(k\alpha - \omega\tau) \tag{7.23}$$

$$y = \beta - \frac{e^{k\beta}}{k}\cos(k\alpha - \omega\tau) \tag{7.24}$$

$$p = -g\left(B + \beta - \frac{1}{2k}e^{2k\beta}\right) \tag{7.25}$$

This is the Gerstner wave. It is a separate solution for every value of B. To obtain the equation for the free surface, we set $p = 0$ to obtain the corresponding value of β. However, since B is arbitrary, this amounts to choosing an arbitrary negative value, β_s, for β. The fluid then occupies $-\infty < \alpha < +\infty$ and $\beta < \beta_s$.

The parametric equations for the free surface are (7.23) and (7.24) with α the parameter, and $\beta = \beta_s$. To get the waveform we set $\tau = 0$ and $\theta = k\alpha$ to obtain

$$x = \theta/k + \frac{e^{k\beta_s}}{k}\sin(\theta) \tag{7.26}$$

$$y = \beta_s - \frac{e^{k\beta_s}}{k}\cos(\theta) \tag{7.27}$$

where now θ is the parameter, and the constant β_s can take any negative value. The cycloid corresponds to $\beta_s = 0$ and represents the Gerstner wave of maximum amplitude.

In 1984 Anatoly A. Abrashkin and Evsei I. Yakubovich showed that the Gerstner wave is but the tip of a very large iceberg; it belongs to an infinite family of Lagrangian solutions. To obtain their result, we first rewrite (7.9) and (7.10) in the form

$$J(\alpha,\beta)\; x_{\tau\tau} = -p_\alpha y_\beta + p_\beta y_\alpha \tag{7.28}$$

$$J(\alpha,\beta)\; y_{\tau\tau} = p_\alpha x_\beta - p_\beta x_\alpha - g\; J(\alpha,\beta) \tag{7.29}$$

Next we guess a form for $x(\alpha,\beta,\tau)$ and $y(\alpha,\beta,\tau)$ that contains adjustable functions and parameters. Eqns (7.9)-(7.11) are solved if we can adjust those functions and parameters so that:

1. The Jacobian in (7.11) is time-independent and does not change sign.

2. A pressure field $p(\alpha, \beta, \tau)$ can be found that satisfies (7.28) and (7.29).

For the given $x(\alpha, \beta, \tau)$ and $y(\alpha, \beta, \tau)$, we solve (7.28) and (7.29) for the derivatives of the pressure field. We find that

$$p_\alpha = -x_\alpha x_{\tau\tau} - y_\alpha(y_{\tau\tau} + g) \tag{7.30}$$

$$p_\beta = -x_\beta x_{\tau\tau} - y_\beta(y_{\tau\tau} + g) \tag{7.31}$$

If these derivatives obey the consistency condition

$$(p_\alpha)_\beta = (p_\beta)_\alpha \tag{7.32}$$

then the pressure field is determined to within an irrelevant constant. Substituting (7.30) and (7.31) into (7.32) we obtain the consistency requirement

$$\frac{\partial}{\partial \tau}\left(\frac{\partial(x_\tau, x)}{\partial(\alpha, \beta)} + \frac{\partial(y_\tau, y)}{\partial(\alpha, \beta)}\right) = 0 \tag{7.33}$$

which integrates to

$$\frac{\partial(x_\tau, x)}{\partial(\alpha, \beta)} + \frac{\partial(y_\tau, y)}{\partial(\alpha, \beta)} = J(\alpha, \beta)\zeta(\alpha, \beta) \tag{7.34}$$

where $\zeta(\alpha, \beta)$ is an arbitrary τ-independent function that may take either sign. Dividing (7.34) by

$$\frac{\partial(x, y)}{\partial(\alpha, \beta)} = J(\alpha, \beta) \tag{7.35}$$

we see that $\zeta(\alpha, \beta)$ is the vorticity. Thus (7.33) is the vorticity equation in Lagrangian coordinates.

The Lagrangian equations are solved if the adjustable functions and parameters in $x(\alpha, \beta, \tau)$ and $y(\alpha, \beta, \tau)$ can be made to satisfy (7.34) and (7.35), where $J(\alpha, \beta)$ is an arbitrary, τ-independent, sign-definite function; and $\zeta(\alpha, \beta)$ is arbitrary and τ-independent, but not necessarily sign-definite. If these criteria are met, the pressure field exists and may be determined from (7.30)-(7.31), but it is not actually necessary to know the pressure field unless it is needed to satisfy a boundary condition. This would be the case if a free surface were present.

The challenge is to choose the adjustable functions and parameters in $x(\alpha, \beta, \tau)$ and $y(\alpha, \beta, \tau)$ in such a way that the left-hand sides of (7.34) and

(7.35) can be made τ-independent. Adopting complex notation, Abrashkin and Yakubovich postulated that

$$z = s + A(\alpha, \beta)e^{i\omega\tau} \tag{7.36}$$

where

$$z = x + iy \quad \text{and} \quad s = \alpha + i\beta, \tag{7.37}$$

$A(\alpha, \beta)$ is complex valued function to be determined, and ω is a real constant to be determined.[1] By direct substitution, they found that the left hand sides of (7.34) and (7.35) are τ-independent if $A(\alpha, \beta)$ satisfies

$$\frac{\partial A}{\partial \alpha} = i\frac{\partial A}{\partial \beta} \tag{7.38}$$

Eqn (7.38) is satisfied if $A(\alpha, \beta) = f(\alpha - i\beta)$, where $f(\cdot)$ is an arbitrary analytic function. Thus

$$z = s + f(s^*)e^{i\omega\tau} \tag{7.39}$$

satisfies the Lagrangian fluid equations, where $f(w)$ is an arbitrary analytic function, and s^* denotes the complex conjugate of s. The fluid particle labeled by s moves in a circle of radius $|f(s^*)|$ centered at the point $z = s$. Thus fluid particles are labeled by the centers of their circular trajectories. From (7.34) and (7.35) we obtain

$$J(\alpha, \beta) = 1 - |f'(s^*)|^2 \tag{7.40}$$

and

$$\zeta(\alpha, \beta) = \frac{-2\omega|f'(s^*)|^2}{1 - |f'(s^*)|^2} \tag{7.41}$$

The Gerstner wave corresponds to $f(w) = e^{Cw}$, where C is a complex constant whose real and imaginary parts determine the amplitude and wavelength of the wave.

We shall investigate the monomials,

$$f(w) = \frac{C}{n}w^n \tag{7.42}$$

[1] The form considered by A&Y is actually more general than (7.36). It involves two frequencies.

where C is a real constant and n is an integer. The fluid motion is given by (7.39) in the form

$$z = s + \frac{C}{n}(s^*)^n e^{i\omega\tau} \tag{7.43}$$

The state of rest corresponds to $C = 0$. The vorticity is

$$\zeta(\alpha, \beta) = \frac{-2\omega C^2(\alpha^2 + \beta^2)^{n-1}}{1 - C^2(\alpha^2 + \beta^2)^{n-1}} \tag{7.44}$$

and

$$J(\alpha, \beta) = 1 - C^2(\alpha^2 + \beta^2)^{n-1} \tag{7.45}$$

We assume that the fluid covers the entire z-plane. However, unless $n = 1$, the *domain* of the labels cannot be the entire s-plane, because that would violate $J > 0$. To avoid this we assume that the labels are confined to the interior of the unit circle in the s-plane if $n \geq 1$, and to the exterior of the s-plane unit circle if $n < 1$. In either case the *range* of the labels does not cover the entire z-plane, and we must take care that the fluid equations are also satisfied in the excluded region.

Suppose $n > 1$. The label domain is

$$\alpha^2 + \beta^2 = |s|^2 < 1 \tag{7.46}$$

To keep $J > 0$, we must choose $C < 1$. Eqn (7.43) maps the label domain (7.46) to a region of physical space whose boundary is the image of

$$|s| = 1 \tag{7.47}$$

To get an equation for this boundary in z-space, we set $z = re^{i\theta}$ and $s = e^{i\mu}$ in (7.43). Each value of μ corresponds to a fluid particle on the boundary (7.47) of the region (7.46). The result,

$$re^{i\theta} = e^{i\mu} + \frac{C}{n}e^{-i\mu n + i\omega\tau} \tag{7.48}$$

is a parametric expression, $r(\mu)$ and $\theta(\mu)$, for the bounding curve (7.47). It follows from (7.48) that

$$r^2 = 1 + \frac{C^2}{n^2} + 2\frac{C}{n}\cos\left((n+1)\mu - \omega\tau\right), \quad 0 < \mu < 2\pi \tag{7.49}$$

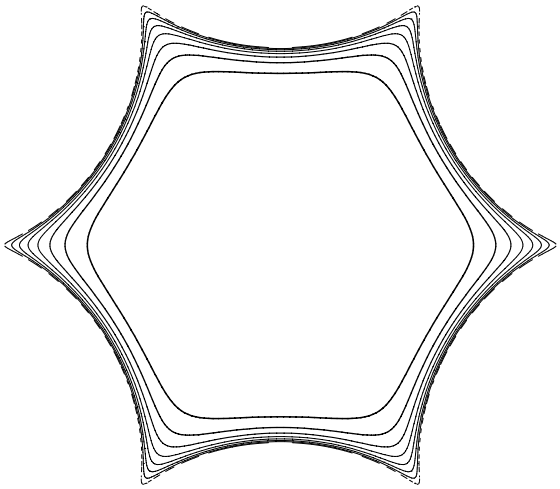

Figure 7.1: Vorticity contours in (x, y)-space for the solution (7.43) with $n = 5$ and $C = 0.9$.

Thus the bounding curve has $n+1$ 'arms'. This bounding curve, and indeed the entire vorticity field (7.44), rotates at angular speed $\omega/(n+1)$ in z-space. The velocity field rotates at this same angular speed, but each fluid particle moves in a perfect circle at angular speed ω around a point that is fixed in z-space and unique to that particle. Figure 7.1 shows the vorticity in z-space—physical space—for the case $n = 5$ and $C = 0.9$. The contours are lines of constant vorticity. The vorticity is zero at the center of the figure, and achieves its greatest magnitude on the outermost (dashed) contour. The *sign* of the vorticity is the same throughout the domain (7.46), and depends on the sign of ω.

It remains to determine the fluid motion in the region outside the bounding curve (7.47). The easiest recourse is to assume that the flow in this exterior region is irrotational. Then the exterior flow satisfies $\nabla^2\psi = 0$, where ψ is the stream function. Although the vorticity is discontinuous at the boundary curve (7.47), the velocity is not. Thus the appropriate boundary condition on ψ is the Neumann boundary condition that the normal derivative ψ_n at the boundary curve be equal to the tangential velocity of the solution (7.43). This determines ψ at every time, and completes the solution on the infinite z-plane.

The case $n = 1$ is exceptional in that the vorticity within the bounding

curve is uniform. This is the solution discovered by Kirchoff in 1876.

The case $n < 1$ proceeds in a similar manner. The solution (7.43) applies in the region exterior to the unit circle (7.47). In the finite region interior to this boundary, it is convenient to assume that the vorticity takes a constant value Γ, which is determined by the consistency condition

$$\oint dl \, \frac{\partial \psi}{\partial n} = \iint dx dy \, \Gamma \tag{7.50}$$

on the normal derivative of ψ. The integrals in (7.50) are around and over the region of the xy-plane corresponding to $\alpha^2 + \beta^2 < 1$.

In the next most complicated case beyond the two we have considered, the labels could be constrained to occupy the annulus

$$1/2 < \alpha^2 + \beta^2 < 1 \tag{7.51}$$

so that there is *both* an exterior region of irrotational flow and an interior region of uniform vorticity. This would be necessary if the arbitrary analytic function $f(w)$ had singularities at both $w = 0$ and $w = \infty$, but it could be done in any case. Note that the precise limits in (7.51) are irrelevant because the values taken by α and β can always be adjusted. Topology is what matters. And the topology could be further generalized to include an arbitrary number of 'holes' in addition to the one at the origin. The result would be a vast Swiss cheese, with holes filled with fluid of uniform vorticity. Could such solutions resemble two-dimensional turbulence, with its isolated coherent vortices?

What are the strange waves that propagate, as in (7.49), around the boundaries of the vortical regions? They are Rossby waves, of course. Moreover, they are Rossby waves living free and wild in their native habitat. These waves are only distantly related to the tame, zoo-animal Rossby waves that depend for their existence on the external parameter named β (which is unrelated to our particle label of the same name). Meteorologists recognized them long ago when they replaced β with $\beta - \bar{u}_{yy}$, where $\bar{u}(y)$ is the 'zonal mean flow'. But it is not only the longitude-averaged latitudinal velocity that posesses a vorticity gradient. Two-dimensional turbulence is *full* of *local* vorticity gradients that point in every direction. Disturbances—particle displacements—propagate at right angles to these gradients at speeds that sometimes exceed the speed of fluid particles—and sometimes do not. In two-dimensional turbulence there is no clear separation between such 'Rossby waves' and the 'mean flows' that transport vorticity.

The example of two-dimensional turbulence demonstrates the severe limitations of wave/mean theory as developed in earlier chapters, and urges us to focus on a dynamics that contains the more essential difficulties of fluid dynamics—the dynamics of two-dimensional turbulence. Understand pure two-dimensional turbulence and you will understand the behaviors of many supposedly 'more complicated' systems. The 'complications' are mere distractions from what should be the main event.

And so, the long good-bye. We are finished with the gravity waves. We wish them well as they and their many admirers leave our stage. Beguiling to some, they hold no fascination for us.

Ours is the harder path. We focus on two-dimensional turbulence, with physics stated most succinctly by the vorticity equation

$$\frac{\partial \zeta}{\partial t} + \frac{\partial(\psi, \zeta)}{\partial(x, y)} = 0 \tag{7.52}$$

where

$$\nabla^2 \psi = \zeta \tag{7.53}$$

Viscosity is missing, and we must account for it eventually. But (7.52) is a challenge by itself. Without the viscosity, there is no such thing as a small-amplitude solution to (7.52). Small compared to what? By re-scaling the time, small amplitudes are made large. Perturbation theory seems useless. What other tricks do we have?

The present chapter contains a flicker of hope. Eqns (7.52) and (7.53) are physically, but not mathematically, equivalent to (7.34) and (7.35). To solve the former we seek $\psi(x, y, t)$. To solve the latter we seek $x(\alpha, \beta, \tau)$ and $y(\alpha, \beta, \tau)$. The advantage of (7.34)-(7.35) is that $J(\alpha, \beta)$ can be any sign-definite function that makes our job easier. This is a consequence of our ability to label the fluid particles in almost any way we want. In other words, it is a consequence of the gauge freedom of fluid mechanics.

Local gauge invariance is the underlying principle of modern physics. Fluid particle relabeling is a local gauge invariance. It is a *gauge* invariance because re-naming the fluid particles makes no difference to how they will behave. It is *local* because the re-naming varies from one fluid particle to the next.

The Eulerian solution $\psi(x, y, t)$ is a three-dimensional manifold in the four-dimensional space spanned by $\{x, y, t, \psi\}$. The Lagrangian solution $x(\alpha, \beta, \tau)$ and $y(\alpha, \beta, \tau)$ is a three-dimensional manifold in the *five*-dimensional

space spanned by $\{\alpha, \beta, \tau, x, y\}$. The extra dimension seems to make a lot of difference. But if a little bit of gauge freedom is this good, what would a lot of it be like? Could fluid dynamics, even turbulence, appear simple when viewed in a space of (say) 26 dimensions?