Hamiltonian Fluid Dynamics

In this last chapter we return to the subject of the first: the fundamental principles of fluid mechanics. In Chapter 1, we derived the equations of fluid motion from Hamilton’s principle of stationary action, emphasizing its logical simplicity and the resulting close correspondence between mechanics and thermodynamics. Now we explore the Hamiltonian approach more fully, discovering its other advantages.¹ The most important of these advantages arise from the correspondence between the symmetry properties of the Lagrangian and the conservation laws of the resulting dynamical equations. We therefore begin with a very brief introduction to symmetry and conservation laws.

1. Symmetry and conservation laws

Noether’s theorem applies to the equations that arise from variational principles like Hamilton’s principle. According to Noether’s theorem: If a variational principle is invariant to a continuous transformation of its dependent and independent variables, then the equations arising from the variational principle possess a divergence-form conservation law. The invariance property is also called a symmetry property. Thus Noether’s theorem connects symmetry properties and conservation laws.

We shall neither state nor prove the general form of Noether’s theorem; to do so would require a rather lengthy digression on continuous groups.² Instead we illustrate the connection between symmetry and conservation laws with a series of increasingly complex and important examples. These examples convey the flavor of the general theory.

Our first example is very simple. Consider a body of mass \( m \) moving in one dimension. The body is attached to the end of a spring with spring-constant \( K \). Let \( x(t) \) be the displacement of the body from its location when the spring is unstretched. Then

\[
\frac{d^2 x}{dt^2} = -\omega_0^2 x, \tag{1.1}
\]

where \( \omega_0^2 = K/m \) is the frequency of the resulting motion. Equation (1.1) conserves energy,

\[
\frac{d}{dt} \left\{ \frac{1}{2} \left( \frac{dx}{dt} \right)^2 + \frac{1}{2} \omega_0^2 x^2 \right\} = 0. \tag{1.2}
\]

Now suppose that the spring-constant \( K \) varies with time in a prescribed manner. The motion is still governed by (1.1), but the coefficient \( \omega_0(t) \) now depends on time. Instead of (1.2) we obtain
\[
\frac{d}{dt} \left\{ \frac{1}{2} \left( \frac{dx}{dt} \right)^2 + \frac{1}{2} \omega_0^2 x^2 \right\} = x^2 \omega_0 \frac{d\omega_0}{dt}.
\] (1.3)

Energy is no longer conserved.

Whether \( \omega_0 \) depends on time or not, the dynamics (1.1) is equivalent to the variational principle

\[
\delta \int_{t_1}^{t_2} L \left( x, \frac{dx}{dt}, t \right) dt = 0,
\] (1.4)

where

\[
L \left( x, \frac{dx}{dt}, t \right) = \frac{1}{2} \left( \frac{dx}{dt} \right)^2 - \omega_0^2 (t) x^2.
\] (1.5)

Here, \( \delta \) denotes the change arising from an infinitesimal variation \( \delta x(t) \) in the location of the body. This variation is arbitrary, but it must vanish at the endpoints in time, \( \delta x(t_1) = \delta x(t_2) = 0 \). The \( t \)-argument of \( L \) corresponds to the time-dependence of \( \omega_0(t) \). For further background, see Chapter 1.

In the case of a constant \( \omega_0 \), energy conservation (1.2) corresponds to the symmetry property that (1.5) is invariant to a uniform translation, \( t \to t + \text{const} \), in the dependent variable. That is, energy conservation corresponds to the time-translation symmetry of the Lagrangian.

To obtain the energy conservation law in its differential form, we consider an infinitesimal change in the time variable, from old time \( t \) to new time

\[
t' = t + \delta t(t),
\] (1.6)

where \( \delta t(t) \) is an infinitesimal function that vanishes at \( t_1 \) and \( t_2 \). Consider two realizations of \( x(t) \). Let the second realization have the same dependence on \( t' \) that the first realization has on \( t \). Thus the two realizations differ only in the time values assigned to corresponding events. If the first realization is \( x = f(t) \), then the second realization is \( x = f(t + \delta t) \). Since \( \omega_0 \) is a constant, the action corresponding to the first realization is

\[
A_i = \int_{t_1}^{t_2} dt \left[ f(t), \frac{d}{dt} f(t) \right].
\] (1.7)

The action corresponding to the second realization is
\[ A_2 = \int_{t_1}^{t_2} dt \int f(t + \delta t) \, \frac{d}{dt} f(t + \delta t) \]

\[ = \int_{t_1}^{t_2} dt \int f(t'), \frac{d}{dt} f(t') = \int_{t_1}^{t_2} dt' \frac{d}{dt} \int f(t'), \frac{d}{dt} \frac{d}{dt} f(t') \]

\[ \int_{t_1}^{t_2} dt' \left( 1 - \frac{d}{dt} \delta t \right) L \left[ f(t'), \left( 1 + \frac{d}{dt} \delta t \right) \frac{d}{dt} f(t') \right] + O(\delta t^2) \]

The integration limits remain fixed because \( t = t' \) at the endpoints. In the final integral, we can replace \( t \) by \( t' \) in the \( O(\delta t) \) terms, and then drop the primes altogether. Then the difference between \( A_1 \) and \( A_2 \) is clearly

\[ \int_{t_1}^{t_2} dt \left[ L - \frac{\partial L}{\partial (dx/dt)} \frac{dx}{dt} \right] \frac{d}{dt} + O(\delta t^2) \]

\[ = \int_{t_1}^{t_2} dt \delta(t) \left[ \frac{\partial L}{\partial (dx/dt)} \frac{dx}{dt} - L \right] + O(\delta t^2) \]

By the variational principle, (1.9) must vanish. Then, since \( \delta(t) \) is arbitrary, we must have

\[ \frac{d}{dt} \left[ \frac{\partial L}{\partial (dx/dt)} \frac{dx}{dt} - L \right] = 0 \]

(1.10)

This is a general statement of energy conservation; with (1.5) it implies (1.2).

If \( \omega_0(t) \) depends on time in a prescribed manner, then (1.5) loses the time-translation symmetry and the corresponding conservation law (1.10). Instead of (1.10), the variation \( \delta(t) \) produces

\[ \frac{d}{dt} \left[ \frac{\partial L}{\partial (dx/dt)} \frac{dx}{dt} - L \right] = -\frac{\partial L}{\partial t} \]

(1.11)

which, with (1.5), implies (1.3).

Now suppose that \( \omega_0(t) \) changes slowly with time, on a time-scale \( T \) that is much longer than \( \omega_0^{-1} \). There is still no time-translation symmetry, but, with a suitable approximation, another conservation law appears.

In the case of constant \( \omega_0 \), the solution of (1.1) is \( x = A \cos(\omega_0 t + \alpha) \), where \( A \) and \( \alpha \) are constants. This suggests that, in the case of slowly varying \( \omega_0(t) \), we may expect solutions of the form

\[ x(t) = A(t) \cos(\theta(t)) \]

(1.12)
where $A(t)$ and 

$$\omega(t) \equiv \frac{d\theta}{dt}$$

are \textit{slowly varying} functions of time. Since $A$ varies slowly compared to $\theta$, 

$$\frac{dx}{dt} \approx -A(t)\frac{d\theta}{dt}\sin\theta(t).$$

(1.14)

Thus, substituting (1.12) and (1.14) into (1.5) we obtain the approximation,

$$L \approx \frac{1}{2}A^2\left(\frac{d\theta}{dt}\right)^2\sin^2\theta - \omega_0^2(t)A^2\cos^2\theta.$$

(1.15)

The variational principle takes the form,

$$\delta \int_{t_1}^{t_2} \frac{1}{2}A^2\left(\frac{d\theta}{dt}\right)^2\sin^2\theta - \omega_0^2(t)A^2\cos^2\theta = 0,$$

(1.16)

for arbitrary $\delta A(t)$ and $\delta \theta(t)$. But if $A$, $\omega_0$, and $d\theta/dt$ all vary slowly compared to $\theta$, then we can accurately replace the trigonometric terms in (1.16) by their averages over a wave-cycle. We obtain

$$\delta \int_{t_1}^{t_2} \frac{1}{2}A^2\left(\frac{d\theta}{dt}\right)^2 - \omega_0^2(t)A^2 = 0.$$  

(1.17)

The integrand of (1.17) is called the \textit{averaged Lagrangian}.

The variational principle (1.17) holds for arbitrary, independent variations of $A(t)$ and $\theta(t)$. But first we note that (1.17) possesses a symmetry property that (1.16) does not, namely a translation-symmetry for $\theta$. This symmetry arises from the approximation leading from (1.16) to (1.17), that is, from the replacement of the rapidly varying terms by their averages over an oscillation period. Because of this symmetry, we expect the phase variation $\delta \theta$ to yield a conservation law. Indeed,

$$\delta \theta : \quad \frac{d}{dt}\left(\frac{d\theta}{A^2 dt}\right) = 0.$$  

(1.18)

On the other hand, the amplitude variation yields

$$\delta A : \quad \frac{d\theta}{dt} = \pm \omega_0(t).$$  

(1.19)
According to (1.19), the body oscillates at a slowly varying frequency of $\omega_0(t)$. The amplitude of the oscillation is nonconstant, but it can be determined from the conservation law (1.18). The energy averaged over an oscillation cycle is

$$E = \frac{1}{2} A^2 \omega_0^2.$$  \hspace{1cm} (1.20)

Thus, according to (1.18), although $E$, $A$ and $\omega_0$ all vary slowly with time, the action

$$W \equiv \frac{E}{\omega_0}$$  \hspace{1cm} (1.21)

is conserved.

Gerald Whitham introduced the averaged Lagrangian method and developed it into a powerful, synthetic tool for the study of linear and nonlinear waves.\(^3\) Of course, results like (1.18) can also be obtained without the use of variational methods — the reader is invited to derive results equivalent to (1.18) and (1.19) directly from (1.1), using WKB theory or the method of his choice — but the Hamiltonian approach has at least two distinct advantages. First, it is remarkably succinct. And second, it leads almost automatically, via Noether’s theorem, to the conservation laws of the system. In non-Hamiltonian approaches, these conservation laws can remain deeply hidden, and are often uncovered only after extensive, unguided manipulations. In the Hamiltonian approach, a conservation law is known to exist as soon as an inspection of the Lagrangian reveals the corresponding symmetry property. In the following sections, we shall see that the advantages of Hamiltonian methods increase dramatically as the system under consideration becomes more complex.

2. The particle-relabeling symmetry property

In the Lagrangian description of a perfect fluid, the dependent variables are the locations,

$$x(a,\tau) = (x(a,b,c,\tau), y(a,b,c,\tau), z(a,b,c,\tau))$$  \hspace{1cm} (2.1)

of the fluid particle labeled by $a=(a,b,c)$ at time $\tau$. The labels remain fixed following the fluid particle, and hence $\partial/\partial \tau \equiv D/Dt$ is the substantial derivative. It is convenient to assign the labels so that

$$da \, db \, dc = d(\text{mass}).$$  \hspace{1cm} (2.2)

Then

$$\alpha = \frac{\partial (x,y,z)}{\partial (a,b,c)}$$  \hspace{1cm} (2.3)
is the specific volume, and the \( \tau \)-derivative of (2.3) yields the continuity equation,

\[
\frac{\partial \alpha}{\partial \tau} = \alpha \nabla \cdot \mathbf{v},
\]

(2.4)

where

\[
\mathbf{v} = \left( \frac{\partial x}{\partial \tau}, \frac{\partial y}{\partial \tau}, \frac{\partial z}{\partial \tau} \right)
\]

(2.5)

is the fluid velocity and \( \nabla = (\partial_x, \partial_y, \partial_z) \) is the gradient operator in \( x \)-space. For further details, refer to Chapter 1.

Readers of Chapter 1 will recall that the Lagrangian

\[
L \left[ \mathbf{x}(\mathbf{a}, \tau) \right] = \int \int \int d\mathbf{a} \left\{ \frac{1}{2} \frac{\partial \mathbf{x}}{\partial \tau} \cdot \frac{\partial \mathbf{x}}{\partial \tau} - E \left( \frac{\partial \mathbf{x}}{\partial \mathbf{a}}, \eta(\mathbf{a}) \right) - \Phi(\mathbf{x}) \right\}
\]

(2.6)

for a perfect fluid is closely analogous to the familiar Lagrangian for a system of discrete particles. In (2.6), the internal energy \( E(\alpha, \eta) \) is a prescribed function of the specific volume \( \alpha \) and the specific entropy \( \eta \). The form of \( E(\alpha, \eta) \) determines the equilibrium thermodynamics of the fluid. \( \Phi(\mathbf{x}) \) is the potential for external forces like gravity. The entropy \( \eta \) depends only on \( \mathbf{a} \) (and not on \( \tau \)), in a manner determined by initial conditions. Hence,

\[
\frac{\partial \eta}{\partial \tau} \equiv \frac{D \eta}{D t} = 0.
\]

(2.7)

Hamilton’s principle states that

\[
\delta \int_{\tau_i}^{\tau_f} d\tau \ L \left[ \mathbf{x}(\mathbf{a}, \tau) \right] = 0
\]

(2.8)

for variations \( \delta \mathbf{x}(\mathbf{a}, \tau) \) that have no normal component at rigid boundaries but are otherwise arbitrary. This yields the momentum equation,

\[
\delta \mathbf{x} : \quad \frac{\partial^2 \mathbf{x}}{\partial \tau^2} = -\alpha \nabla p - \nabla \Phi
\]

(2.9)

with

\[
p = -\frac{\partial E(\alpha, \eta)}{\partial \alpha},
\]

(2.10)
and the correct boundary conditions at free and rigid boundaries. For complete details, refer to Chapter 1.

In this chapter, we consider variations that vanish at the extremities of the fluid, and therefore yield the dynamical equation but not its boundary conditions. This avoids some fussy calculations at no real loss. In the most frequently considered case of a fluid confined by rigid boundaries, the only boundary condition is a simple one: no flow of fluid particles through the boundary.

We begin by examining the symmetry properties of (2.6). As in the example of Section 1, the time-translation symmetry of (2.6) corresponds to the conservation of energy,

\[
\frac{d}{dt} \iiint da \left\{ \frac{\partial x}{\partial \tau} \cdot \frac{\partial x}{\partial \tau} + E + \Phi \right\} = 0.
\] (2.11)

The presence of \( \Phi(x) \) (which could incorporate the boundary conditions via an infinite repelling force at rigid boundaries) generally spoils the space-translation symmetry of (2.6). But if, for example, \( \Phi \) represents a force (like gravity) acting only in one direction, and there are no boundaries in a second direction, then the momentum in this second direction is conserved. If, to give another example, \( \Phi(x) \) (and the boundaries) have rotational symmetry, then angular momentum is conserved.

All of this is too familiar to be very exciting. But consider the symmetry property corresponding to variations in the particle labels \((a,b,c)\) that do not affect the \(E\) or \(\Phi\) in (2.6). First suppose that the fluid is homentropic. Then the particle labels enter (2.6) only through the Jacobian (2.3) in \(E(\alpha)\). Let

\[
\begin{align*}
\delta a &= a + \delta a(a,b,c,\tau) \\
\delta b &= b + \delta b(a,b,c,\tau) \\
\delta c &= c + \delta c(a,b,c,\tau)
\end{align*}
\] (2.12)

be a new set of particle labels with the same value of (2.3) as the old labels. That is,

\[
\frac{\partial (a',b',c')}{\partial (x,y,z)} = \frac{\partial (a,b,c)}{\partial (x,y,z)}.
\] (2.13)

The condition (2.13) implies that

\[
\frac{\partial \delta a}{\partial a} + \frac{\partial \delta b}{\partial b} + \frac{\partial \delta c}{\partial c} = 0,
\] (2.14)

to \(O(\delta a^2)\). Thus the particle-label variation is nondivergent in \(a\)-space.

Now consider two realizations of the motion. In the second realization, \(x\) has the same dependence on \(a'\) that it has on \(a\) in the first realization. Thus the two realizations
differ only in the label values assigned to corresponding fluid particles. Since this relabeling affects neither $E$ nor $\Phi$, it corresponds to the action-variation

$$
\int d\tau \int \int \int d\textbf{a} \left\{ \frac{\partial \textbf{x}}{\partial \tau} \cdot \frac{\partial \delta \textbf{x}}{\partial \tau} \right\},
$$

(2.15)

which, according to Hamilton’s principle, must vanish. Here,

$$
\delta \frac{\partial \textbf{x}}{\partial \tau} = \left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a - \left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a.
$$

(2.16)

By the chain rule,

$$
\left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a = \left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a + \left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a - \left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a + \left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a.
$$

(2.17)

Thus,

$$
\delta \frac{\partial \textbf{x}}{\partial \tau} = -\left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a + \left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a - \left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a + \left. \frac{\partial \textbf{x}}{\partial \tau} \right|_a + O(\delta a^2).
$$

(2.18)

Substituting (2.18) into (2.15), and requiring (2.15) to vanish, we obtain

$$
-\int d\tau \int \int \int d\textbf{a} \frac{\partial x_i}{\partial \tau} \frac{\partial x_j}{\partial \tau} \frac{\partial}{\partial a_j} \delta a_i = 0,
$$

(2.19)

where $(x,y,z)=(x_1,x_2,x_3)$, $(a,b,c)=(a_1,a_2,a_3)$, and repeated subscripts denote summation. Now let

$$
A_j \equiv \frac{\partial x_i}{\partial \tau} \frac{\partial x_j}{\partial a_i}.
$$

(2.20)

That is,

$$
A \equiv (A_1, A_2, A_3) = u \nabla_a x + v \nabla_a y + w \nabla_a z,
$$

(2.21)

where $\nabla_a = (\partial / \partial a, \partial / \partial b, \partial / \partial c)$ is the gradient operator in $a$-space. Then, because $\delta a$ vanish at the endpoints in time, (2.19) becomes
\[ \int d\tau \iiint da \frac{\partial A}{\partial \tau} \cdot \delta a = 0. \] (2.22)

Now (2.22) holds only for particle-label variations \( \delta a(a, \tau) \) satisfying (2.14), and (2.14) implies that

\[ \delta a = \nabla_a \times \delta \Gamma(a, \tau) \] (2.23)

for some (arbitrary) \( \delta \Gamma(a, \tau) \). Thus (2.22) becomes

\[ \int d\tau \iiint da \frac{\partial}{\partial \tau} (\nabla_a \times A) \cdot \delta \Gamma = 0. \] (2.24)

Then, since \( \delta \Gamma \) is arbitrary,

\[ \frac{\partial}{\partial \tau} (\nabla_a \times A) = 0. \] (2.25)

This is the conservation law corresponding to the particle-relabeling symmetry in homentropic flow. Readers of Chapter 4 (Section 3) have already seen (2.25). In Chapter 4 we defined \( A \) by

\[ \mathbf{v} = A \nabla a + B \nabla b + C \nabla c, \] (2.26)

but (2.26) is equivalent to (2.21).

The conservation law (2.25) is the general vorticity law for homentropic flow; all the other vorticity laws are direct consequences of (2.25). If, for example, \( \theta(a,b,c) \) is any scalar conserved on fluid particles (\( \partial \theta / \partial \tau = 0 \)), then, because \( \partial / \partial \tau \) commutes with \( \nabla_a \), (2.25) implies that

\[ \frac{\partial}{\partial \tau} \left[ (\nabla_a \times A) \cdot \nabla_a \theta \right] = 0. \] (2.27)

But, as shown in Chapter 4,

\[ (\nabla_a \times A) \cdot \nabla_a \theta = \frac{(\nabla \times \mathbf{v}) \cdot \nabla \theta}{\rho}. \] (2.28)

Thus (2.27) is Ertel’s theorem for homentropic flow.

By (2.25) and Stokes’s theorem,

\[ \frac{d}{dt} \oint A \cdot da = 0, \] (2.29)

where the integration is around a fixed loop in \( a \)-space. But
\[ A \cdot da = v \cdot dx \]  

(2.30)

by (2.20). Thus (2.29) is equivalent to

\[ \frac{d}{dt} \oint v \cdot dx = 0 , \]  

(2.31)

where the integration is around a moving loop that always contains the same fluid particles. This is Kelvin’s circulation theorem.

All of these results apply to homentropic flow, in which the internal energy \( E \) depends only on \( \alpha \). Now we consider the general case of nonhomentropic flow, for which \( E=E(\alpha, \eta) \). The dependence of \( \eta(a) \) on \( a \) partly destroys the particle-relabeling symmetry property, but the last two terms in (2.6) remain invariant to a relabeling of fluid particles that lie in the same constant-entropy surface. In the nonhomentropic case, it is convenient to let the entropy itself be one of the particle labels, say \( c=\eta \). Then \( E+\Phi \) is invariant to particle relabelings (2.12) satisfying \( \delta \mathbf{c}=0 \) and (2.14), which reduces to

\[ \frac{\partial \delta a}{\partial a} + \frac{\partial \delta b}{\partial b} = 0 . \]  

(2.32)

By (2.32), the particle-label variations satisfy

\[ \delta a = - \frac{\partial}{\partial b} \delta \psi(a, \tau), \quad \delta b = \frac{\partial}{\partial a} \delta \psi(a, \tau), \quad \delta c = 0 , \]  

(2.33)

where \( \delta \psi \) is an arbitrary infinitesimal function. The variations (2.33) are a subset of (2.23). Substituting (2.33) into (2.22), we obtain

\[ \int d\tau \int \int da \oint d\psi \frac{\partial}{\partial \tau} \left( \frac{\partial B}{\partial a} - \frac{\partial A}{\partial b} \right) = 0 . \]  

(2.34)

Thus only the \( c \)-component of \( \nabla_a \times A \) is conserved,

\[ \frac{\partial}{\partial \tau} \left[ (\nabla_a \times A) \cdot \nabla_a c \right] = 0 . \]  

(2.35)

Since \( c=\eta \), (2.35) is equivalent to

\[ \frac{\partial}{\partial \tau} \left[ \frac{(\nabla \times v) \cdot \nabla \eta}{\rho} \right] = 0 . \]  

(2.36)

This is the general form of Ertel’s theorem.
The particle-relabeling symmetry property is unique to fluid mechanics. It has no analogue in discrete-particle mechanics, where the particle labels cannot be varied continuously. Similarly, it does not exist in general continuua (e.g. elastic solids), in which the internal energy $E$ depends separately on all of the derivatives $\frac{\partial x_i}{\partial a_j}$. The essence of a fluid is that these derivatives enter the internal energy only through the Jacobian (2.3). We shall see that the particle-relabeling symmetry property is also responsible for the existence of a (closed) Eulerian formulation of fluid mechanics.

3. **Sound waves in one dimension: the conservation of wave action**

It is often useful to regard fluid motion as the sum of a mean flow and a disturbance therefrom. Then interest attaches to the interactions between the mean flow and the disturbance. The most useful statements about these interactions take the form of conservation laws. When the conservation laws are derived directly from the conventional Eulerian equations, the manipulations required are often tedious and unrevealing. However, because the conservation laws reflect symmetry properties of the underlying Lagrangian, Hamiltonian methods offer a simple and motivated derivation. In the next two sections, we illustrate these ideas with the simplest example available to us: a sound wave propagating in a one-dimensional flow.

Let $x(a, \tau)$ be the (one-dimensional) fluid motion, and let

$$x(a, \tau) = X(a, \tau) + \xi(X, T), \quad (3.1)$$

where $X(a, \tau)$ is the mean flow, and $\xi(X(a, \tau), \tau)$ is the displacement at time $\tau$ of the fluid particle labeled by $a$ from the location it would have if it had moved with the mean flow. Here $T=\tau$, but $\partial / \partial T$ will imply that $X$ is held fixed. The $\tau$-derivative of (3.1) is

$$\frac{\partial x}{\partial \tau} = \frac{\partial X}{\partial \tau} + \left( \frac{\partial}{\partial T} + \frac{\partial X}{\partial \tau} \right) \xi(X, T). \quad (3.2)$$

That is,

$$u = U + D\xi, \quad (3.3)$$

where

$$U \equiv \frac{\partial X}{\partial \tau} \quad (3.4)$$

is the mean velocity and

$$D \equiv \frac{\partial}{\partial T} + U \frac{\partial}{\partial X} \quad (3.5)$$

is the time derivative moving with the mean flow.
We assume that the flow is homentropic. Then Hamilton’s principle requires that

\[
L_{\mathbf{x}(a, \tau)} = \int d\tau \int da \left\{ \frac{1}{2} \left( \frac{\partial x}{\partial \tau} \right)^2 - E \left( \frac{\partial x}{\partial a} \right) \right\}
\]

be stationary with respect to variations \( \delta x(a, \tau) \) in the time-dependent mapping from \( a \)-space to \( x \)-space. Substituting (3.1) and (3.2) into (3.6), we obtain

\[
L_{[X(a, \tau), \xi(X,T)]} = \int d\tau \int da \left\{ \frac{1}{2} \left( U + D\xi \right)^2 - E \left( V + V \frac{\partial \xi}{\partial X} \right) \right\}
\]

where

\[
V \equiv \frac{\partial X}{\partial a}
\]

is the specific volume of the mean flow. Hamilton’s principle now requires that (3.7) be stationary with respect to variations in the composite mapping from \( a \)-space into \( X \)-space and from \( X \)-space into \( x \)-space. Since the variations, \( \delta X(a, \tau) \) and \( \delta \xi(X,T) \), in these two mappings are independent, Hamilton’s principle now yields two dynamical equations, reflecting the many possible ways of dividing a single flow into a mean flow and a disturbance. This redundancy allows us to place restrictions on one of the mappings.

Suppose that the disturbance takes the form of wave,

\[
\xi(X,T) = A(X,T) \cos \theta(X,T),
\]

whose amplitude \( A \), wavenumber

\[
k \equiv \frac{\partial \theta}{\partial X},
\]

and frequency

\[
\omega \equiv -\frac{\partial \theta}{\partial T}
\]

all vary slowly, in space and time, compared to \( \theta(X,T) \). We also assume that the wave amplitude \( A \) is small. Then

\[
E \left( V + V \frac{\partial \xi}{\partial X} \right) = E(V) + E'(V) V \frac{\partial \xi}{\partial X} + \frac{1}{2} E''(V) \left( V \frac{\partial \xi}{\partial X} \right)^2 + O(\xi^3).
\]

Substituting (3.12) and (3.9) into (3.7), and averaging the rapidly varying trigonometric terms over a wave cycle, we obtain the averaged Lagrangian.
\[ L = L_1[X(a, \tau)] + L_2[X(a, \tau), A(X, T), \theta(X, T)], \quad (3.13) \]

where
\[ L_1 = \iint d\tau da \left\{ \frac{1}{2} U^2 - E(V) \right\} \quad (3.14) \]

and
\[ L_2 = \iint d\tau da \left\{ (\omega - Uk)^2 - c^2 k^2 \right\}. \quad (3.15) \]

Here, \( U \) and \( V \) are merely abbreviations for (3.4) and (3.8), \( \omega \) and \( k \) are abbreviations for (3.10) and (3.11), and
\[ c^2 \equiv V^2 E'(V) \quad (3.16) \]

will turn out to be the sound speed. We see that \( L_1 \) depends only on the mean flow, while \( L_2 \) depends on both the mean flow and the wave field (and hence couples the two together).

We obtain the equations for the wave field by varying \( A(X, T) \) and \( \theta(X, T) \). These variations affect only \( L_2 \). The amplitude variation yields the dispersion relationship,
\[ \delta A: \quad (\omega - Uk)^2 = c^2 k^2. \quad (3.17) \]

We choose
\[ \omega = Uk + ck, \quad (3.18) \]

corresponding to a wave propagating toward positive \( x \) with respect to the mean flow. (To change the direction of propagation, select the other root of (3.17).) The phase variation yields
\[ \delta L_2 = \int dT dX \frac{\partial a}{\partial X} \left\{ \frac{1}{2} A^2 \left[ (\omega - Uk) \left(-\frac{\partial \theta}{\partial T} - U \frac{\partial \theta}{\partial X}\right) - c^2 k \frac{\partial \theta}{\partial X} \right] \right\} \quad (3.19) \]

leading to the conservation law
\[ \frac{\partial}{\partial T} (W) + \frac{\partial}{\partial X} [(U + c)W] = 0, \quad (3.20) \]

for wave action,
\[ W \equiv \frac{E_r}{\omega_r} \] (3.21)

where

\[ E_r = \frac{1}{2} \bar{\rho} A^2 \omega_r^2 \] (3.22)

is the wave energy measured in a reference frame moving with the mean flow, \( \bar{\rho} \equiv \partial a / \partial X = V^{-1} \) is the mean density, and

\[ \omega_r = \omega - Uk = ck \] (3.23)

is the wave frequency relative to the mean flow. Once again, action conservation corresponds to the \( \theta \)-translation symmetry of the Lagrangian.

We obtain evolution equations for the slowly varying wavenumber and frequency from the definitions (3.10) and (3.11). By (3.10) and (3.11),

\[ \frac{\partial k}{\partial T} + \frac{\partial \omega}{\partial X} = 0. \] (3.24)

By (3.24) and (3.18),

\[ \left[ \frac{\partial}{\partial T} + (U + c) \frac{\partial}{\partial X} \right] k = -k \frac{\partial}{\partial X} (U + c). \] (3.25)

Similarly,

\[ \frac{\partial \omega}{\partial T} = (U + c) \frac{\partial k}{\partial T} + k \frac{\partial}{\partial T} (U + c) \]

\[ = -(U + c) \frac{\partial \omega}{\partial X} + k \frac{\partial}{\partial T} (U + c) \] (3.26)

implies that

\[ \left[ \frac{\partial}{\partial T} + (U + c) \frac{\partial}{\partial X} \right] \omega = +k \frac{\partial}{\partial T} (U + c). \] (3.27)

Thus \( k \) and \( \omega \) propagate at speed \( c \) relative to the mean flow, and experience refractive changes (on account of the right-hand sides of (3.25) and (3.27)) if the mean flow varies in space or time. This completes the description of the slowly varying wave field.

4. **Sound waves in one dimension: the equations for the mean flow**
To obtain the equation for the mean flow \(X(a,\tau)\), we require that (3.13) be stationary with respect to variations \(\delta X(a,\tau)\) in the mapping from \(a\)-space to \(X\)-space, holding the mapping from \(X\)-space to \(x\)-space fixed. This means we fix the dependence of \(A(X,T)\) and \(\theta(X,T)\) on their arguments. However, the variations \(\delta X(a,\tau)\) still cause a change in \(A(X(a,\tau),T)\) and \(\theta(X(a,\tau),T)\), because the value of the \(X\)-argument changes. Thus, for example,

\[
\delta X(a,\tau) : \quad \delta A(X,T) = \frac{\partial A}{\partial X} \delta X(a,\tau). \tag{4.1}
\]

The variations in the mean flow affect both \(L_1\) and \(L_2\). The \(L_1\)-variation takes the familiar form

\[
\delta L_1 = \int d\tau \int da \left\{ -\frac{\partial^2 X}{\partial \tau^2} - V \frac{\partial P}{\partial X} \right\} \delta X(a,\tau), \tag{4.2}
\]

where

\[
P = -E'(V) \tag{4.3}
\]

is the pressure associated with the mean flow. The \(L_2\)-variation contains the effect of the waves on the mean flow. We find that

\[
\delta L_2 = \int d\tau \int da \left\{ \frac{1}{2} A \delta A \left[ (\omega - Uk)^2 - c^2 k^2 \right] \right.

\[
+ \frac{1}{2} A^2 \left[ (\omega - Uk)(\delta \omega - k \delta U - U \delta k) - ck(c \delta k + k \delta c) \right] \right\} \tag{4.4}
\]

By the dispersion relationship (3.18), the coefficient of \(\delta A\) in (4.4) vanishes. We write the remaining terms as

\[
\delta L_2 = \int d\tau \int da \left\{ \frac{1}{2} A^2 \left[ \frac{\partial \omega}{\partial X} - (U + c) \frac{\partial k}{\partial X} \right] \delta X - k \frac{\partial \delta X}{\partial \tau} - kc' \frac{\partial \delta X}{\partial a} \right\}, \tag{4.5}
\]

where \(c' = dc(V)/dV\). The \(\partial \omega/\partial X\)- and \(\partial k/\partial X\)-terms in (4.5) are terms like (4.1); they arise from changes in the spatial argument of \(\omega\) and \(k\), respectively. Integrating (4.5) by parts, we obtain

\[
\delta L_2 = \int d\tau \int da \left\{ \frac{1}{2} A^2 c k \left( \frac{\partial \omega}{\partial X} - (U + c) \frac{\partial k}{\partial X} \right) + \frac{\partial}{\partial \tau} \left( \frac{1}{2} A^2 k^2 c' \right) + \frac{\partial}{\partial a} \left( \frac{1}{2} A^2 k^2 c' \right) \right\} \delta X \tag{4.6}
\]

Thus, by (4.2) and (4.6), Hamilton’s principle implies that
This is the equation for the mean flow. Using the definition (3.21) of wave action and the dispersion relationship (3.18), we can rewrite (4.7) in the form

\[ \frac{\partial}{\partial \tau} \left( U - \frac{Wk}{\bar{\rho}} \right) = -\frac{1}{\bar{\rho}} \frac{\partial}{\partial X} \left( P - \frac{1}{2} A^2 k^2 c^2 \right) + \frac{Wk}{\bar{\rho}} \frac{\partial}{\partial X} (U + c). \]  

(4.8)

This form of the mean-flow equation, which arises naturally from the variational principle, seems to be the most convenient form from which to draw conclusions about the interactions between the mean flow and the waves. However, it is somewhat more conventional to write (4.8) in a form in which all the effects of the waves on the mean flow appear as the divergence of a radiation stress.

To obtain the radiation-stress form of (4.8), multiply (4.8) by the mean density \( \bar{\rho} \), and use the continuity equation,

\[ \frac{\partial \bar{\rho}}{\partial \tau} + \bar{\rho} \frac{\partial U}{\partial X} = 0, \]  

(4.9)

for the mean flow. The result is

\[ \frac{\partial}{\partial T} (\bar{\rho} U) + \frac{\partial}{\partial X} (\bar{\rho} U^2) + \frac{\partial P}{\partial X} = \bar{\rho} \frac{\partial}{\partial \tau} \left( \frac{Wk}{\bar{\rho}} \right) + \frac{\partial}{\partial X} \left( \frac{1}{2} A^2 k^2 c^2 \right) + Wk \frac{\partial}{\partial X} (U + c), \]  

(4.10)

in which the purely mean-flow terms have been collected on the left-hand side. Then use the action equation (3.20), the refraction equation (3.25), and the mean continuity equation (4.9) to eliminate the time-derivatives on the right-hand side of (4.10). After some manipulation, we obtain

\[ \frac{\partial}{\partial T} (\bar{\rho} U) + \frac{\partial}{\partial X} (\bar{\rho} U^2) + \frac{\partial P}{\partial X} = \frac{\partial R}{\partial X}, \]  

(4.11)

where

\[ R = \frac{1}{2} A^2 k^2 c c' - \frac{1}{2} \bar{\rho} A^2 k^2 c^2 \]  

(4.12)

is the radiation stress.

5. **Sound waves in three dimensions:**
   - particle-relabeling and non-acceleration
All of these results generalize easily to three space dimensions. The calculations are straight-forward, and we omit some of the intermediate steps. Readers who have carefully studied the preceding sections will have no trouble supplying the details.

In three dimensions, Hamilton’s principle requires that

\[
L \equiv \int dt \int \int da \left\{ \frac{\partial \mathbf{X}}{\partial \tau} \cdot \frac{\partial \mathbf{X}}{\partial \tau} - E \left( \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right) \right\}
\]

be stationary with respect to variations \( \delta \mathbf{X}(a, \tau) \) in the fluid-particle locations. Once again, we divide the motion into a mean flow and a disturbance,

\[
\mathbf{x}(a, \tau) = \mathbf{X}(a, \tau) + \xi(X, T)
\]

and set

\[
\frac{\partial \mathbf{X}}{\partial \tau} = \frac{\partial \mathbf{X}}{\partial \tau} + \left( \frac{\partial}{\partial T} + \frac{\partial \mathbf{X}}{\partial \tau} \cdot \nabla \mathbf{X} \right) \xi = \mathbf{U} + \mathbf{D} \xi,
\]

where \( \mathbf{U} = \partial \mathbf{X}/\partial \tau \) is the mean velocity, \( \mathbf{D} = \partial / \partial T + \mathbf{U} \cdot \nabla \mathbf{X} \), and \( \nabla \mathbf{X} = (\partial \mathbf{X}/\partial X, \partial \mathbf{X}/\partial Y, \partial \mathbf{X}/\partial Z) \). Next, we expand the internal energy,

\[
E \left( \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right) = E(V) + V \mathcal{E}'(V) \left\{ \nabla \mathbf{X} \cdot \xi + \frac{\partial \eta, \xi}{\partial (Y, Z)} + \frac{\partial (\xi, \xi)}{\partial (X, Z)} + \frac{\partial (\xi, \eta)}{\partial (X, Y)} \right\}
\]

\[
+ \frac{1}{2} c^2 (\nabla \mathbf{X} \cdot \xi)^2 + O(\xi^3)
\]

in powers of \( \xi = (\xi, \eta, \zeta) \). Here, \( V = \partial \mathbf{X}/\partial \mathbf{a} \) and \( c^2 = V^2 E''(V) \) as before. If the disturbance \( \xi(X, T) \) takes the form of a small-amplitude, slowly-varying wave,

\[
\xi = Re \left( A e^{i \theta} \right),
\]

then the average (over a wave cycle) of all the terms proportional to \( E'(V) \) in (5.4) vanishes, and we obtain the averaged Lagrangian

\[
L = L_1[\mathbf{X}(a, \tau)] + L_2[\mathbf{X}(a, \tau), A(X, T), \theta(X, T)],
\]

where

\[
L_1 = \int dt \int \int da \left\{ \frac{\partial \mathbf{X}}{\partial \tau} \cdot \frac{\partial \mathbf{X}}{\partial \tau} - E \left( \frac{\partial \mathbf{X}}{\partial \mathbf{a}} \right) \right\}
\]

and
\[
L_2 = \int d\tau \iiint d\mathbf{a} \left\{ \frac{1}{2} (\omega - \mathbf{U} \cdot \mathbf{k})^2 \mathbf{A} \cdot \mathbf{A}^* - \frac{1}{2} c^2 (\mathbf{k} \cdot \mathbf{A})(\mathbf{k} \cdot \mathbf{A}^*) \right\}.
\]

(5.8)

Here,
\[
\mathbf{k} \equiv \nabla_x \theta \quad \text{and} \quad \omega \equiv -\frac{\partial \theta}{\partial T}
\]

(5.9)

are the slowly varying wavenumber and frequency of the wave. Once again, \(L_1\) depends only on the mean flow, while \(L_2\) depends on both the mean flow and the slowly varying wave.

We obtain the equations for the wave field by varying \(\mathbf{A}(\mathbf{X},T)\) and \(\theta(\mathbf{X},T)\). The amplitude variations yield
\[
\delta \mathbf{A} : \quad \mathbf{A} = A \frac{\mathbf{k}}{|\mathbf{k}|} \quad \text{and} \quad \omega_r \equiv \omega - \mathbf{U} \cdot \mathbf{k} = ck,
\]

(5.10)

where now \(k = |\mathbf{k}|\), and we adopt the convention that the wave propagates in the direction of \(\mathbf{k}\). With no loss in generality, we take \(A\) to be real. Then (5.8) takes the form
\[
L_2 = \int d\tau \iiint d\mathbf{a} \left\{ \frac{1}{2} (\omega - \mathbf{U} \cdot \mathbf{k})^2 \mathbf{A}^2 - \frac{1}{2} c^2 k^2 A^2 \right\}.
\]

(5.11)

The phase variation yields the equation
\[
\delta \theta : \quad \frac{\partial W}{\partial T} + \nabla_x \cdot \left[ \left( \mathbf{U} + c \frac{\mathbf{k}}{k} \right) W \right] = 0
\]

(5.12)

for the conservation of wave action,
\[
W = \frac{1}{2} \rho A^2 ck = \frac{E_r}{\omega_r}.
\]

(5.13)

Once again, we obtain the evolution equation for \(\mathbf{k}\) and from the definitions (5.9) and the dispersion relationship (5.10b), finding that
\[
\frac{\partial \mathbf{k}}{\partial T} + \nabla_x (\mathbf{U} \cdot \mathbf{k} + ck) = 0.
\]

(5.14)

To get the equations for the mean flow, we require \(L\) to be stationary with respect to arbitrary variations \(\delta \mathbf{X}(\mathbf{a},\tau)\) in the mean flow. We find that
\[
\delta L_2 = \int d\tau \iiint d\mathbf{a} \left\{ \frac{1}{2} A^2 \omega_r \left[ \delta \omega + \mathbf{U} \cdot \delta \mathbf{k} - \mathbf{k} \cdot \frac{\partial \delta \mathbf{X}}{\partial \tau} - \frac{1}{2} A^2 c k \delta \mathbf{X} \frac{\partial (\mathbf{X})}{\partial (\mathbf{a})} \right] \right\} \frac{\partial \delta \mathbf{X}}{\partial \mathbf{a}},
\]

(5.15)
leading to
\[
\delta X : \quad \frac{\partial}{\partial \tau} \left( U_i - \frac{WK_i}{\rho} \right) = -\frac{1}{\rho} \frac{\partial}{\partial X_i} \left( P - \frac{1}{2} A^2 \kappa^2 \chi' \right) + \frac{W}{\rho} \left( k_j \frac{\partial U_j}{\partial X_i} + k \frac{\partial c}{\partial X_i} \right),
\]
(5.16)
in which repeated subscripts denote summation. Once again, we may use (5.12), (5.14) and the three-dimensional analogue of (4.9) to rewrite (5.16) in the form
\[
\frac{\partial}{\partial T} \rho U_i (\text{4.9}) + \frac{\partial}{\partial X_i} \rho U_i U_j (\text{4.9}) + \frac{\partial P}{\partial X_i} = \frac{\partial R_{ij}}{\partial X_j},
\]
(5.17)
where
\[
R_{ij} = \frac{1}{2} A^2 \left( \chi' k^2 \delta_{ij} - \rho c^2 k_i k_j \right)
\]
(5.18)
is the radiation stress, but an interesting result follows more directly from (5.16). From (5.16) we can show that
\[
\frac{\partial}{\partial \tau} \left( U - \frac{WK}{\rho} \right) \cdot dX = 0,
\]
(5.19)
where the integration is around a closed loop of fluid particles moving at the velocity of the mean flow. However, we obtain (5.19) in a more direct and motivated fashion by considering the particle-relabeling symmetry of L. Thus consider particle-label variations \( \delta a(a, \tau) \) that leave \( \partial X_i/\partial a_j \) unchanged. These variations do not affect \( A(X,T) \) and \( \theta(X,T) \) at all. Thus
\[
\delta L = \int d\tau \iiint da \left\{ \frac{\partial X_i}{\partial \tau} \cdot \delta \frac{\partial X_i}{\partial \tau} - \frac{1}{2} A^2 \omega \cdot k \cdot \delta \frac{\partial X_i}{\partial \tau} \right\}
\]
(5.20)
The rest goes as in Section 2. We find that
\[
\frac{\partial}{\partial \tau} (\nabla_a \times A) = 0
\]
(5.21)
where now (cf. (2.20))
\[
A_j \equiv \left( \frac{\partial X_i}{\partial \tau} - \frac{WK_i}{\rho} \right) \frac{\partial X_i}{\partial a_j}.
\]
(5.22)
Thus (cf. (2.30))

\[ A \cdot da = \left\{ \frac{\partial X}{\partial \tau} - \frac{Wk}{\bar{\rho}} \right\} dX \]  

(5.23)

The circulation theorem (5.19) follows from (5.21) and (5.23) in the same way that (2.31) followed from (2.25) and (2.30). We also obtain the potential vorticity law (cf. (2.27-28))

\[ \frac{\partial}{\partial \tau} \left[ \nabla_X \times (U - \frac{Wk}{\bar{\rho}}) \cdot \nabla_X \Theta \right] = 0 \]  

(5.24)

where \( \Theta \) is any conserved scalar \( (\partial \Theta / \partial \tau = 0) \). However, (5.21) and (5.23) are perhaps themselves the most succinct and illuminating statement of what has come to be called a non-acceleration theorem.

Consider an initially quiescent fluid \( (A=0) \) into which waves begin to propagate. Motion ensues, but it must satisfy (5.21). According to (5.21-22), a mean flow is present only where waves are present (wave action \( W \neq 0 \)), and it disappears as soon as the wave action propagates away. Thus without dissipation, no permanent mean flow can arise. But consider what happens if some local wave dissipation suddenly grabs some of the wave action \( W \) and removes it from the fluid, leaving the mean flow \( \partial X / \partial \tau \) alone. Then the conservation law (5.21) guarantees that a permanent mean flow will be left after the remaining waves propagate away. Readers of Chapter 6 (Section 1) have already met a non-acceleration theorem of this type, in connection with the generation of a quasigeostrophic mean flow over topography.

This example of sound waves propagating in a mean flow is certainly rather special, but our treatment illustrates the really fundamental ideas that would occur in a more sweeping discussion of this subject. However, before leaving the subject of waves and mean flows, we mention several lines along which the theory presented here can be substantially generalized.

First, the wave amplitudes need not be small. In fact, Whitham originally developed the averaged Lagrangian method as a means for handling nonlinear waves. To consider waves of arbitrary amplitude, one simply omits the step of truncating the averaged Lagrangian at the quadratic order in the wave amplitude. The dispersion relation acquires an amplitude-dependence, and the wave action is no longer \( E_\nu / \omega_\nu \), but the variational method automatically supplies the needed generalizations.

Second, the waves need not be slowly varying, and, in fact, they need not even be waves! One simply introduces an additional dependent variable, the ensemble parameter \( \mu \), into the disturbance field,

\[ \xi = \xi(X,T,\mu). \]  

(5.25)

The parameter \( \mu \) identifies the members of an ensemble of flow realizations. Each member of the ensemble evolves independently of the others, and the Lagrangian of the
meta-system is simply the $\mu$-integral (or ensemble average) of the individual Lagrangians. This meta-system has an obvious $\mu$-translation symmetry, and the resulting conservation law is a very general statement of action conservation. However, to use this general law one must usually identify the ensemble average with another kind of average (such as the average over the wave-cycle of a single wavetrain) thus partly abandoning the generality.

Finally, the mean-flow fields $U(X,T)$ and $V(X,T)$ can themselves have symmetry. By Noether’s theorem, these symmetries also correspond to conservation laws. If the mean fields have time-translation symmetry, then the resulting conservation law is a statement of pseudoenergy conservation. If the mean fields are invariant to translation in a particular direction, then the corresponding pseudomomentum is conserved. If the mean fields have symmetry, then it is natural to identify the average in the corresponding space-time direction with the ensemble average of the disturbance. Then the generalized action acquires a concrete meaning.

Finally, we should mention that, in the case of a symmetry corresponding to the translation of an independent variable such as $a, X, T$ or $\mu$, the energy-momentum-tensor formulas may be useful. These formulas express the result of a varying the independent variables in an arbitrary Lagrangian. Memorize them, and you will never have to do another variation. But for those, like me, who enjoy taking variations, the energy-momentum-tensor formalism is a nuisance to remember.

6. Approximations and constraints

Thus far, we have considered only the full Lagrangian corresponding to the exact equations for a perfect fluid. However, in many cases we are interested in dynamical equations that represent an approximation to the exact equations. For example, theorists often use Euler’s equations to study flows in which sound waves are thought to be unimportant. Euler’s equations are a useful approximation, because they filter out the sound waves.

In nearly every case, the (conservative form of the) approximate dynamical equations are equivalent to a variational principle. In most cases, we can view the Lagrangian in this variational principle as an approximation to the exact Lagrangian for the perfect fluid. In many cases, the approximation corresponds to a constraint. In this section, we give simple examples of approximate fluid equations obtained from constrained Lagrangians. Later we will suggest that useful new approximations can be obtained by attaching constraints to the Lagrangian.

There are two methods for constraining variational principles. The first and most general is Lagrange’s method of undetermined multipliers. However, in some cases, we can avoid Lagrange multipliers by solving the constraint, that is, by using the constraint to remove some of the variables. That is the second method. Our examples illustrate both methods. We begin with a brief explanation of Lagrange’s method.

First consider the problem of finding the maximum or minimum (or inflection points) of a function $f(x_1, x_2, x_3, ..., x_N)$ of $N$ variables. As is well known, the extrema occur at the points where

$$ \nabla f = 0. \quad (6.1) $$

Here, $\nabla$ is the gradient operator in $N$-dimensional space. Thus (6.1) represents $N$ equations. According to (6.1), the extrema occur at the points where $f$ is stationary with respect to infinitesimal variations $\delta x_i$ in its arguments.

But suppose we want to find the extrema of $f$ subject to the $M$ constraints that

$$\phi_m(x_1, x_2, \ldots, x_N) = 0, \quad m = 1, \ldots, M .$$

The constraints (6.2) define an $N-M$ dimensional hypersurface in $N$-dimensional space. In this case, the condition that (6.1) hold at an extremum is overly restrictive. It is only necessary that $\nabla f$ have no projection on the hypersurface (6.2). That is, $\nabla f$ must lie in the $M$-dimensional subspace of vectors locally perpendicular to the hypersurface. This subspace is spanned by the $M$ vectors $\nabla \phi_m$. Thus

$$\nabla f = -\sum_m \lambda_m \nabla \phi_m ,$$

at the extrema on the hypersurface. The $\lambda_m$ are $M$ constants that remain to be determined.

To determine the extrema (or stationary points) of $f$ on the hypersurface (6.2), we solve the $N$ equations (6.3) and the $M$ equations (6.2) for $(x_1, x_2, \ldots, x_N)$ and $(\lambda_1, \lambda_2, \ldots, \lambda_M)$. However, the equations (6.3) also arise from the requirement that the $x_i$-derivatives of

$$F(x_1, \ldots, x_N, \lambda_1, \ldots, \lambda_M) \equiv f + \sum_m \lambda_m \phi_m$$

vanish, while (6.2) arise from the requirement that the $\lambda_m$-derivatives of (6.4) vanish. Hence we may say that the problem of finding the stationary points of $f(x_1, \ldots, x_N)$ subject to the constraints $\phi_m=0$ is equivalent to finding the stationary points of $F(x_1, \ldots, x_N, \lambda_1, \ldots, \lambda_M)$ with no constraints.

We obtain Euler’s equations of fluid motion by imposing the constraints

$$\frac{\partial(x,y,z)}{\partial(a,b,c)} = \alpha_0$$

on the Lagrangian for a perfect fluid. The constraints (6.5) require the specific volume to have the uniform value $\alpha_0$. The $\tau$-derivative of (6.5) yields

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 .$$

Since $\alpha$ is uniform, the specific internal energy $E(\alpha_0, \eta(a))$ depends only on the fluid-particle identity $a$, and hence the total internal energy
\[ \iiint da \ E(\alpha_0, \eta(a)) \]  

(6.7)

is unaffected by variations \( \delta x(a, \tau) \) in the fluid-particle locations. The fluid Lagrangian thus becomes

\[ L[x(a, \tau)] = \iiint da \left\{ \frac{1}{2} \frac{\partial x}{\partial \tau} \cdot \frac{\partial x}{\partial \tau} - \Phi(x) + \lambda(a, \tau) \left( \frac{\partial(x)}{\partial(a)} - \alpha_0 \right) \right\}, \]  

(6.8)

where \( \lambda(a, \tau) \) is the Lagrange-multiplier field corresponding to the constraints (6.5).

There are a continuous infinity of these constraints, one for every \( a \) and \( \tau \). Hamilton’s principle now states that

\[ \delta \int L \, dt = 0 \]  

(6.9)

for arbitrary infinitesimal variations \( \delta x(a, \tau) \) and \( \delta \lambda(a, \tau) \) in the particle locations and Lagrange multipliers. The Lagrange-multiplier variations yield (6.5) (and hence (6.6)) at every particle and time. The particle-location variations yield

\[ \delta x : \quad \frac{\partial^2 x}{\partial \tau^2} = -\alpha_0 \nabla \lambda - \nabla \Phi \]  

(6.10)

by steps similar to those in Chapter 1. Thus \( \lambda \) turns out to be the (dynamic) pressure. It is determined by (6.6) in the same way that the \( \lambda_m \)'s in (6.3) are determined by the \( M \) constraints (6.2).

Our next example can be viewed as a further constraint on the Lagrangian for a uniform-density fluid. We now suppose that the fluid has a free surface and lies above a flat, rigid lower boundary, and we constrain the fluid to move in vertical columns. This time, however, we avoid Lagrange multipliers by building the constraints into our description of the fluid. We do this by assuming that the horizontal locations,

\[ x(a, b, \tau), \ y(a, b, \tau) \]  

(6.11)

depend on only two of the particle labels. Thus every fluid particle with the same value of \( (a, b) \) lies in the same vertical column, and, since \( (u,v) = (\partial x/\partial \tau, \partial y/\partial \tau) \), this material column remains vertical. The uniform-density constraint (6.5) becomes

\[ \frac{\partial(x, y)}{\partial(a, b)} \frac{\partial z}{\partial c} = \alpha_0, \]  

(6.12)

because only \( z(a, b, c, \tau) \) depends on \( c \). Directly from (6.12),
We assign the labels so that \( c = 0 \) at the rigid bottom \( z = 0 \), and \( c = H_0 \) (constant) at the free surface. Since the fluid particles on these boundaries remain there, these assignments hold for all time. Then the \( \text{const} \) in (6.13) vanishes and the fluid depth is

\[
h \equiv \frac{\partial(a,b)}{\partial(x,y)} \alpha_0 H_0. \tag{6.14}\]

Thus,

\[
z = \frac{c}{H_0} h, \tag{6.15}\]

and

\[
\frac{\partial z}{\partial \tau} = \frac{c}{H_0} \frac{\partial h}{\partial \tau}. \tag{6.16}\]

We build the constraints into the Lagrangian by regarding the horizontal locations (6.11) as functions of the column-label \( (a,b) \) only, and by using (6.15) and (6.16) to eliminate \( z \) in favor of \( x \) and \( y \). Since

\[
\int \int da \ db \int_0^{H_0} dc \left\{ \frac{1}{2} \left( \frac{\partial x}{\partial \tau} \right)^2 + \frac{1}{2} \left( \frac{\partial y}{\partial \tau} \right)^2 \right\} = \frac{1}{2} H_0 \int \int da \ db \left\{ \left( \frac{\partial x}{\partial \tau} \right)^2 + \left( \frac{\partial y}{\partial \tau} \right)^2 \right\}, \tag{6.17}\]

\[
\int \int da \ db \int_0^{H_0} dc \left\{ \frac{1}{2} \left( \frac{\partial z}{\partial \tau} \right)^2 \right\} = \int \int da \ db \int_0^{H_0} dc \left\{ \frac{1}{2} \left( \frac{c}{H_0} \frac{\partial h}{\partial \tau} \right)^2 \right\} = \frac{1}{2} H_0 \int \int da \ db \left\{ \frac{1}{2} \left( \frac{\partial h}{\partial \tau} \right)^2 \right\}, \tag{6.18}\]

and

\[
\int \int da \ db \int_0^{H_0} dc \{ gz \} = \int \int da \ db \int_0^{H_0} dc \left\{ g \frac{c}{H_0} h \right\} = \frac{1}{2} H_0 \int \int da \ db \{ gh \}, \tag{6.19}\]

the fluid Lagrangian becomes

\[
L[x(a,b,\tau), y(a,b,\tau)] = \frac{1}{2} \int \int da \ db \left\{ \left( \frac{\partial x}{\partial \tau} \right)^2 + \left( \frac{\partial y}{\partial \tau} \right)^2 + \frac{1}{2} \left( \frac{\partial h}{\partial \tau} \right)^2 - gh \right\}. \tag{6.20}\]
(apart from an irrelevant constant factor). The Lagrangian (6.20) depends only on the horizontal locations of the labeled fluid columns; the symbol \( h \) is merely an abbreviation for the right-hand side of (6.14).

We obtain the approximate dynamical equations by requiring that \( \delta \int L d\tau = 0 \) for arbitrary variations \( \delta x, \delta y \) in the horizontal locations of the fluid columns. The \( \delta x \)-variations yield

\[
\delta x : \quad \delta L = \int \int da db \left\{ \frac{\partial x}{\partial \tau} \frac{\partial x}{\partial \tau} + \frac{1}{2} \frac{\partial h}{\partial \tau} \frac{\partial h}{\partial \tau} - \frac{1}{2} g \frac{\partial h}{\partial \tau} \right\}
\]

\[
= \int \int da db \left\{ -\frac{\partial^2 x}{\partial \tau^2} \delta x - \frac{1}{2} \frac{\partial^2 h}{\partial \tau^2} \delta h - \frac{1}{2} g \delta h \right\}
\]

(6.21)

But for any quantity \( F \),

\[
\int \int da db \{ F \delta h \} = \int \int da db \left\{ -Fh^2 \delta \left( \frac{1}{h} \right) \right\} = \int \int da db \left\{ -Fh^2 \delta \left( \frac{1}{\alpha_0 H_0} \frac{\partial (x, y)}{\partial (a, b)} \right) \right\}
\]

\[
= \frac{1}{\alpha_0 H_0} \int \int da db \left\{ -Fh^2 \frac{\partial (x, y)}{\partial (a, b)} \right\} = \frac{1}{\alpha_0 H_0} \int \int da db \left\{ \delta x \frac{\partial (Fh^2, y)}{\partial (a, b)} \right\}
\]

\[
= \int \int da db \left\{ \delta x \frac{1}{h} \frac{\partial}{\partial x} (Fh^2) \right\}
\]

(6.22)

Thus by (6.21-22) and the corresponding equations for \( \delta y \), we obtain the equation of motion,

\[
\delta x : \quad \frac{Du}{Dt} = -g\nabla h - \frac{1}{3h} \nabla \left( h^2 \frac{D^2 h}{Dt^2} \right)
\]

(6.23)

where \( u \) is the horizontal velocity and \( \nabla \) the horizontal gradient operator.

If, in a further approximation, we neglect the kinetic energy arising from the vertical velocity and thus use the Lagrangian

\[
L[x(a, b, \tau), y(a, b, \tau)] = \frac{1}{4} \int \int da db \left\{ \left( \frac{\partial x}{\partial \tau} \right)^2 + \left( \frac{\partial y}{\partial \tau} \right)^2 - gh \right\}
\]

(6.24)

instead of (6.20), we obtain

\[
\delta x : \quad \frac{Du}{Dt} = -g \nabla h
\]

(6.25)

instead of (6.23).\(^9\) In either case, the \( \tau \)-derivative of the definition (6.14) yields
\[
\frac{Dh}{Dt} + h \nabla \cdot \mathbf{u} = 0. \tag{6.26}
\]

Equations (6.25) and (6.26) are the shallow-water equations. Equations (6.23) and (6.26) are the equations for a thin layer of fluid proposed by Green and Naghdi. The Green-Naghdi equations are more general than the shallow-water equations, because they take the vertical acceleration into account. They are also more general than the usual Boussinesq equations for surface waves in a shallow fluid in that they apply to flows with vorticity. Both the shallow-water and Green-Naghdi equations may be subjected to further approximations, corresponding to further constraints on (6.20) and (6.24).

There are two important advantages to deriving approximate dynamical equations by this method of applying the approximations directly to the Lagrangian. First, the conservation laws survive if the approximations respect the corresponding symmetry properties. Second, transformations to new dependent and independent variables, in which the approximate dynamics takes its simplest mathematical form, automatically suggest themselves. The present examples illustrate the first advantage very well; we illustrate the second advantage in Section 13.

The Lagrangians (6.20) and (6.24) inherit the time-translation and particle-relabeling symmetry properties of the exact Lagrangian for a perfect fluid. In the case of (6.20) and (6.24), the relabeling symmetry corresponds to a relabeling of fluid columns that does not affect the Jacobian in (6.14). Thus, by Noether’s theorem, the shallow-water and Green-Naghdi equations exactly conserve an approximate form of the energy and potential vorticity. In the case of the shallow-water equations, the time-translation symmetry leads to the conservation of energy

\[
\frac{1}{2} \int \int \! dadb \left\{ \left( \frac{\partial x}{\partial \tau} \right)^2 + \left( \frac{\partial y}{\partial \tau} \right)^2 + gh \right\}, \tag{6.27}
\]

and the column-relabeling symmetry leads (by steps similar to those in Section 2) to the conservation of potential vorticity,

\[
\frac{D}{Dt} \left( \frac{\zeta}{h} \right) = 0, \tag{6.28}
\]

where \( \zeta \equiv v_x - u_y \) is the relative vorticity and \( \mathbf{u} = (u, v) \). In the case of the Green-Naghdi equations, these two symmetries lead to the conservation of energy,

\[
\frac{1}{2} \int \int \! dadb \left\{ \left( \frac{\partial x}{\partial \tau} \right)^2 + \left( \frac{\partial y}{\partial \tau} \right)^2 + \frac{1}{3} \left( \frac{\partial h}{\partial \tau} \right)^2 + gh \right\}, \tag{6.29}
\]

and potential vorticity,
\[ \frac{D}{Dt} \left( \zeta + \frac{1}{2} J \frac{Dh}{Dt} \right) = 0, \]  

where \( J \) is the horizontal Jacobian. The kinetic energy in (6.29) includes a contribution from the vertical velocity. Similarly, it can be shown that (6.30) is a closer approximation to the general Ertel invariant than is (6.28). The conservation laws survive because our approximations to the Lagrangian do not violate the time-translation and particle-relabeling symmetries. On the other hand, approximations applied directly to the equations of motion sometimes result in the loss of conservation laws, even when these approximations are based upon a very careful scaling analysis. The retention of conservation laws, which seems to be important, is a primary advantage of approximation methods based upon Hamilton’s principle.

7. The canonical equations

We have been using Hamilton’s principle in the form

\[ \delta \int_{t_1}^{t_2} L(q_i, \dot{q}_i) dt = 0, \]  

for arbitrary independent variations \( \delta q(t) \) in the \( N \) generalized coordinates of the system. The variations vanish at the endpoints in time, \( \delta q_i(t_1) = \delta q_i(t_2) = 0 \). In the fluid system, the \( q_i \) correspond to the locations of labeled fluid particles, and the subscript \( i \) represents both the particle-label and the directional component.

The variational principle (7.1) yields the Euler-Lagrange equations,

\[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) = \frac{\partial L}{\partial q_i}. \]  

This leads us to define the generalized momenta

\[ p_i = \frac{\partial L}{\partial \dot{q}_i}, \]  

and the Hamiltonian,

\[ H = \sum_i p_i \dot{q}_i - L. \]  

From (7.3) and (7.4) it follows that

\[ dH = \sum_i \left\{ \dot{q}_i dp_i + p_i d\dot{q}_i - \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i \right\} = \sum_i \left\{ \dot{q}_i dp_i - \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i \right\}. \]  

VII-27
Now, if we can solve (7.3) for \( \dot{q}_i(p,q) \), then we can regard \( H=H(p,q) \). (By \((p,q)\), we mean the whole set, \((p_1,...,p_N,q_1,...,q_N)\).) Then (7.5) implies that
\[
\frac{\partial H}{\partial p_i} = q_i, \quad \frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i}.
\] (7.6)

By (7.2) and (7.3), these take the form of Hamilton’s canonical equations,
\[
\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}.
\] (7.7)

The dynamics (7.7) is equivalent to (7.2).

The canonical equations (7.7) continue to hold if \( L(q,\dot{q},t) \), and hence \( H(p_i;q_i,t) \), contains an explicit time-dependence, like that arising from \( \omega_0(t) \) in Section 1. Just as the (nonrelativistic) Lagrangian \( L \) is always the difference between the kinetic and potential energies, the Hamiltonian \( H \) always turns out to be equal to the energy of the system. That is,
\[
E = H(p,q,t),
\] (7.8)

where \( E \) is the energy of the system. From (7.7) it follows that
\[
\frac{dE}{dt} = \frac{\partial H}{\partial p_i} \frac{dp_i}{dt} + \frac{\partial H}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}.
\] (7.9)

Thus \( E \) is conserved if the \( H \) has no explicit time-dependence.

It is easy to show that the canonical equations (7.7) are equivalent to the extended form of Hamilton’s principle,
\[
\delta \int_h^b dt \left\{ p_i \frac{dq_i}{dt} - H(p,q,t) \right\} = 0,
\] (7.10)

where \( \delta \) now stands for arbitrary independent variations in the \( p_i \)'s and \( q_i \)'s. The \( q_i \)-variations (but not necessarily the \( p_i \)-variations) must vanish at the endpoints in time. If \( H \) has no explicit time-dependence, then the time-translation symmetry of (7.10) leads to the energy-conservation law, \( dH/dt=0 \).

How does this work in the case of the perfect fluid? The perfect-fluid Lagrangian is
\[
L[x(a,\tau)] = \int \int \int da \left\{ \frac{\dot{x}}{\dot{\tau}} \cdot \frac{\dot{x}}{\dot{\tau}} - E \left( \frac{\partial x}{\partial a} , \eta(a) \right) - \Phi(x) \right\},
\] (7.11)
where $x(a, \tau)$ is analogous to $q_i(t)$. Therefore, in analogy with (7.3), the generalized momenta are

$$p(a, \tau) = \frac{\delta L}{\delta (\frac{\partial x}{\partial \tau})} = \frac{\partial x}{\partial \tau}, \tag{7.12}$$

Here $\delta F/\delta f$ denotes the functional derivative, defined by

$$\delta F[f(a)] = \int \int \int da \frac{\delta F}{\delta f(a)} \delta f(a), \tag{7.13}$$

where $F[f]$ is any functional of the function $f(a)$. In analogy with (7.4), the Hamiltonian is

$$H[u(a, \tau), x(a, \tau)] = \int \int \int da \left\{ u \frac{\partial x}{\partial \tau} - L \right\} = \int \int \int da \left\{ \frac{1}{2} u \cdot u + E \left( \frac{\partial(x)}{\partial(a)} \cdot \eta(a) \right) + \Phi(x) \right\}. \tag{7.14}$$

Thus $H$ is indeed the energy. The canonical equations take the form

$$\frac{\partial u}{\partial \tau} = -\frac{\delta H}{\delta x}, \quad \frac{\partial x}{\partial \tau} = \frac{\delta H}{\delta u}. \tag{7.15}$$

These are equivalent to Hamilton’s principle in the form

$$\delta \int d\tau \left\{ \int \int \int da \ u \frac{\partial x}{\partial \tau} - H[u, x] \right\} = 0, \tag{7.16}$$

for arbitrary independent variations, $\delta u(a, \tau)$ and $\delta x(a, \tau)$, in the velocities and locations of labeled fluid particles. The time- and particle-relabeling symmetries of $H$ yield the same energy and potential-vorticity conservation laws as before.

For reasons which begin to appear in Section 9, the canonical equations (7.7) and the extended form (7.10) of Hamilton’s principle are usually preferred over (7.1) and (7.2). However, merely by writing (7.7), we have reached an important crossroads of mathematical physics, and it is worthwhile to pause and look a ways in each direction before taking up our main path. This we do in the remainder of this section. Readers who are in too much of a hurry to admire the scenery may skip ahead to Section 8. The remainder of this section is admittedly primarily cultural, but, on the other hand, nothing promotes learning better than establishing a connection between something new and something entirely familiar.

We can regard the canonical equations (7.7) as the equations for a trajectory in the $2N$-dimensional phase space spanned by the $p_i$’s and $q_i$’s. This was the viewpoint adopted in Chapter 5. By (7.7), the $2N$-dimensional flow in phase space is nondivergent,
\[
\sum_{i=1}^{N} \left( \frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = 0.
\] (7.17)

In fact, the flow in each two-dimensional \((p_i, q_i)\)-subspace of phase space is also nondivergent.

Alternatively, we can regard (7.7) as the ray equations for slowly-varying waves propagating through the \(N\)-dimensional configuration space spanned by the \(q_i\)'s. This wave-interpretation of (7.7) was, in fact, a primary motivation for Hamilton's work, and most mechanics books contain a chapter about it. However, it is very important to emphasize that the hypothetical waves described by (7.7) have nothing to do with the waves that may actually be present in the physical system represented by \(H(p,q)\).

In the usual explanation of the analogy between mechanics and optics, one starts with mechanics and moves in the direction of wave theory. However, readers with a background in fluid mechanics are more likely to appreciate the analogy in reverse, by beginning from the equations for a slowly varying wavetrain in \(N\)-dimensional space.

In the simplest approach to slowly varying waves, one simply assumes that the slowly varying frequency and wavenumber obey the same (locally evaluated) dispersion relation as do plane waves in a uniform medium. Let

\[
\psi(x_1, x_2, \ldots, x_N, t)(7.18)
\]

be the dependent variable in \(N\)-dimensional space, and temporarily suppose that the medium is uniform. Then plane waves of the form

\[
\psi = A_0 e^{i(k_0 \cdot x - \omega_0 t)}(7.19)
\]

(where \(A_0, k_0, \) and \(\omega_0\) are constants, and \(k_0\) and \(x\) are \(N\)-dimensional) obey a dispersion relation of the form

\[
\omega_0 = \Omega(k_0 ; D_0), (7.20)
\]

where \(D_0\) is a (constant) parameter (or set of parameters) that depends on the medium. The dispersion relation (7.20) comes from the physics of \(\psi\); we take it as given.

Now suppose that the solution takes the form of a slowly varying wave,

\[
\psi = A(x,t) e^{\theta(x,t)}(7.21)
\]

By slowly varying, we mean that the amplitude \(A\), wavenumber

\[
k \equiv \frac{\partial \theta}{\partial x}, (7.22)
\]

and frequency
\[ \omega \equiv -\frac{\partial \theta}{\partial t} \quad (7.23) \]

all vary slowly compared to \( \theta(x,t) \). This slow variation could arise from the waves themselves (as in the dispersion from a distant source) or from slow variations in the medium, through the parameter(s) \( D(x,t) \). In either case, we assume that

\[ \omega(x,t) = \Omega(k(x,t); D(x,t)), \quad (7.24) \]

where \( \Omega \) is the same function as in (7.20). That is, the slowly varying frequency and wavevector obey the same dispersion relation as the plane wave. This approximation, often called geometrical optics, also arises at the first order of WKB theory.

We can use the ansatz (7.24) to determine \( \theta(x,t) \) in two different but equivalent ways. In the first way, we substitute the definitions (7.22) and (7.23) into (7.24) to obtain a first-order partial differential equation for \( \theta(x,t) \),

\[ -\frac{\partial \theta}{\partial t} = \Omega \left( \frac{\partial \theta}{\partial x}; D(x,t) \right). \quad (7.25) \]

In the second way, we use (7.22), (7.23), and (7.24) to form ray equations for \( k(x,t) \) and \( \omega(x,t) \). To form the equation for \( k \), we write

\[ \frac{\partial k_i}{\partial t} = -\frac{\partial \omega}{\partial x_i} = -\frac{\partial}{\partial x_i} \Omega(k; D) = -\frac{\partial \Omega}{\partial k_j} \frac{\partial k_j}{\partial x_i} - \frac{\partial \Omega}{\partial D} \frac{\partial D}{\partial x_i} \quad (7.26) \]

where repeated subscripts are summed from 1 to \( N \). Then, defining the group velocity,

\[ C_i \equiv \frac{\partial \Omega}{\partial k_i}, \quad (7.27) \]

we rewrite (7.26) as

\[ \left( \frac{\partial}{\partial \xi} + C_j \frac{\partial}{\partial x_j} \right) k_i = -\frac{\partial \Omega}{\partial x_i}, \quad (7.28) \]

where by

\[ \frac{\partial \Omega}{\partial x_i} \equiv \frac{\partial \Omega}{\partial D} \frac{\partial D}{\partial x_i}, \quad (7.29) \]
we mean the derivative arising only from changes in the medium.

According to (7.28), an observer moving at the group velocity sees a constant $k$ if the medium doesn’t vary with $x$. This inspires us to write the single equation (7.28) as two equations,

$$\frac{dk_i}{dt} = -\frac{\partial \Omega}{\partial x_i}, \quad \text{and} \quad \frac{dx_i}{dt} = \frac{\partial \Omega}{\partial k_i}.$$  \hspace{1cm} (7.30)

To solve (7.30), we pick any point $x$ and time $t$, and step (7.30) forward in time, finding the equation $x(t)$ of the ray, and the corresponding wavenumber $k(t)$ along it. At every time, we can obtain the frequency $\omega(t)$ from (7.24), or (less simply) from

$$\frac{d\omega}{dt} = \frac{\partial \Omega}{\partial t}.$$  \hspace{1cm} (7.31)

The derivation of (7.31) is similar to that of (7.30a); again, $\partial \Omega/\partial t$ denotes the time-change arising from the medium. After finding the wavenumbers and frequencies along many rays, we know $k(x,t)$ and $\omega(x,t)$, and we can then determine $\theta(x,t)$ from (7.22-23). (This also gives us a general method for solving first-order partial differential equations in $N+1$ dimensions, which can always be written in the general form (7.25).)

Now we come to the main point: (7.30) are analogous to (7.7). More generally, every Hamiltonian system is analogous to a field of slowly varying waves with a dispersion relation determined by the Hamiltonian. The analogy is between

$$x_i \leftrightarrow q_i \quad \omega \leftrightarrow E$$  \hspace{1cm} (7.32)

In particular, (7.8) is analogous to the dispersion relation (7.24).

Just as the ray equations (7.30) imply the existence of a phase $\theta(x,t)$ satisfying (7.22) and (7.23), the canonical equations (7.7) imply the existence of a function $S(q,t)$ satisfying

$$p_i = \frac{\partial S}{\partial q_i} \quad \text{and} \quad E = -\frac{\partial S}{\partial t}.$$  \hspace{1cm} (7.33)

The function $S(q,t)$ is called Hamilton’s principal function. Substituting (7.33) into (7.8) yields

$$\frac{\partial S}{\partial t} = H\left(\frac{\partial S}{\partial q}, q, t\right),$$  \hspace{1cm} (7.34)
which is the analogue of (7.25). Equation (7.34) is called the *Hamilton-Jacobi equation*. Just as the solution of (7.25) is equivalent to the (global) solution of the ray equations (7.30), the solution of (7.34) is equivalent to the solution of the canonical equations (7.7). And just as (7.7) arise from the variational principle (7.10), the ray equations (7.30) are also equivalent to a variational principle.

In 1926, Erwin Schrödinger asked himself whether the analogy between the canonical equations and the ray equations for slowly varying waves might be more than an analogy. Suppose they are actually equivalent. That is, suppose that the canonical equations (7.7) are the ray equations corresponding to some underlying wave-equation for \( \psi(q,t) \). If the canonical equations correspond to a mechanical system, say a particle of mass \( m \) at \( q \) in a force-potential of \( V(q,t) \), then \( q \) has the same units as \( x \), but \( p \) has the units of \( m \, dq/dt \), which are different from those of \( k \). The units of \( E \) also differ from the units of \( \omega \). The most economical assumption reconciling the units is that

\[
p_i = \hbar k_i \quad \text{and} \quad E = \hbar \omega , \tag{7.35}
\]

where \( \hbar \) is a constant (Planck’s constant) with the appropriate dimensions. The canonical equations are now equivalent to the ray equations with

\[
\Omega(k,x,t) = \frac{1}{\hbar} H(\hbar k, x, t) . \tag{7.36}
\]

The underlying equation for \( \psi(x,t) \) can be deduced from the dispersion relation

\[
\omega = \frac{1}{\hbar} H(\hbar k, x, t) \tag{7.37}
\]

by replacing

\[
-i \omega \rightarrow \frac{\partial}{\partial \hbar} \quad \text{and} \quad i k \rightarrow \nabla . \tag{7.38}
\]

For a single particle moving in three dimensions,

\[
H(p,q,t) = \frac{1}{2m} p \cdot p + V(q,t) , \tag{7.39}
\]

and we obtain the \( \psi \)-equation

\[
\hbar \frac{\partial \psi}{\partial t} = \left[ -\frac{\hbar^2}{2m} \nabla^2 + V \right] \psi . \tag{7.40}
\]

This is the simplest form of Schrödinger’s equation. Of course, nothing in this derivation gives the value of \( \hbar \) or the *interpretation* of \( \psi \) (which Schrödinger never accepted).
Now we recur to the other interpretation of the canonical equations, and regard (7.7) as the equations for a $2N$-dimensional trajectory in phase space. We consider a collection of moving phase-space points, each point moving at the velocity $(\dot{p}, \dot{q})$ given by (7.7). Each moving point represents an evolving physical system. Let

$$P(p, q, t) dp \, dq$$  \hspace{1cm} (7.41)

be the number of points in phase-space volume $dpdq$ at time $t$. Since the points representing systems can neither be created nor destroyed,

$$\frac{\partial P}{\partial t} + \frac{\partial}{\partial p_i}(\dot{p}_i, P) + \frac{\partial}{\partial q_i}(\dot{q}_i, P) = 0.$$  \hspace{1cm} (7.42)

Equation (7.42) is the $N$-dimensional analogue of the continuity equation,

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} \cdot (\nabla \rho) = 0,$$  \hspace{1cm} (7.43)

expressing the conservation of fluid particles in ordinary three-dimensional space. Since, by (7.17), the flow in phase space is nondivergent, (7.42) reduces to Liouville’s equation,

$$\frac{\partial P}{\partial t} + \dot{p}_i \frac{\partial P}{\partial p_i} + \dot{q}_i \frac{\partial P}{\partial q_i} = 0.$$  \hspace{1cm} (7.44)

In Chapter 5, we saw how Liouville’s equation forms the basis for equilibrium statistical mechanics.

To find the wave-analogue of (7.42) and (7.44), imagine that the slowly varying wave-field can be represented as a superposition of wave-packets. Each wave-packet is characterized by a location $x(t)$ in $N$-dimensional space and by a (carrier) wavenumber $k(t)$. Thus each wave-packet corresponds to a point in $2N$-dimensional $x$-$k$ space. The wave-packets are the analogues of the moving points in phase space. Let

$$n(k, x, t) dk \, dx$$  \hspace{1cm} (7.45)

be the number of wave-packets in $2N$-dimensional volume $dk \, dx$ at time $t$. Then, in complete analogy with (7.42),

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial k_i}(\dot{k}_i, n) + \frac{\partial}{\partial x_i}(\dot{x}_i, n) = 0,$$  \hspace{1cm} (7.46)

for as long as each wave-packet maintains its identity. By the ray equations (7.30),

$$\frac{\partial k_i}{\partial k_i} + \frac{\partial \dot{x}_i}{\partial x_i} = 0.$$  \hspace{1cm} (7.47)
Hence (7.46) reduces to the analogue of (7.44),

\[
\frac{\partial n}{\partial t} + \dot{k}_i \frac{\partial n}{\partial k_i} + \dot{x}_i \frac{\partial n}{\partial x_i} = 0. 
\]  

(7.48)

Now let

\[ E(k, x, t) \, dk \, dx \]  

(7.49)

be the energy within \( dk \) of \( k \) and \( dx \) of \( x \). That is, let \( E(k, x, t) \) be the slowly-varying wavenumber spectrum of the \( N \)-dimensional waves. Suppose we decompose the wave-field into packets in such a way that each wave-packet initially has the same amount of energy. In principle, this is always possible; the decomposition into wave-packets is very non-unique. If there is no \( N \)-dimensional mean flow, then each wave-packet conserves its energy. In that case, \( E(k, x, t) \) is proportional to \( n(k, x, t) \), and (7.46) and (7.48) imply that

\[
\frac{\partial E}{\partial t} + \dot{k} \cdot \frac{\partial E}{\partial k} + \dot{x} \frac{\partial E}{\partial x} = 0 
\]  

(7.50)

and

\[
\frac{\partial E}{\partial t} + \dot{k} \cdot \frac{\partial E}{\partial k} + \dot{x} \frac{\partial E}{\partial x} = 0. 
\]  

(7.51)

This is the evolution equation for the wavenumber spectrum of a slowly varying wave field. If the spectrum is monochromatic, that is, if

\[ E(k, x, t) = \delta(k - k_0(x, t))E(x, t), \]  

(7.52)

then the integral of (7.50) over all \( N \) wavenumber dimensions yields the familiar equation,

\[
\frac{\partial E(x, t)}{\partial t} + \dot{C}(x, t)E(x, t) = 0 
\]  

(7.53)

for the energy in a slowly varying wavetrain. Here

\[ C \equiv \frac{\partial \Omega(k_0, x, t)}{\partial k} \]  

(7.54)

is the local group velocity evaluated at the carrier wavenumber. If the medium of the slowly varying wave-field varies in space or time, then the energy of each wave-packet is not generally conserved. However, if the medium varies slowly, each wave-packet
conserves its wave-action, and (7.50-54) hold with the energy spectrum replaced by the action spectrum.

8. **Eulerian forms of Hamilton’s principle**

We can regard the fluid motion as a time-dependent mapping,

\[ x = x(a, \tau), \]  

from \( a \)-space into \( x \)-space. Hamilton’s principle requires that the action,

\[
\int d\tau \int \int da \left\{ \frac{\partial x}{\partial \tau} \cdot \frac{\partial x}{\partial \tau} - E \left( \frac{\partial (x)}{\partial (a)} \cdot \eta(a) \right) - \Phi(x) \right\}. \tag{8.2}
\]

be stationary with respect to arbitrary, infinitesimal variations \( \delta x(a, \tau) \) in this mapping. However, each forward mapping (8.1) uniquely determines the inverse mapping

\[ a = a(x, t) \]  

from \( x \)-space back into \( a \)-space. Thus Hamilton’s principle must be equivalent to the statement that the action be stationary for arbitrary variations \( \delta a(x, t) \) in this inverse mapping. This simple interchange between dependent and independent variables leads to the various **Eulerian forms** of Hamilton’s principle.

By this reasoning, Hamilton’s principle must be equivalent to the requirement that

\[
\int dt \int \int dx \left\{ \frac{\partial a}{\partial x} \left( \frac{\partial x}{\partial \tau} \cdot \frac{\partial x}{\partial \tau} - E \left( \frac{\partial (x)}{\partial (a)} \cdot \eta(a) \right) - \Phi(x) \right) \right\}. \tag{8.4}
\]

be stationary with respect to arbitrary variations \( \delta a(x, t) \). To carry out these variations, we must express the integrand of (8.4) entirely in terms of \( a(x, t) \) and its derivatives. To express the velocity

\[ v \equiv \frac{\partial x}{\partial \tau} \]  

as derivatives of \( a \), we solve the three equations

\[
\left( \frac{\partial}{\partial \tau} + v \cdot \nabla \right) a_i = 0, \quad i = 1, 2, 3
\]  

for the three components of \( v \) and substitute the results back into (8.4). Here \( a = (a_1, a_2, a_3) = (a, b, c) \). Equation (8.6) just restates the fact that \( a \) are particle labels. Equivalently, we can append the three equations (8.6) as constraints on (8.4), and then
vary \( a(x,t) \) and \( v(x,t) \) independently. Thus Hamilton’s principle is equivalent to the requirement that
\[
\int dt \int \int d^3x \frac{\partial}{\partial(x)} \left\{ \frac{1}{2} v \cdot v - E \left( \frac{\partial}{\partial(a)} \cdot \eta(a) \right) - \Phi(x) - A \cdot \frac{Da}{Dt} \right\} \tag{8.7}
\]
be stationary for arbitrary variations \( \delta a(x,t) \), \( \delta v(x,t) \) and \( \delta A(x,t) \). Here \( A=(A_1, A_2, A_3)=(A,B,C) \) are the Lagrange multipliers corresponding to the Lin constraints (8.6).

But look what a mess we have made! The density
\[
\rho = \frac{\partial}{\partial(x)} \tag{8.8}
\]
now multiplies every term in (8.7), so the variations \( \delta a \) are going to be a complicated affair. Furthermore, the variations of \( a \) now affect the entropy \( \eta(a(x,t)) \).

To mitigate these sources of complexity, we make two further modifications to (8.7). First, we take the entropy itself to be one of our fluid-particle labels, say \( c=\eta \). (This won’t work if the fluid is homentropic, but if the fluid is homentropic we won’t need to do it!) Second, we append the time-derivative of (8.8), namely
\[
\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0, \tag{8.9}
\]
as a fourth constraint on the Lagrangian. This allows us to vary \( \rho \) independently, in the same way that the Lin constraints (8.6) allow us to vary \( v \) independently. Hamilton’s principle now requires that
\[
\int dt \int \int d^3x \left\{ \frac{1}{2} \rho v \cdot v - \rho E \left( \frac{1}{\rho}, \eta \right) - \rho \Phi(x) - \rho A \cdot \frac{Da}{Dt} + \phi \left( \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) \right) \right\} \tag{8.10}
\]
be stationary for arbitrary variations in \( \rho(x,t) \), \( v(x,t) \), \( \eta(x,t) \), \( A(x,t) \), \( a(x,t) \) and \( \phi(x,t) \) (with \( c=\eta \)). Here \( \phi \) is the Lagrange multiplier corresponding to the constraint (8.9). However, the four constraints (8.6) and (8.9) are not independent. In fact, (8.6) imply (8.9), the time-derivative of (8.8). Hence, one of these four constraints can be dropped, and, if our purpose is to cast Hamilton’s principle into a form involving as many familiar Eulerian variables as possible, then it is sensible to keep the \( \phi \)-constraint and drop one of the \( A \)-constraints. Thus, dropping the \( B \)-constraint, we obtain (after an integration by parts)
\[
\int dt \int \int d^3x \left\{ \frac{1}{2} \rho v \cdot v - \rho E \left( \frac{1}{\rho}, \eta \right) - \rho \Phi(x) - \rho A \frac{Da}{Dt} - \rho C \frac{D\eta}{Dt} - \rho D \frac{D\phi}{Dt} \right\}. \tag{8.11}
\]
Hamilton’s principle now requires that (8.11) be stationary for variations in \( \rho(x,t), v(x,t), \eta(x,t), A(x,t), a(x,t), C(x,t) \) and \( \phi(x,t) \).

This is almost as good as anyone could do! However, we can use the equations,

\[
\delta v : \quad v = A\nabla a + C\nabla \eta + \nabla \phi, \tag{8.12}
\]

obtained by varying \( v \), to eliminate \( v \) from (8.11). After some cancellations (and apart from an irrelevant sign-change in the whole expression), we obtain

\[
\int dt \left\{ \int \int \int dx \left( \rho A \frac{\partial a}{\partial t} + \rho C \frac{\partial \eta}{\partial t} + \rho \frac{\partial \phi}{\partial t} \right) + H \right\}, \tag{8.13}
\]

where

\[
H[\rho, A,a,C,\eta,\phi] = \int \int \int dx \left\{ \rho v \cdot v + \rho E \left( \frac{1}{\rho} \right) + \rho \Phi \right\}, \tag{8.14}
\]

and \( v \) in (8.14) is simply an abbreviation for the right-hand side of (8.12). Hamilton’s principle now requires that (8.13) be stationary with respect to variations of \( \rho(x,t), A(x,t), a(x,t), C(x,t), \eta(x,t) \) and \( \phi(x,t) \).

Note that (8.13) fits the extended form (7.10) of Hamilton’s principle with \( a, \eta, \) and \( \phi \) corresponding to the generalized coordinates \( q_i(t) \), and \( -\rho A, -\rho C \), and \( -\rho \) corresponding to the generalized momenta \( p_i(t) \). Thus (8.13) is the \( x \)-space analogue of the integral in (7.16). Note, too, that (8.13) has as many fields — six — as does the corresponding particle-mechanics form of Hamilton’s principle (for which the fields are \( x(a,t) \) and \( u(a,t) \)).

It is a healthy exercise in variational calculus to show that the variations of (8.13) yield the equations for a perfect fluid. We find that

\[
\delta A : \quad \frac{DA}{Dt} = 0, \quad \delta a : \quad \frac{DA}{Dt} = 0, \tag{8.15}
\]

\[
\delta C : \quad \frac{DC}{Dt} = 0, \quad \delta \eta : \quad \frac{DC}{Dt} = \frac{\partial}{\partial \eta} E \left( \frac{1}{\rho} \right) \equiv T, \tag{8.16}
\]

\[
\delta \phi : \quad \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0, \tag{8.17}
\]

and

\[
\delta \rho : \quad A \frac{\partial a}{\partial t} + C \frac{\partial \eta}{\partial t} + \frac{\partial \phi}{\partial t} + \frac{1}{2} v \cdot v + \Phi + E + \frac{\rho}{\rho} = 0, \tag{8.18}
\]
where
\[ p = -\frac{\partial}{\partial \alpha} E(\alpha, \eta) \tag{8.19} \]
is the pressure and \( \alpha \equiv \rho^{-1} \). The right-hand side of (8.16b) is the temperature. We can view (8.18) as a generalized Bernoulli equation.

To obtain the momentum equation from (8.15-19), we take the time-derivative of (8.12) and substitute from (8.15), (8.16), and (8.18). First note that (8.12) implies that
\[ \omega = \nabla \times v = \nabla A \times \nabla a + \nabla C \times \nabla \eta, \tag{8.20} \]
and hence
\[ \omega \times v = (v \cdot \nabla A) \nabla a - (v \cdot \nabla a) \nabla A + (v \cdot \nabla C) \nabla \eta - (v \cdot \nabla \eta) \nabla C. \tag{8.21} \]
The time-derivative of (8.12) is
\[
\frac{\partial v}{\partial t} = \frac{\partial A}{\partial t} \nabla a + \frac{\partial C}{\partial t} \nabla \eta + A \frac{\partial a}{\partial t} + C \frac{\partial \eta}{\partial t} + \nabla \frac{\partial \phi}{\partial t} \\
= \frac{\partial A}{\partial t} \nabla a + \frac{\partial C}{\partial t} \nabla \eta - \frac{\partial a}{\partial t} \nabla A - \frac{\partial \eta}{\partial t} \nabla C + \nabla \left( A \frac{\partial a}{\partial t} + C \frac{\partial \eta}{\partial t} + \frac{\partial \phi}{\partial t} \right) \\
= -(v \cdot \nabla A) \nabla a - (v \cdot \nabla C - T) \nabla \eta + (v \cdot \nabla a) \nabla A + (v \cdot \nabla \eta) \nabla C \\
- \nabla \left( \frac{1}{2} v \cdot v + \Phi + E + \frac{p}{\rho} \right) \\ = -(\omega \times v) - \frac{1}{\rho} \nabla p - \nabla \Phi. \tag{8.22} \]
By (8.21), (8.16b) and (8.19), this is
\[
\frac{\partial v}{\partial t} = -(\omega \times v) - \frac{1}{\rho} \nabla p - \nabla \Phi. \tag{8.23} \]

The Lagrangian (8.13) obviously has time-translation symmetry; the energy (8.14) is correspondingly conserved. But where in (8.13-14) is the particle-relabeling symmetry property? It seems to have disappeared, but it is actually present as a gauge symmetry of the Hamiltonian (8.14). In the general case of non-homentropic fluid, the four velocity potentials \( A, a, C, \) and \( \phi \) occur only in the expression (8.12) for the three components of \( v \). It must therefore be possible to vary these four potentials in a way that is not detectable by the Hamiltonian (8.14); this is a symmetry property.

We therefore consider variations of (8.13) in which
\[
\delta \rho = \delta \eta = 0, \tag{8.24} \]
and in which the remaining variations \( \delta a, \delta A, \delta C, \delta \phi \) do not affect the form
\[ A da + C d\eta + d\phi, \quad (8.25) \]

where \( d \) denotes the change arising from an arbitrary infinitesimal displacement in \( x \)-space. Variations having this property do not affect the velocity (8.12) and Hamiltonian (8.14). Now,

\[
\delta (A da + C d\eta + d\phi) = \delta A a + A \delta a + \delta C d\eta + d(\delta \phi)
\]

\[
= \delta A a - A \delta a + \delta C d\eta + d(\delta \phi + A \delta a)
\]

We seek variations for which (8.26) vanishes. Regarding \((A, a, \eta)\) as curvilinear coordinates in \( x \)-space, we see that (8.26) vanishes if

\[
\delta A = \frac{\partial G}{\partial a}, \quad \delta a = -\frac{\partial G}{\partial A}, \quad \delta C = \frac{\partial G}{\partial \eta}
\]

(8.27)

and

\[
\delta \phi = -G - A \delta a = -G + A \frac{\partial G}{\partial A}
\]

(8.28)

where \( G(A,a,\eta,\tau) \) is an arbitrary infinitesimal function. (Readers who are familiar with classical mechanics will recognize a close correspondence between this development and the classical theory of canonical transformations. Readers familiar with group theory will see a connection with the theory of symplectic transformations.)

For variations satisfying (8.24), (8.27) and (8.28), Hamilton’s principle implies that

\[
\int dt \left[ \iiint dx \left( \rho \frac{\partial \delta a}{\partial t} + \rho A \frac{\partial \delta a}{\partial t} + \rho \delta C \frac{\partial \eta}{\partial t} + \rho \frac{\partial \delta \phi}{\partial t} \right) \right] = 0.
\]

(8.29)

However, by (8.27) and (8.28) (and because the variations vanish at the endpoints in time), (8.29) is equivalent to

\[
\int dt \iiint dx \rho \frac{\partial}{\partial \tau} G(A,a,\eta,\tau) = 0.
\]

(8.30)

But (8.30) is in turn equivalent to

\[
0 = \int d\tau \iiint dA da d\eta \frac{\partial (x,y,z)}{\partial (A,a,\eta)} \rho \frac{\partial}{\partial \tau} G(A,a,\eta,\tau)
\]

\[
= -\int d\tau \iiint dA da d\eta \ G(A,a,\eta,\tau) \frac{\partial}{\partial \tau} \left[ \frac{\partial (x,y,z)}{\partial (A,a,\eta)} \rho \right]
\]

(8.31)
Then, since $G$ is an arbitrary function, we must have

$$\frac{DQ}{Dt} = 0,$$  \hspace{1cm} (8.32)

where

$$Q = \frac{1}{\rho} \frac{\partial(Aa,\eta)}{\partial (x,y,z)} = \frac{1}{\rho} (\nabla A \times \nabla a) \cdot \nabla \eta.$$  \hspace{1cm} (8.33)

By (8.20),

$$Q = \frac{1}{\rho} \omega \cdot \nabla \eta.$$  \hspace{1cm} (8.34)

As expected, the gauge symmetry corresponds to the conservation of potential vorticity.

Readers to whom the action (8.13-14) seems but a clumsy equivalent to the particle-mechanics version (7.16) share a strong prejudice of the writer. But (8.13-14) and its numerous alternative forms have become so firmly established that no introduction to Hamiltonian fluid mechanics could omit them entirely. Moreover, there are special situations in which the Eulerian forms of Hamilton’s principle are much simpler than the general particle-mechanics form; these are the cases of vanishing (or severely constrained) vorticity. For example, if we impose the constraints of uniform $a, A, C, \eta$ on (8.13-14), we obtain

$$\int dt \left[ \int \int \int d\mathbf{x} \rho \left( \frac{\partial \phi}{\partial t} + \frac{1}{2} \nabla \phi \cdot \nabla \phi + E \left( \frac{1}{\rho} \right) + \Phi (\mathbf{x}) \right) \right].$$  \hspace{1cm} (8.35)

(Since, when the fluid is homentropic, these constraints are consistent with (8.15-16), they represent a specialization rather than an approximation.) Requiring that (8.35) be stationary with respect to variations of (the remaining variables) $\rho$ and $\phi$ yields the equations for irrotational flow. Of course, irrotational flows represent only a subset of the solutions to the general perfect-fluid equations, but if you are interested in irrotational flows, then you cannot do better than (8.35).

In the general case of unconstrained vorticity, (8.20) shows that $a, A, C, \eta$ are vorticity labeling coordinates, scalars that move with the fluid but carry all the information needed to calculate the vorticity field. Thus it is not surprising that the elimination of these variables corresponds to the constraint of irrotational flow, or that the gauge symmetry associated with them corresponds to Ertel’s theorem.

9. The geometrical view of dynamics

In Section 7, we saw how the canonical equations form the basis for a helpful analogy between Hamiltonian dynamics and the theory of slowly varying waves. In this section,
we see how the same canonical equations form the basis for a geometrical formulation of dynamics, a formulation in which all the ingredients are geometrical objects with well-defined properties. It is clearly impossible, within a few pages, to develop this subject with complete thoroughness and rigor. However, it is quite possible, and very worthwhile, to sketch the main themes and ideas before going on to concrete examples that illustrate the enormous advantages of the geometrical viewpoint.

What is a geometrical object? Students first encounter vectors as \( n \)-tuples of numbers. Then, gradually, they come to regard vectors as geometrical objects (arrows!) having a physical meaning and reality that are independent of the numbers used to describe them. The notation of vector analysis (dot products, cross products, etc.) emphasizes this independent reality. The same vector corresponds to different \( n \)-tuples in different coordinate systems, but it is, after all, the same vector. To express Hamiltonian dynamics in a geometrical language, we must generalize the concept of a vector, and we must add a few more geometrical objects to our repertoire. But the goal is the same as in elementary vector analysis, namely, to express physical reality in the starkest and clearest way, in a way that de-emphasizes the role of coordinates and numbers.

Because of their simplicity, the canonical equations,

\[
\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad \frac{dq_i}{dt} = +\frac{\partial H}{\partial p_i}, \quad i = 1, 2, \ldots, N, \tag{9.1}
\]

are the ideal departure point for a geometrical theory of Hamiltonian dynamics. We begin by defining the Poisson bracket,

\[
\{A, B\} \equiv \sum_{i=1}^{N} \left( \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right), \tag{9.2}
\]

of any two functions \( A(p,q) \) and \( B(p,q) \). In terms of the Poisson bracket, the canonical equations take the form

\[
\frac{dp_i}{dt} = \{p_i, H\}, \quad \frac{dq_i}{dt} = \{q_i, H\}. \tag{9.3}
\]

More generally,

\[
\frac{dF}{dt} = \{F, H\}, \tag{9.4}
\]

where \( F(p,q) \) is any function of \( p \) and \( q \).

Equation (9.4) is the general evolution equation for a Hamiltonian system. It contains two geometrical objects: the Poisson bracket operator \( \{\ ,\ \} \) and the Hamiltonian \( H(p,q) \), a scalar. We are interested in the geometrical properties of these objects, that is, the properties that survive transformation from the canonical coordinates \((p,q)\) to arbitrary new coordinates.

With a view to coordinate transformations, we now define
\[ z = (z^1, z^2, \ldots, z^{2N}) \equiv (q_1, q_2, \ldots, q_N, p_1, p_2, \ldots, p_N). \] (9.5)

The superscript notation for coordinates is a well-established convention of tensor analysis. In terms of \( z \), the canonical equations (9.1) take the form

\[ \frac{dz^i}{dt} = J^{ij} \frac{\partial H}{\partial z^j}, \] (9.6)

where

\[ J = \begin{bmatrix} 0_N & I_N \\ -I_N & 0_N \end{bmatrix}, \] (9.7)

and repeated subscripts are summed from 1 to \( 2N \). In (9.7), \( 0_N \) is the \( N \times N \) matrix of zeroes, and \( I_N \) is the \( N \times N \) identity matrix. The definition (9.2) takes the form

\[ \{A, B\} \equiv \frac{\partial A}{\partial \bar{z}^i} J^{ij} \frac{\partial B}{\partial z^j}. \] (9.8)

The significance of (9.4) and (9.8) is that they are covariant with respect to transformations

\[ \bar{z}^i = \bar{z}^i(z) \] (9.9)

from the canonical coordinates \( z \) to new coordinates \( \bar{z} \). In this section we are interested in coordinate transformations that are invertible. We thus require that the transformation (9.9) obey

\[ \det \left( \frac{\partial \bar{z}^i}{\partial z^j} \right) \neq 0, \] (9.10)

but apart from (9.10) the new coordinates are arbitrary. In the following section, we consider non-invertible transformations. By covariant we mean that the equations (9.4) and (9.8) take the same form in the new coordinates as in the old coordinates provided that \( H \) follows the transformation rule for a scalar, and that \( J^{ij} \) follows the transformation rule for a contravariant tensor. We now pause to explain these terms.\(^{16}\)

In tensor analysis there are two kinds of vectors: contravariant vectors (or simply vectors), with components denoted by superscripts, following the transformation rule,

\[ v^i = \frac{\partial \bar{z}^i}{\partial z^j} v^j, \] (9.11)
and covariant vectors (or simply covectors), with components denoted by subscripts, following the transformation rule,

\[ \tilde{v}_i = \frac{\partial z^j}{\partial \tilde{z}^i} v_j. \]  

(9.12)

The overbars denote the vector (or covector) components in the new, \( \tilde{z} \)-coordinates. A tensor is contravariant if it transforms like the product of contravariant vectors. To see that \( J^{ij} \) is a contravariant tensor (and hence deserves its superscripts), we transform the definition (9.8) into new coordinates. Using the chain rule, we find that

\[ \{A, B\} = \frac{\partial A}{\partial \tilde{z}^m} \frac{\partial m}{\partial \tilde{z}^i} J^{ij} \frac{\partial B}{\partial \tilde{z}^n} \frac{\partial n}{\partial \tilde{z}^j} = \frac{\partial A}{\partial \tilde{z}^m} \frac{\partial B}{\partial \tilde{z}^n} \tilde{J}^{mn}, \]  

(9.13)

where

\[ \tilde{J}^{mn} = \frac{\partial \tilde{z}^m}{\partial \tilde{z}^i} J^{ij} \frac{\partial \tilde{z}^n}{\partial \tilde{z}^j}. \]  

(9.14)

The transformation rule (9.14) for \( J \) clearly follows the rule (9.11) in each index. Hence \( J \) is a contravariant tensor.

Now we state the geometrical properties of \( H \) and \{ , \}. For \( H \) this is easy; a scalar is the simplest of geometrical objects. The Hamiltonian \( \tilde{H} \) in new coordinates has the same value as the Hamiltonian \( H \) at the corresponding point in the old coordinates.

The geometrical properties of the Poisson bracket \{ , \} are more interesting. There are three:

(a) Nonsingularity. For coordinate transformations satisfying (9.10), the following property holds in the new coordinates: If \( A \) and \( B \) are independent, nonconstant functions, then \( \{A, B\} \) cannot vanish uniformly,

\[ \{A, B\} \neq 0. \]  

(9.15)

(b) Antisymmetry. In all coordinates,

\[ \{A, B\} = -\{B, A\}. \]  

(9.16)

(c) Jacobi property. In all coordinates,

\[ \{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0. \]  

(9.17)

To establish these three properties, it is only necessary to show that they hold in the canonical coordinates \( (p,q) \) (an easy matter), and that the properties (9.15-17) are themselves covariant.

First we restate the properties (9.15-17) in terms of \( J \), the contravariant tensor appearing in the definition (9.8) of \{ , \}. In terms of \( J \), the properties are:
(a) $J$ is nonsingular,
\[
\det(J^{ij}) \neq 0.
\] (9.18)

(b) $J$ is antisymmetric,
\[
J^{ij} = -J^{ji}.
\] (9.19)

(c) $J$ obeys the Jacobi relation,
\[
j^{im} \frac{\partial J^{jk}}{\partial x^m} + J^{im} \frac{\partial J^{ki}}{\partial x^m} + J^{km} \frac{\partial J^{ij}}{\partial x^m} = 0.
\] (9.20)

To obtain (9.20), substitute the definition (9.8) into (9.17).

Now, it is obvious that properties (a-c) hold in the canonical coordinates (9.5), in which the components of $J^{ij}$ are constants. To establish covariance, we must show that if (a-c) hold in one set of coordinates, then they hold in any other set of coordinates. We show this by transforming (9.18-20) into new coordinates, using the rule (9.14) for the transformation of $J$. We find that (9.18) holds in new coordinates satisfying (9.10), while (9.19) and (9.20) hold in all new coordinates. Readers familiar with tensor analysis will recognize that the three terms in (9.20) are individually noncovariant; they contain ordinary, noncovariant derivatives. However, as the reader will verify, the worrisome terms all cancel out, and (9.20) holds with $J$ replaced by $\tilde{J}$.

The dynamical equation (9.4,9.8), with the properties (a-c), is all there is to Hamiltonian dynamics. That is, given a scalar $H(z)$ and a contravariant tensor $J^{ij}(z)$ with the properties (9.18-20), we have a Hamiltonian system. Everything that can be proved about general Hamiltonian systems can be proved, using tensor analysis, from the properties (9.18-20). In particular, if $J$ satisfies (9.18-20), then, according to Darboux’s theorem, it is always possible to transform the dynamics to canonical coordinates in which $J$ takes the simple form (9.7).17 Moreover, the canonical coordinates are non-unique; different sets of canonical coordinates are related to one another by canonical transformations. A canonical transformation is a transformation that leaves the components of $J$ unchanged. In particular, if $J$ has the form (9.7) corresponding to canonical coordinates, then $J$ has this same simple form after a canonical transformation to new (canonical) coordinates.

There is an interesting analogy between the geometry of phase space as determined by $J^{ij}$ (sometimes called symplectic geometry), and the Euclidean geometry of ordinary space as determined by the metric tensor $g_{ij}$. Just as canonical coordinates exist if $J^{ij}$ satisfies (9.18-20), Cartesian coordinates exist if $g_{ij}$ is nonsingular, symmetric, and implies vanishing curvature. In particular, (9.20) is analogous to the vanishing of the curvature tensor. The metric tensor takes its simplest form, $g_{ij} = I$ (where $I$ is the identity tensor), in Cartesian coordinates. The different sets of Cartesian coordinates are related by orthogonal transformations. Thus canonical transformations in phase space are analogous to rigid rotations in ordinary Euclidean space.
The geometrical viewpoint suggests a general strategy for analyzing Hamiltonian systems: to seek the coordinates $z$ in which the two geometrical objects $H(z)$ and $J^{ij}(z)$ take their simplest mathematical form. The tensor $J^{ij}(z)$ takes its simplest form in canonical coordinates; it is thus useless to look for anything simpler than (9.7). On the other hand, the scalar $H(z)$ may not take a particularly simple form in any set of canonical coordinates. In that case we may be forced to compromise between simplifying $J^{ij}$ and simplifying $H$.

Once again, the analogy with Euclidean geometry is illuminating. Suppose we want to compute the area and volume of a solid in ordinary space. The solid is analogous to the Hamiltonian. Suppose the solid is a sphere. Then the solid is most easily described in spherical coordinates, which are non-Cartesian. However, the metric tensor (the analogue of $J$) takes its simplest form in Cartesian coordinates, and if we use spherical coordinates to conform to the solid, the formulas for area and volume will be slightly more complicated than they would be in Cartesian coordinates.

This is exactly the sort of compromise we face in Hamiltonian fluid dynamics. The Poisson bracket takes its simplest form in Lagrangian variables, but the Hamiltonian takes its simplest form in the more familiar Eulerian variables. If one chooses to simplify $H$, one pays the price in $\{ , \}$, but, as the next few sections will show, the results sometimes justify the expense.

10. Non-canonical Hamiltonian Dynamics

Let

$$\frac{d\dot{z}^i}{dt} = J^{ij} \frac{\partial H}{\partial z^j}, \quad i = 1, \ldots, 2N \quad (10.1)$$

be the equations for a Hamiltonian system with Hamiltonian $H(z)$. The $2N \times 2N$ symplectic tensor $J^{ij}$ has the general properties (9.18-20). In particular, $J^{ij}$ is nonsingular. Now, let

$$\{\zeta^\alpha\}, \quad \alpha = 1, 2, \ldots, n < 2N \quad (10.2)$$

be a subset of the $2N$ coordinates in (10.1), and suppose there exists a closed set of evolution equations involving only the members of this subset:

$$\frac{d\dot{z}^\alpha}{dt} = f^\alpha(\dot{z}^\beta), \quad \alpha = 1, \ldots, n. \quad (10.3)$$

We use Roman superscripts to denote the complete set of $2N$ coordinates, and Greek superscripts to denote the coordinates in the subset. We let $\{\zeta^i\}$ denote the members of $\{z^i\}$ that are not in $\{z^\alpha\}$.

Now, (10.3) must be equivalent to
\[
\frac{dz^\alpha}{dt} = J^{\alpha\beta} \frac{\partial H}{\partial z^\beta}, \quad \alpha = 1, \ldots, n
\]  

(10.4)

where \( J^{\alpha\beta} \) is an \( n \times n \) tensor whose components are a subset of the components of \( J^{ij} \). In the derivatives in (10.4), the \( \zeta^i \) are always held fixed. By hypothesis, the right-hand side of (10.4) is independent of the \( \zeta^i \). Now we slightly strengthen that hypothesis, and assume that both \( H \) and \( J^{\alpha\beta} \) depend only on the \( z^\alpha \). The dependence of the Hamiltonian on the subset \( \{z^\alpha\} \) of dynamical variables could be viewed as the original motivation for the division of \( \{z^i\} \) into subsets.

Now we forget all about the \( \zeta^i \) and regard (10.4) as an evolution equation in the \( n \)-dimensional phase space spanned by the \( z^\alpha \). Is this system Hamiltonian? That depends on the properties of the \( J^{\alpha\beta} \). It is easy to see that \( J^{\alpha\beta} \) inherits the antisymmetry property (9.19) from \( J^{ij} \). Similarly, \( J^{\alpha\beta} \) must have the Jacobi property. To see this, just let the superscripts \( i,j,k \) in (9.20) take the Greek values \( \alpha, \beta, \gamma \), and remember that \( \partial J^{\beta\gamma}/\partial z^m = 0 \) for non-Greek \( m \).

The nonsingularity property (9.18) has a different fate; in general, the determinant of \( J^{\alpha\beta} \) will vanish. This always happens when \( n \) is odd, because the determinant of an odd-order antisymmetric matrix always vanishes.

If \( J^{\alpha\beta} \) is singular, then the hypothesis of Darboux’s theorem is not satisfied, and canonical coordinates do not exist. It is then impossible to transform \( J^{\alpha\beta} \) into the \( (n \times n) \) analogue of the form (9.7). However, a generalization of Darboux’s theorem allows \( J^{\alpha\beta} \) to be put into the form

\[
\begin{bmatrix}
J^{\alpha\beta}
\end{bmatrix} = \begin{bmatrix}
0_M & I_M & 0 \\
-I_M & 0_M & 0 \\
0 & 0 & 0_{n-2M}
\end{bmatrix}
\]  

(10.5)

where \( 2M < n \) is the rank of \( J^{\alpha\beta} \). (The coordinates \( z^\alpha \) in which \( J^{\alpha\beta} \) takes its simplest form (10.5) are still typically not the coordinates in which \( H(z^\alpha) \) is simplest.)

Systems governed by equations of the form (10.4) in which \( J^{\alpha\beta} \) is antisymmetric, obeys the Jacobi property, but is singular, are called non-canonical Hamiltonian systems. In fluid mechanics, noncanonical Hamiltonian systems arise from the transformation from Lagrangian variables to Eulerian variables. The Lagrangian variables,

\[
u(a, \tau), \quad x(a, \tau),
\]  

(10.6)

are the velocities and locations of marked fluid particles. The Eulerian variables

\[
v(x, t), \quad \rho(x, t), \quad \eta(x, t),
\]  

(10.7)

are the velocity, density and entropy at fixed locations. The Lagrangian variables (10.6) are the analogues of the complete set of coordinates \( \{z^i\} \) in (10.1). If we know (10.6), we know everything there is to know about the motion of the fluid. The Eulerian
variables (10.7) are the analogues of \( \{ z^\alpha \} \) in (10.4). If we know (10.7), then we know everything we need to compute the Hamiltonian,

\[
H = \iiint \mathbf{dx} \left\{ \rho \mathbf{v} \cdot \mathbf{v} + \rho \mathbf{E} \left( \frac{1}{\rho} \right) + \rho \Phi(\mathbf{x}) \right\},
\]

(10.8)

but our knowledge of the fluid motion is incomplete; we cannot say which fluid particle went where. However, as we are already well aware, we can write the equations of fluid motion in a form that involves only the Eulerian variables (10.7). This is possible because the physics doesn’t care how we label the fluid particles. Thus the same particle-relabeling symmetry property that leads to the general vorticity-conservation law (2.25) also accounts for the (reduced) Eulerian formulation of fluid mechanics.

If \( J^{\alpha\beta} \) is singular with corank \( K \), then there are \( K \) independent covectors \( v^{(k)}_\beta \) such that

\[
J^{\alpha\beta} v^{(k)}_\beta = 0, \quad k = 1, \ldots, K.
\]

(10.9)

In the coordinates for which \( J^{\alpha\beta} \) takes the form (10.5), these covectors are just the last \( K \) unit vectors. In any set of coordinates,

\[
v^{(k)}_\beta = \frac{\partial C^{(k)}}{\partial z^\beta}, \quad k = 1, \ldots, K,
\]

(10.10)

where \( \{ C^{(k)} \} \) is a set of \( K \) independent scalars. That is, the null covectors of \( J^{\alpha\beta} \) are gradients. This has important consequences.

To prove (10.10), we must prove that

\[
J^{\alpha\beta} \frac{\partial C}{\partial z^\beta} = 0, \quad \alpha = 1, \ldots, n
\]

(10.11)

has \( K \) independent solutions, \( C^{(k)}(z^\alpha) \).

For this we use a standard result from the theory of first-order differential equations. Let

\[
D^{(\alpha)} \equiv J^{\alpha\beta} \frac{\partial}{\partial z^\beta}
\]

(10.12)

be the directional derivative in the direction of

\[
(J^{\alpha_1}, J^{\alpha_2}, \ldots, J^{\alpha_n}).
\]

(10.13)

(Mathematicians prefer to call \( D^{(\alpha)} \) a tangent vector.) Then (10.11) is equivalent to

\[
D^{(\alpha)} C = 0, \quad \alpha = 1, \ldots, n.
\]

(10.14)
If $J^{\alpha\beta}$ is nonsingular, then the vectors (10.13) point in $n$ different directions, and $C=const$ is the only solution of (10.14). If $J^{\alpha\beta}$ is singular with corank $K$, then (10.13) locally span $n-K$ dimensions. This does not necessarily mean that (10.14) has nontrivial solutions, because the integral curves corresponding to the $D^{(\alpha)}$ could still fill up the $n$-dimensional space. By the integral curve corresponding to $D^{(\alpha)}$, we simply mean the path,

$$\frac{dz^\beta}{dt} = J^{\alpha\beta},$$

(10.15)

traced by a particle always moving in the direction of (10.13). However, according to the Frobenius theorem, these integral curves all lie on an $(n-K)$-dimensional surface,

$$C^{(1)} = C^{(2)} = \cdots = C^{(K)} = 0,$$

(10.16)

if the $D^{(\alpha)}$ have the property,

$$D^{(\alpha)}D^{(\beta)} - D^{(\beta)}D^{(\alpha)} = \sum_\gamma f_\gamma D^{(\gamma)},$$

(10.17)

that the commutator of any two is a linear combination of the rest. The $f_\gamma$ are any functions of location in phase space. The property (10.17) is nontrivial, because the $D^{(\alpha)}$ do not span all $n$ dimensions. The $D^{(\alpha)}$ defined by (10.12) satisfy (10.17) on account of the Jacobi property,

$$J^{\alpha\delta} \frac{\partial J^{\beta\nu}}{\partial z^\delta} + J^{\beta\delta} \frac{\partial J^{\alpha\nu}}{\partial z^\delta} + J^{\gamma\delta} \frac{\partial J^{\alpha\beta}}{\partial z^\delta} = 0.$$

(10.18)

Using (10.18), we find that (10.17) holds with

$$f_\gamma = \frac{\partial J^{\gamma\beta}}{\partial z^\gamma}$$

(10.19)

Hence there exist $K$ independent $C^{(k)}$ satisfying (10.11). Equivalently, the general solution of (10.11) is

$$C = F(C^{(1)}, C^{(2)}, \ldots, C^{(K)})$$

(10.20)

where each $C^{(k)}$ is arbitrary by an additive constant. The functions $C^{(k)}$ are called Casimirs.

Now we summarize our results. Every Hamiltonian system consists of a phase space with two geometrical objects, a bilinear operator $\{ , \}$ called the Poisson bracket, and the Hamiltonian $H$, a scalar. The Poisson bracket always has the antisymmetry property,

$$\{A,B\} = -\{B,A\},$$

(10.21)
and the Jacobi property,

$$\{A,\{B,C\}\} + \{B,\{C,A\}\} + \{C,\{A,B\}\} = 0.$$  \hspace{1cm} (10.22)

However, it may happen that the Poisson bracket is singular. Then there exists a set of Casimir functions for which

$$\{C,A\} = 0$$ \hspace{1cm} (10.23)

for \textit{any} function \(A\). In particular,

$$\{C,H\} = 0,$$ \hspace{1cm} (10.24)

and hence \(C\) is conserved by the dynamics. In a sense, this conservation law is trivial, because (10.24) holds for any \(H\) whatsoever.

Singular Poisson brackets typically arise from a transformation from canonical coordinates (in which the bracket is nonsingular) to a \textit{reduced} set of (fewer) coordinates in which the dynamics comprises a fewer number of equations but is nevertheless closed. Then (although we shall not prove this general fact) the Casimirs are the conserved quantities corresponding to the symmetries that permit the reduction. In the case of the fluid, the reduction corresponds to a transformation from Lagrangian to Eulerian variables. The Eulerian equations form a closed set on account of the particle-relabeling symmetry property of the full, Lagrangian dynamics. Therefore we shall not be surprised to find that the Casimirs of the fluid include the potential vorticity.
11. Poisson brackets for fluids

Now we compute the Poisson brackets for the fluid, in the reduced phase-space of Eulerian variables. There are two general ways to proceed. We can start in Lagrangian variables, in which the bracket is known and takes a simple form, and simply transform everything into Eulerian variables; or, we can obtain the bracket directly in Eulerian coordinates by a mixture of inspiration and guesswork. The first method is fool-proof, but it can be rather tedious. The second method is usually the quickest, but it contains an element of risk. We illustrate both methods.

As an example of the first method — direct transformation from Lagrangian to Eulerian variables — we consider the general equations for an unbounded perfect fluid in one dimension. The generalization to three dimensions is straightforward. In Lagrangian variables, the velocities \( u(a, \tau) \) and locations \( x(a, \tau) \) of labeled fluid particles constitute canonical coordinates (see Section 7). Hence the analogue of (9.2) is

\[
\{A, B\} = \int da \left( \frac{\delta A}{\delta x(a)} \frac{\delta B}{\delta u(a)} - \frac{\delta A}{\delta u(a)} \frac{\delta B}{\delta x(a)} \right),
\]

(11.1)

where \( A \) and \( B \) are any two functionals of \( u(a) \) and \( x(a) \). (We suppress the time-argument wherever possible; all functional derivatives are taken at a fixed time.) We assume for simplicity that the fluid is homentropic. Then the Hamiltonian is

\[
H[u(a), x(a)] = \int da \left\{ \frac{1}{2} u(a)^2 + E \left( \frac{dx}{da} \right) + \Phi(x(a)) \right\}.
\]

(11.2)

The general evolution equation is

\[
\frac{dF}{dt} = \{F, H\},
\]

(11.3)

where \( F[u(a), x(a)] \) is any functional of the canonical coordinates. To verify (11.3), it suffices to consider the cases \( F = u(a) \) and \( F = x(a) \). For the latter, we obtain

\[
\frac{\delta x(a, \tau)}{\partial \tau} = \{x(a), H\} = \int da' \left( \frac{\delta x(a)}{\delta x(a')} \frac{\delta H}{\delta u(a')} - \frac{\delta x(a)}{\delta u(a')} \frac{\delta H}{\delta x(a')} \right)
\]

\[
= \int da' \left( \frac{\delta x(a)}{\delta x(a')} \frac{\delta H}{\delta u(a')} \right)
\]

(11.4)

since \( x(a) \) and \( u(a) \) are independent coordinates. Since

\[
x(a) = \int da' x(a') \delta(a - a'),
\]

(11.5)

we have
\[
\frac{\delta x(a)}{\delta x(a')} = \delta(a - a').
\] (11.6)

By (11.2),
\[
\frac{\delta H}{\delta u(a')} = u(a').
\] (11.7)

Thus (11.4) becomes
\[
\frac{\partial x(a, \tau)}{\partial \tau} = \int da' \delta(a - a')u(a') = u(a, \tau).
\] (11.8)

By similar steps,
\[
\frac{\partial u(a, \tau)}{\partial \tau} = \{u(a), H\} - \int da' \left( -\frac{\delta u(a)}{\delta u(a')} \frac{\delta H}{\delta x(a')} \right) = -\frac{dx}{da} \frac{\partial p}{\partial x} - \frac{\partial \Phi}{\partial x},
\] (11.9)

where \( p = -dE(\alpha)/d\alpha \).

To transform (11.1) into Eulerian coordinates \( u(x) \) and \( \rho(x) \), we use the chain rule for functional derivatives,
\[
\frac{\delta A}{\delta x(a)} = \int dx' \left\{ \frac{\delta A}{\delta u(x')} \frac{\delta u(x')}{\delta x(a)} + \frac{\delta A}{\delta \rho(x')} \frac{\delta \rho(x')}{\delta x(a)} \right\}.
\] (11.10)

Similarly,
\[
\frac{\delta A}{\delta u(a)} = \int dx' \left\{ \frac{\delta A}{\delta u(x')} \frac{\delta u(x')}{\delta u(a)} + \frac{\delta A}{\delta \rho(x')} \frac{\delta \rho(x')}{\delta u(a)} \right\}.
\] (11.11)

To compute the functional derivatives of \( u(x') \) and \( \rho(x') \), we must express \( u(x') \) and \( \rho(x') \) as functionals of \( x(a) \) and \( u(a) \). From
\[
u(x') = \int da \ u(a) \delta(a - a'),
\] (11.12)

(where \( a' \) is the label corresponding to \( x' \)), we find that
\[
\frac{\delta u(x')}{\delta x(a)} = 0 \quad \text{and} \quad \frac{\delta u(x')}{\delta u(a)} = \delta(a - a').
\] (11.13)
\[ \rho(x') = \int dx \, \rho(x) \delta(x - x') = \int da \, \delta(x(a) - x(a')) \]  

implies that

\[ \frac{\delta \rho(x')}{\delta x(a)} = \delta' (x - x') \quad \text{and} \quad \frac{\delta \rho(x')}{\delta u(a)} = 0, \]  

where \( \delta' \) is the derivative of the delta function. Thus (11.10) and (11.11) become

\[ \frac{\delta A}{\delta x(a)} = \int dx' \frac{\delta A}{\delta \rho(x')} \delta'(x - x') = \int dx' \frac{\partial}{\partial x} \left( \frac{\delta A}{\delta \rho(x)} \right) \]  

and

\[ \frac{\delta A}{\delta u(a)} = \int dx' \frac{\delta A}{\delta u(x')} \delta(a - u') = \int da' \frac{1}{\rho'} \frac{\delta A}{\delta \rho(\rho)} \delta(a - u') = \frac{\delta A}{\rho \delta \rho(\rho)}. \]  

Substituting (11.16) and (11.17) into (11.1), we finally obtain the Poisson bracket

\[ \{A, B\} = \int dx \left( \frac{\partial}{\partial x} \frac{\delta A}{\delta \rho} \frac{\delta B}{\delta u} - \frac{\partial}{\partial x} \frac{\delta B}{\delta \rho} \frac{\delta A}{\delta u} \right) \]  

for any two Eulerian functionals \( A[\rho(x), u(x)] \) and \( B[\rho(x), u(x)] \).

The bracket (11.18) is obviously antisymmetric. It is much less obvious that it satisfies the Jacobi property. However, (11.18) must satisfy the Jacobi property because (11.1) obviously does, and because the Jacobi property is covariant, that is, unaffected by transformations of the variables.

We can check (11.18) by showing that it gives the correct evolution equations for \( u(x) \) and \( \rho(x) \). Writing the Hamiltonian

\[ H = \int dx \, \rho(x) \left( \frac{1}{2} u(x)^2 + E \left( \frac{1}{\rho(x)} \right) + \Phi(x) \right) \]  

as a functional of the Eulerian variables, we find that

\[ \frac{\delta H}{\delta \rho(x)} = \frac{1}{2} u^2 + E + \Phi(x) - \frac{1}{\rho} E' \]  

and

VII-53
We also have
\[
\frac{\delta \rho(x)}{\delta \rho(x')} = \delta(x - x') \quad \text{and} \quad \frac{\delta \rho(x)}{\delta u(x')} = 0.
\]  
(11.22)

Thus,
\[
\frac{\partial}{\partial t} \rho(x) = \{\rho, H\} = \int dx' \left( \frac{\partial}{\partial x'} \left( \delta(x - x') \right) \rho u(x') \right) = -\frac{\partial}{\partial x} (\rho u),
\]  
(11.23)

as expected. By similar steps we obtain the expected equation for \( \partial u / \partial t \).

For a general, three-dimensional non-homentropic perfect fluid, (11.18) generalizes to the bracket obtained by Morrison and Greene,\(^{21}\)
\[
\{A, B\} = \iiint dx \left[ \nabla \left( \frac{\partial A}{\partial \rho} \right) \cdot \frac{\partial B}{\partial v} - \nabla \left( \frac{\partial B}{\partial \rho} \right) \cdot \frac{\partial A}{\partial v} + \frac{\nabla \times v}{\rho} \cdot \frac{\partial A}{\partial v} \times \frac{\partial B}{\partial v} + \frac{\nabla \eta}{\rho} \cdot \left( \frac{\partial A}{\partial \eta} \frac{\partial B}{\partial \eta} - \frac{\partial B}{\partial \eta} \frac{\partial A}{\partial \eta} \right) \right]
\]  
(11.24)

Next, as an example of the method of guessing, we determine the Poisson bracket corresponding to the quasigeostrophic equation,
\[
\frac{\partial}{\partial t} \nabla^2 \psi + J(\psi, \nabla^2 \psi + h(x,y)) = 0,
\]  
(11.25)

where \( \psi \) is the streamfunction for the flow, \( J(\cdot, \cdot) \) is the horizontal Jacobian operator, and the prescribed function \( h(x,y) \) is the bottom topography (or Coriolis parameter); see Chapters 2 and 6. To avoid subtleties connected with boundary conditions, we continue to assume that the flow is spatially periodic or unbounded (with \( \psi, h \to 0 \) at infinity). Let the (reduced) Eulerian phase space be the space of functions \( \psi(x,y) \) that are periodic or vanish at infinity, or, alternatively, the space of all \( \zeta(x,y) = \nabla^2 \psi \) (which, along with the periodic or quiescent boundary conditions, determines \( \psi \)). Let \( A[\zeta(x,y)] \) be an arbitrary functional. Then
\[
\frac{dA}{dt} = \iint dx \, dy \, \frac{\delta A}{\delta \zeta} \frac{\partial \zeta}{\partial t} = -\iint dx \, dy \, \frac{\delta A}{\delta \zeta} J(\psi, \zeta + h)
\]  
\[
= \iint dx \, dy \, q J \left( \psi, \frac{\delta A}{\delta \zeta} \right)
\]  
(11.26)
where

\[ q = \zeta + h \]  \hspace{1cm} (11.27)

is the potential vorticity. The conserved energy is

\[ H = \iint dx\, dy \ \frac{1}{2} \nabla \psi \cdot \nabla \psi. \]  \hspace{1cm} (11.28)

Hence

\[ \delta H = \iint dx\, dy \ \nabla \psi \cdot \nabla \delta \psi = -\iint dx\, dy \ \psi \ \delta \zeta, \]  \hspace{1cm} (11.29)

and thus

\[ \frac{\delta H}{\delta \zeta} = -\psi. \]  \hspace{1cm} (11.30)

Therefore, we can write (11.26) in the form,

\[ \frac{dA}{dt} = \iint dx\, dy \ q \left( \frac{\delta A}{\delta \zeta}, \frac{\delta H}{\delta \zeta} \right). \]  \hspace{1cm} (11.31)

On the other hand, we know that

\[ \frac{dA}{dt} = \{A,H\}. \]  \hspace{1cm} (11.32)

Comparing (11.31) and (11.32), we guess that

\[ \{A,B\} = \iint dx\, dy \ q \left( \frac{\delta A}{\delta \zeta}, \frac{\delta B}{\delta \zeta} \right) \]  \hspace{1cm} (11.33)

for any two functionals A and B. This is a guess, because we have really only established (11.33) for the special case \( B=H \). To show that our guess is correct, we must verify that the bracket (11.33) satisfies the antisymmetry and Jacobi properties. The antisymmetry property follows from the antisymmetry property of the Jacobian in the integrand of (11.33). The Jacobi property is somewhat harder to see, but it too follows from the corresponding property of \( J(,),\) (and our assumptions about the boundary conditions).

The canonical bracket (11.1) is nonsingular, but both (11.18) and (11.33) are singular brackets. In the case of (11.33), it is especially easy to determine the Casimirs, that is, the functionals C for which \( \{C,B\}=0 \) for any B. By (11.33) and our boundary assumptions,
\[
\{C,B\} = -\iint dx\,dy \, J\left(\frac{\delta C}{\delta \zeta}, q \frac{\delta B}{\delta \zeta}\right),
\]
(11.34)

and (11.34) must vanish for any \(B\). It follows that the coefficient of \(\delta B/\delta \zeta\) in the integrand of (11.34) must vanish. Hence

\[
\frac{\delta C}{\delta \zeta} = F'(q),
\]
(11.35)

where \(F'\) is an arbitrary function (not necessarily single-valued) and the prime denotes the derivative. It follows that

\[
C = \iint dx\,dy \, F(q),
\]
(11.36)

where \(F(q)\) is an arbitrary function of the potential vorticity. The Casimirs of the general perfect-fluid bracket (11.24) evidently take the form

\[
C = \iiint dx\,\rho F(\eta, q)
\]
(11.37)

where \(F\) is an arbitrary function of two variables, and

\[
q = \frac{(\nabla \times \mathbf{v}) \cdot \nabla \eta}{\rho}
\]
(11.38)

is the potential vorticity.\(^{22}\)

If the fluid is bounded by a rigid wall at which \(\psi\) vanishes, then the Poisson bracket (11.33) contains an additional, boundary-integral term whose form is very hard to guess.\(^{23}\) However, the Casimirs still take the general form (11.36). This situation seems to be typical: boundary conditions greatly complicate Poisson brackets. Why, then, are brackets so popular? The answer seems to be that many people dislike the Lagrangian description of fluid mechanics, and want to base their deductions on a purely Eulerian description of the flow. Brackets offer a way to do this. The Eulerian bracket is a machine that operates on any two Eulerian functionals. Along with the Hamiltonian of the system (which also depends only on Eulerian variables), the bracket provides a complete and abstractly geometric description of the whole dynamics. The geometric nature of this description is important because many important dynamical conclusions depend not on the detailed form of the bracket, but on its general, geometrical properties. The next section illustrates this very well.

12. Pseudoenergy, stability and available energy

Again we consider a discrete Hamiltonian system described by the coordinates \(\{\mathbf{z}^i(t), i = 1\ to\ n\}\), and governed by
\[ \frac{dz^j}{dt} = J^{ij} \frac{\partial H}{\partial z^j}. \]  

(12.1)

where \( H(z) \) is the Hamiltonian, and the symplectic tensor \( J^{ij}(z) \) is antisymmetric and obeys the Jacobi property (9.20). Now, however, we regard the \( z^j \) as the analogues of the Eulerian fluid variables \( \rho(x,t), v(x,t), \) etc. The superscript on \( z^j(t) \) is analogous to the argument \( x \) of \( \rho(x,t) \). The tensor \( J^{ij} \) is singular, with null covectors equal to the gradients of \( K \) Casimirs \( C^{(k)}(z) \),

\[ J^{ij} \frac{\partial C^{(k)}}{\partial z^j} = 0. \]  

(12.2)

Let \( z_0 = (z_0^1, z_0^2, \ldots, z_0^n) \) be a fixed point of (12.1). That is, let

\[ \frac{dz^j}{dt} = J^{ij} \frac{\partial H}{\partial z^j} = 0 \quad \text{at} \quad z = z_0. \]  

(12.3)

The fixed point \( z_0 \) corresponds to a steady solution of the Eulerian fluid equations. If \( J \) were nonsingular, then (12.3) would imply that

\[ \frac{\partial H}{\partial z^j} = 0 \quad \text{at} \quad z = z_0. \]  

(12.4)

That is, the fixed point would be a stationary point of the energy \( H \). However, in the present case of a singular \( J \), (12.3) only implies that \( \partial H/\partial z \) lies in the null space of \( J \). Hence,

\[ \frac{\partial H}{\partial z^j} = -\sum_{k=1}^K \lambda_k \frac{\partial C^{(k)}}{\partial z^j} \quad \text{at} \quad z = z_0, \]  

(12.5)

where \( \{\lambda_k\} \) is a set of \( K \) constants. The constants \( \{\lambda_k\} \) depend on the particular \( z_0 \). According to (12.5), every fixed point of (12.1) is a stationary point of

\[ I(z) \equiv H(z) + \sum_k \lambda_k C^{(k)}(z), \]  

(12.6)

for some \( \{\lambda_k\} \). That is,

\[ \frac{\partial I}{\partial z^j} = 0 \quad \text{at} \quad z = z_0. \]  

(12.7)
If the stationary point \( z_0 \) is an extremum of \( I(z) \), that is, a local maximum or minimum of (12.6), then the fixed point \( z_0 \) is stable in the sense that if \( z \) is sufficiently near \( z_0 \), then it always stays within a neighborhood of \( z_0 \).

Now let

\[
z = z_0 + \Delta z,
\]

where

\[
\Delta z = (\Delta z^1, \Delta z^2, \ldots, \Delta z^n)
\]

is the disturbance amplitude, the departure of \( z \) from its value \( z_0 \) in the steady state. Consider the pseudoenergy, a measure of the disturbance activity. Since \( dI(z)/dt=0 \), then \( d\Delta I/dt=0 \); the pseudoenergy is conserved. For small \( \Delta z \), the Taylor-series expansion of (12.10) yields

\[
\Delta I(\Delta z; z_0) \equiv I(z_0 + \Delta z) - I(z_0),
\]

a measure of the disturbance activity. Since \( dI(z)/dt=0 \), then \( d\Delta I/dt=0 \); the pseudoenergy is conserved. For small \( \Delta z \), the Taylor-series expansion of (12.10) yields

\[
\Delta I(\Delta z; z_0) \equiv H_L + O((\Delta z)^3),
\]

where

\[
H_L = \frac{1}{2} \frac{\partial^2 I}{\partial z^i \partial z^j}(z_0) \Delta z^i \Delta z^j.
\]

Thus, on account of (12.7), the conserved disturbance activity (12.10) is second order in the disturbance-amplitude \( \Delta z \).

In examining the dynamics near the fixed point \( z=z_0 \), one frequently considers the linear approximation to (12.1), that is, the approximation obtained by expanding (12.1) in powers of \( \Delta z \) and then simply throwing away all the \( O((\Delta z)^2) \) terms. We shall show that these linear equations are Hamiltonian and that the Hamiltonian is (12.12). By (12.1) and (12.3),

\[
\frac{d}{dt}(z_0^i + \Delta z^i) = J^{ij}(z_0 + \Delta z) \frac{\partial H}{\partial z^j}(z_0 + \Delta z)
\]

\[
= \left[ J^{ij}(z_0) + \frac{\partial J^{ij}}{\partial \Delta z^m}(z_0) \Delta z^m \right] \left[ \frac{\partial H}{\partial z^j}(z_0) + \frac{\partial^2 H}{\partial z^i \partial z^m}(z_0) \Delta z^m \right] + O((\Delta z)^2)
\]

\[
= \left[ \frac{\partial J^{ij}}{\partial \Delta z^m} \frac{\partial H}{\partial z^m} + J^{ij} \frac{\partial^2 H}{\partial z^i \partial z^m}(z_0) \right] \Delta z^m + O((\Delta z)^2)
\]
The zero subscript on the square brackets means that enclosed expression is evaluated at $z = z_0$. But

$$
\frac{\partial J^{ij}}{\partial z^m} \frac{\partial H}{\partial z^j} = - \sum_k \lambda_k \frac{\partial C^{(k)}}{\partial z^j} = J^{ij} \sum_k \lambda_k \frac{\partial^2 C^{(k)}}{\partial z^m \partial z^j}
$$

by (12.5) and (12.2). Thus the linear dynamics is

$$
\frac{d}{dt} \Delta z^i = J^{ij}(z_0) \frac{\partial H_L}{\partial \Delta z^j},
$$

where $H_L$ is given by (12.12).

The linear dynamics (12.15) exactly conserve $H_L$. Moreover, $H_L$ is the leading term in the pseudoenergy (12.10), which is conserved by the exact nonlinear dynamics (12.1). Interestingly, however, the Hamiltonian $H$ of the exact dynamics differs from the Hamiltonian $H_L$ of the linear dynamics by terms that are order one in $\Delta z$. Hence the exact energy is not a close approximation to the energy (12.12) conserved by the linear dynamics; the two energies differ by terms involving the Casimirs. This is the general resolution of a long-standing paradox regarding linear approximations to the nonlinear, Eulerian fluid equations: the energy conserved by the linear equations bears an ambiguous relationship to the energy conserved by the exact, nonlinear equations. Well-known examples include the linear equations for sound waves and for internal gravity waves. To resolve the paradox, one always needs to invoke additional conservation laws, which turn out to be the conservation laws for the Casimirs of the system. We come back to this below. First, however, we return to the question of stability.

If the linear Hamiltonian (12.12) is a sign-definite quadratic form, then the fixed point $z_0$ is stable according to the linear dynamics (12.15). Moreover, since $H_L$ is the leading-order term in (12.11), $z_0$ is also stable according to the general (nonlinear) dynamics (12.1) provided that $\Delta z$ is sufficiently small. Interestingly, however, the latter conclusion does not survive the passage to the continuum limit, in which $\{\rho(x), v(x), \eta(x)\}$ replace $\{z_i\}$. In the continuum limit, the number $n$ of phase-space dimensions becomes uncountably infinite (corresponding to the uncountably infinite number of locations $x$ in the fluid), and although the $O((\Delta z)^3)$ are smaller than the $H_L$-terms in (12.11), there are (roughly speaking) infinitely more of them. Nevertheless, the form of (12.10) sometimes suggests how to construct a useful bound on $\Delta z$.

Now we apply these ideas to fluids, using the two examples given in the previous section. First consider the quasi-geostrophic system (11.25) with Hamiltonian (11.28), bracket (11.33), and general Casimir (11.36). Let $\psi_0(x)$ be a particular steady solution of the dynamical equation (11.25). Then, in analogy with (12.7), $\psi_0(x)$ must be a stationary point of the functional,

$$
I[\psi(x)] = H + C = \iint d\mathbf{x} \left\{ \frac{1}{2} \nabla \psi \cdot \nabla \psi + F(q) \right\},
$$

(12.16)
where \( q = \nabla^2 \psi + h \) is the potential vorticity. Indeed,

\[
\delta I = \iint d\mathbf{x} \left\{ (-\psi + F'(q)) \delta \zeta \right\} = 0
\]

implies that

\[
\psi_0 = F'(\nabla^2 \psi_0 + h) \equiv \Psi(q_0).
\]

The function \( \Psi \), which determines the steady state, is analogous to the set \( \{ \lambda_k \} \) of constants in (12.6). Now let

\[
\psi(x,t) = \psi_0(x) + \psi'(x,t),
\]

where \( \psi' \) is analogous to \( \Delta z \). The analogue of (12.10) is

\[
\Delta I = I[\psi_0 + \psi'] - I[\psi_0] = \iint d\mathbf{x} \left\{ \frac{1}{2} \nabla(\psi_0 + \psi') \cdot \nabla(\psi_0 + \psi') - \frac{1}{2} \nabla \psi_0 \cdot \nabla \psi_0 + F(q_0 + q') - F(q_0) \right\} = H_L + O((\psi')^3)
\]

where

\[
H_L = \iint d\mathbf{x} \left\{ \frac{1}{2} \nabla \psi' \cdot \nabla \psi' + \frac{1}{2} F''(q_0) (\nabla^2 \psi')^2 \right\}.
\]

Because of (12.17), (12.20) is second-order in the disturbance streamfunction \( \psi' \). It is easy to check that the linear dynamics,

\[
\frac{\partial \nabla^2 \psi'}{\partial t} + J(\psi_0, \nabla^2 \psi') + J(\psi', \nabla^2 \psi_0 + h) = 0,
\]

conserves (12.21). According to linear dynamics, the steady state \( \psi_0(x) \) is stable if (12.21) is positive-definite, that is, if

\[
F''(q_0) = \frac{d^2 \Psi}{dq_0^2} > 0.
\]

This is Arnol’d’s theorem.25

None of these steps depends on the precise form (11.33) of the bracket, and, in fact, (12.16-23) continue to hold if the periodic boundary is replaced by a simply-connected rigid boundary at which \( \psi = 0 \). For the rigid boundary, the bracket (11.33) is incorrect, but the Casimirs still take the form (11.36). However, in the case of a simply connected
rigid boundary, (12.21) is negative-definite if $F''(q_0)$ is sufficiently negative. Specifically, if
\[
\frac{d\Psi}{dq_0} < -\frac{1}{k_{\text{min}}}^2,
\]
(12.24)
where $k_{\text{min}}$ is the lowest wavenumber in the usual eigenfunction expansion, then $H_\ell$ is negative-definite and $\psi_0(x)$ is stable according to the linear dynamics (12.22). This is Arnol’d’s second theorem.

Once again, stability according to linear dynamics does not imply stability according to the exact nonlinear dynamics, but the form of $\Delta I$ suggests a strategy for constructing a norm in which the disturbance remains bounded. By (12.20),
\[
\Delta I = \int \int dx \left\{ \frac{1}{2} \nabla \psi' \cdot \nabla \psi' + \nabla \psi_0 \cdot \nabla \psi' + F(q_0 + q') - F(q_0) \right\}
\]
(12.25)
This is the same as
\[
\Delta I = \int \int dx \left\{ \frac{1}{2} \nabla \psi' \cdot \nabla \psi' - \Psi(q_0) \nabla^2 \psi' + \int_{q_0}^{q_0 + q'} \Psi(q) dq \right\}
\]
(12.26)
Thus $\Delta I$ is positive if $d\Psi(q)/dq>0$. If, moreover,
\[
0 < c_1 < \frac{d\Psi}{dq} < c_2,
\]
(12.27)
where $c_1$ and $c_2$ are constants, then the steady state is stable in the sense that
\[
\|\psi'(t)\|^2 < \frac{c_2}{c_1} \|\psi'(0)\|^2,
\]
(12.28)
where
\[
\|\psi'\|^2 \equiv \int \int dx \left\{ \nabla \psi' \cdot \nabla \psi' + c_1(q')^2 \right\}.
\]
(12.29)
The business of obtaining such stability bounds can become very intricate. However, even in cases where it is difficult to construct bounds, the concept of pseudoenergy
illuminates the conservation laws that arise from linear approximations. This brings us to our second example.

Consider the perfect-fluid equations, with Hamiltonian

\[
H = \iiint d\mathbf{x} \left\{ \frac{1}{2} \mathbf{v} \cdot \mathbf{v} + \rho E \left( \frac{1}{\rho} , \eta \right) + \rho \Phi(\mathbf{x}) \right\},
\]

(12.30)
bracket (11.24), and Casimirs

\[
C = \iiint d\mathbf{x} \rho F(\eta,q),
\]

(12.31)

where \(q\) is the potential vorticity (11.38). We assume for simplicity that there is no external potential, \(\Phi \equiv 0\). Then

\[
\frac{\delta H}{\delta \rho} = \frac{1}{2} \mathbf{v} \cdot \mathbf{v} + E + \frac{P}{\rho}, \quad p \equiv -\frac{\partial E(\alpha, \eta)}{\partial \alpha},
\]

\[
\frac{\delta H}{\delta \mathbf{v}} = \rho \mathbf{v},
\]

(12.32)

\[
\frac{\delta H}{\delta \eta} = \rho T, \quad T \equiv \frac{\partial E(\alpha, \eta)}{\partial \eta}
\]

and

\[
\frac{\delta C}{\delta \rho} = F - q F_q,
\]

\[
\frac{\delta C}{\delta \mathbf{v}} = \nabla \times (F_q \nabla \eta),
\]

(12.33)

\[
\frac{\delta C}{\delta \eta} = \rho F_\eta - (\nabla \times \mathbf{v}) \cdot \nabla F_q
\]

For a given steady state \(\{\rho_0(\mathbf{x}), \mathbf{v}_0(\mathbf{x}), \eta_0(\mathbf{x})\}\), the first step is to determine the corresponding \(F(\eta, q)\). In analogy with (12.7), we require that

\[
\frac{\delta C}{\delta \rho} = -\frac{\delta H}{\delta \rho}, \quad \frac{\delta C}{\delta \mathbf{v}} = -\frac{\delta H}{\delta \mathbf{v}}, \quad \frac{\delta C}{\delta \eta} = -\frac{\delta H}{\delta \eta},
\]

(12.34)

at the steady state. Let the steady state be a state of rest, \(\mathbf{v}_0 \equiv 0\). Then (12.34b) is satisfied if \(F\) depends only on \(\eta\), and (12.34a,c) become

\[
F(\eta_0) = -E \left( \frac{1}{\rho_0}, \eta_0 \right) + \frac{1}{\rho_0} E_\alpha \left( \frac{1}{\rho_0}, \eta_0 \right)
\]

(12.35)
and

\[ F_\eta(\eta_0) = -E_\eta \left( \frac{1}{\rho_0}, \eta_0 \right) \]  \hspace{1cm} (12.36)

The function

\[ F(\eta) = -E_\rho \left( \frac{1}{\rho_0}, \eta \right) + \frac{1}{\rho_0} E_\alpha \left( \frac{1}{\rho_0}, \eta_0 \right) \equiv -E_\rho \left( \frac{1}{\rho_0}, \eta \right) - \frac{\rho_0}{\rho_0} \]  \hspace{1cm} (12.37)

satisfies both (12.35) and (12.36). By (12.37) and (12.31), the pseudoenergy is

\[ I[\rho(x), v(x), \eta(x)] = H + C = \iiint dx \left\{ \frac{1}{2} \rho v \cdot v + \rho E_\rho \left( \frac{1}{\rho}, \eta \right) - \rho E_\rho \left( \frac{1}{\rho_0}, \eta_0 \right) - \frac{\rho}{\rho_0} \rho_0 \right\}, \]  \hspace{1cm} (12.38)

and the disturbance activity is

\[ \Delta I = I[\rho(x), v(x), \eta(x)] - I[\rho_0(x), v_0(x), \eta_0(x)] \]
\[ = \iiint dx \left\{ \frac{1}{2} \rho v \cdot v + \rho E_\rho \left( \frac{1}{\rho}, \eta \right) - \rho E_\rho \left( \frac{1}{\rho_0}, \eta_0 \right) - \frac{(\rho - \rho_0)}{\rho_0} \rho_0 \right\} \]  \hspace{1cm} (12.39)

Now suppose that the fluid is homentropic, \( \eta = \eta_0 = \text{const} \), and let

\[ \rho'(x,t) = \rho(x,t) - \rho_0(x), \quad v'(x,t) \equiv v(x,t) \]  \hspace{1cm} (12.40)

be a small departure from the state of rest. Then, expanding (12.39) about the state of rest, and disregarding terms of cubic order in the primes, we obtain

\[ \Delta I = \iiint dx \left\{ \frac{1}{2} \rho_0 v' \cdot v' + \frac{1}{2} \frac{1}{\rho_0} E_{\text{res}} \left( \frac{1}{\rho_0}, \eta_0 \right) (\rho')^2 \right\} = \iiint dx \left\{ \frac{1}{2} \rho_0 v' \cdot v' + \frac{1}{2} \frac{c^2}{\rho_0} (\rho')^2 \right\} \]  \hspace{1cm} (12.41)

where \( c \) is the sound speed in the state of rest. This is the familiar form of the energy conserved by the linear equations for sound waves. The \( \rho' \)-term in (12.41) can never be negative, and any decrease in this term must be balanced by a gain in the kinetic energy. Therefore, it is sensible to call this term the available internal energy. Similarly, if we had retained the gravitational potential \( \Phi \), we would have obtained an available potential energy term in (12.41). The linear dynamics conserve (12.41), while the general, nonlinear dynamics conserve (12.39). Again, both are second-order in the disturbance fields, but the latter is nonquadratic. The pseudoenergy (12.39) offers a means of generalizing the concept of available energy to the general, nonlinear dynamics.
However, in the general nonlinear case, the available energy is even more dependent on the choice of the reference state $z_0$.

If the reference state has additional symmetries (besides, or perhaps instead of, time-independence), then these methods generalize considerably. We offer only a glimpse. To explain the essential idea, we recur to the discrete Hamiltonian system (12.1). Let the dynamical variables be $z^i(t;\varepsilon)$, where $\varepsilon$ is a parameter that controls their transformation (or variation). Let $\varepsilon=0$ correspond to the unvaried dynamics, that is, to the identity transformation. Consider the transformations defined by solutions of the ordinary differential equations

$$
\frac{dz'^i}{d\varepsilon} = \{z'^i, G\}.
$$

(12.42)

The function $G$ is called the generator of the transformation. Thus Casimirs generate null transformations. From (12.1) we see that the Hamiltonian generates a transformation that corresponds to the actual physical evolution of the system; the transformation parameter is ordinary time.

Now suppose that the transformation generated by $G$ causes no variation in the Hamiltonian $H$. Then since

$$
0 = \frac{dH}{d\varepsilon} = \frac{\partial H}{\partial z'^i} \frac{dz'^i}{d\varepsilon} = \frac{\partial H}{\partial z'^i} \{z'^i, G\} = \{H, G\},
$$

(12.43)

it follows that $G$ is conserved,$^{27}$

$$
\frac{dG}{dt} = 0.
$$

(12.44)

If, for example, $H$ is translation-invariant, then $G$ is the corresponding momentum.

Now we return to the consideration of a particular fixed point $z_0$. Suppose that the steady solution represented by this fixed point is unaffected by the transformation generated by $G$ in the sense that

$$
\{z, G\} = 0 \quad \text{at} \quad z = z_0.
$$

(12.45)

If $G$ generates a translation, then (12.45) states that the steady flow represented by $z_0$ is translation-invariant. In the same way that (12.3) implies (12.5), (12.45) implies that $z_0$ is a stationary point of

$$
I(z) \equiv G(z) + \sum_k \lambda_k C^{(k)}(z).
$$

(12.46)

More generally, $z_0$ is a stationary point of
\[ I(z) \equiv H(z) + \lambda_0 G(z) + \sum_k \lambda_k C^{(k)}(z). \] (12.47)

Since \( H \), \( G \) and \( C^{(k)} \) are all conserved, (12.47) is also conserved. The disturbance activity \( \Delta I \) based on (12.47) is clearly a generalization of that considered before.

The most general \( I(z) \) contains the generator of every symmetry of \( z_0 \), plus all the Casimirs. This general \( I(z) \) offers the greatest number of possibilities in the search for stability bounds and useful expressions for available energy. However, to achieve maximal flexibility, one must be willing to consider reference states with lots of symmetry. A thorough discussion of this subject would fill a book by itself, and is well beyond our scope.²⁸

13. Dynamical approximations: the semigeostrophic equations

Suppose we have a fluid system whose governing dynamical equations are (in their conservative form) Hamiltonian. These equations might be the general equations for a perfect fluid, but they could also be any known (but Hamiltonian) approximation thereto. Suppose further that this Hamiltonian dynamics is canonical in the sense that a variational principle exists. The latter assumption is not actually restrictive, because every set of fluid equations ought to have an underlying Lagrangian formulation, in which the dependent variables are the locations of marked fluid particles. Only a Lagrangian formulation offers a truly complete description of the fluid motion.

Now suppose that we want to make a further approximation to this Hamiltonian system, and that we want this further approximation to be Hamiltonian too. A typical motivation for this further approximation might be that the original equations (which we henceforth call exact) include more physical phenomena than we need to consider, and that the extra phenomena, though physically unimportant, greatly complicate the analysis. For example, the exact perfect-fluid equations admit sound waves, but sound waves are unimportant contributors to the flow depicted on weather maps. Hence, numerical weather-prediction models use approximate dynamical equations that filter out sound waves, thereby avoiding the very short time-step needed to resolve them.

In most cases, inertia-gravity waves do not contribute significantly to weather, but the search for approximate equations that filter out inertia-gravity waves has had a long and somewhat tortuous history. Certainly, the quasigeostrophic equations (Chapter 2) are one possibility; they are relatively simple and admit no inertia-gravity waves. But quasigeostrophy comes at a high price: The quasigeostrophic equations require that the flow remain close to a prescribed state of rest in the sense that the isopycnals (and the top and bottom boundaries) must be nearly flat. Neither the ocean nor the atmosphere satisfies this artificial restriction very well.

Section 6 offered a general method for making dynamical approximations to Hamiltonian systems. Given an “exact” set of dynamical equations arising from a variational principle, we form the approximation by attaching constraints to the variational principle. The constraints reduce the number of degrees of freedom and, thereby, the range of physical phenomena described by the approximation. Section 6 emphasized one very important advantage of this approximation method: The resulting approximate dynamics inherit the conservation laws of the exact dynamics if the
constraints respect the corresponding symmetry property. In this section, we illustrate another important advantage of the method, namely, that it suggests transformations to new dependent and independent variables in which the approximate dynamics takes its simplest mathematical form. The motivation for these transformations rests on the geometrical viewpoint developed in Section 9.

In Section 6 we obtained the Euler equations (which support no sound waves) from the exact perfect-fluid equations by attaching the constraint (6.5) of uniform density on the particle locations \( x(a, \tau) \). In this section, we obtain the semigeostrophic equations (which support no inertia-gravity waves) by attaching constraints corresponding to geostrophic balance. These geostrophic constraints involve both the locations \( x(a, \tau) \) and the velocities \( u(a, \tau) \) of marked fluid particles. Hence, we must use the extended form of Hamilton’s principle, in which \( x(a, \tau) \) and \( u(a, \tau) \) are independent variables.

Meteorologists and oceanographers often regard the primitive equations (that is, the perfect-fluid equations with the Boussinesq and hydrostatic approximations — see Sections 16 and 17 of Chapter 2) as the departure point for further approximations. The primitive equations filter out sound waves but admit both inertia-gravity waves and nearly geostrophic motions. The primitive equations are closely related to the shallow-water equations (see Section 18 of Chapter 2). In Chapter 2, we frequently used the shallow-water equations as a simple paradigm for the primitive equations.

In this section, we derive the semigeostrophic approximation to the shallow-water equations. The semigeostrophic equations filter out inertia-gravity waves but accurately describe the nearly geostrophic motions. The semigeostrophic approximation to the primitive equations proceeds by similar steps, but it is easier to grasp the essential ideas by working with the simpler shallow-water equations.

In Section 6, we obtained the variational principle for the nonrotating shallow-water equations by constraining the fluid to have a uniform density and to move in vertical columns. The generalization to rotating coordinates is easy. The rotating shallow-water equations are equivalent to the requirement that

\[
\delta \int d\tau \, L = 0 , \tag{13.1}
\]

where

\[
L[u(a, \tau), x(a, \tau)] = \int \int da \left\{ (u - R) \frac{\partial x}{\partial \tau} + (v + P) \frac{\partial y}{\partial \tau} \right\} - H \quad \tag{13.2}
\]

is the Lagrangian, and

\[
H = \frac{1}{2} \int \int da \left\{ u^2 + v^2 + g \frac{\partial (a, b)}{\partial (x, y)} \right\} \quad \tag{13.3}
\]

is the Hamiltonian. In (13.1) \( \delta \) stands for arbitrary variations in the locations \( \delta x(a, b, \tau) = (\delta x, \delta y) \) and velocities \( \delta u(a, b, \tau) = (\delta u, \delta v) \) of the fluid column labeled by \( a = (a, b) \) at time \( \tau \). The prescribed functions \( R(x, y) \) and \( P(x, y) \) are any two functions satisfying
\[ \frac{\partial R}{\partial y} + \frac{\partial P}{\partial x} = f(x,y), \quad (13.4) \]

where \( f(x,y) \) is the Coriolis parameter. For generality, we allow \( f \) to depend arbitrarily on location. If \( f = 2\Omega \) (constant), then we can take \( R = \Omega y \) and \( P = \Omega x \). The particle labels are assigned so that the fluid depth is

\[ h = \frac{\partial (a,b)}{\partial (x,y)}. \quad (13.5) \]

(This differs slightly from the labeling convention in Section 6.) Once again, the \( \tau \)-derivative of (13.5) yields

\[ \frac{Dh}{Dt} + h \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0, \quad (13.6) \]

where \( D/Dt = \partial/\partial t + u \partial/\partial x + v \partial/\partial y \). Hamilton’s principle (13.1) yields

\[ \delta u : \quad u = \frac{\partial x}{\partial \tau}, \quad \delta v : \quad v = \frac{\partial y}{\partial \tau}, \quad (13.7) \]

and

\[ \delta x : \quad \frac{\partial u}{\partial \tau} - f \frac{\partial v}{\partial \tau} = -g \frac{\partial h}{\partial x}, \quad \delta y : \quad \frac{\partial v}{\partial \tau} + f \frac{\partial u}{\partial \tau} = -g \frac{\partial h}{\partial y}. \quad (13.8) \]

The Hamiltonian (13.3) has the time-translation and particle-relabeling symmetry properties corresponding to the conservation of energy (13.3) and potential vorticity,

\[ \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} + f \frac{h}{h} = \frac{\partial}{\partial x} \left( -R(x,y) \frac{\partial x}{\partial \tau} + P(x,y) \frac{\partial y}{\partial \tau} - \frac{1}{2} g \frac{\partial (a,b)}{\partial (x,y)} \right), \quad (13.9) \]

on fluid columns.

We are interested in nearly geostrophic flow. If we attach the constraints

\[ u = v = 0 \quad (13.10) \]

to (13.1), that is, if we simply drop the \( u \)- and \( v \)-terms from (13.2) and (13.3), then the Lagrangian

\[ L[x(a,\tau)] = \int \int da \left\{ -R(x,y) \frac{\partial x}{\partial \tau} + P(x,y) \frac{\partial y}{\partial \tau} - \frac{1}{2} g \frac{\partial (a,b)}{\partial (x,y)} \right\} \quad (13.11) \]
depends only on the fluid-column locations. The dimensionality of phase-space has been reduced by half. Hamilton’s principle now yields

\[
\delta x: \quad -f \frac{\partial y}{\partial \tau} = -g \frac{\partial h}{\partial x}, \quad \delta y: \quad +f \frac{\partial x}{\partial \tau} = -g \frac{\partial h}{\partial y}.
\]  

(13.12)

Thus the dynamics corresponding to (13.11) is (13.6) and (13.12). We lose \(u\) and the equations (13.7) relating \(u\) to \(\partial x/\partial \tau\), but \(\partial x/\partial \tau\) retains its meaning as the true velocity of massive fluid columns. The Hamiltonian in (13.11) has the time-translation and particle-relabeling symmetries of the exact Hamiltonian (13.3). By the methods of Section 2, the energy \(\frac{1}{2} \int \int da \, gh\), and potential vorticity \(f/h\) are correspondingly conserved.

Equations (13.6) and (13.12), which filter out inertia-gravity waves, are the shallow-water analogues of the planetary geostrophic equations (Chapter 3, Section 1), that is, the primitive equations without inertia. However, the complete neglect of inertia is often too inaccurate. Thus we need a better approximation to (13.7-8) than (13.12). If, instead of (13.10), we impose the constraints

\[
u = v_G \equiv g \frac{\partial h}{\partial x}, \quad v = v_G \equiv f \frac{\partial h}{\partial y},
\]  

(13.13)

then the Lagrangian becomes

\[
L[x(a,\tau)] = \int \int da \left\{ (u_G - R) \frac{\partial x}{\partial \tau} + (v_G + P) \frac{\partial y}{\partial \tau} \right\} - H,
\]  

(13.14)

with Hamiltonian,

\[
H[x(a,\tau)] = \frac{1}{2} \int \int da \left\{ u_G^2 + v_G^2 + g \frac{\partial (a,b)}{\partial (x,y)} \right\}.
\]  

(13.15)

Since the geostrophic velocity \(u_G[x(a,\tau)]\) is a functional of the particle locations, (13.14) and (13.15) are also functionals of the particle locations. The motivation for (13.13) is (13.12).

The \(\delta x(a,\tau)\) and \(\delta y(a,\tau)\) variations of (13.14) yield momentum equations that filter out inertia-gravity waves, but are intermediate between (13.8) and (13.12) in physical content.\(^{30}\) (The precise form of these equations does not concern us here. In this section, we are mainly interested in the philosophy behind the Hamiltonian approximation theory.) By the time-translation and particle-relabeling symmetry properties of (13.15), the dynamics based on (13.14) conserves the energy (13.15) and the potential vorticity

\[
\frac{\partial y_G}{\partial x} - \frac{\partial u_G}{\partial y} + f
\]  

(13.16)
on fluid columns. Note that the potential vorticity (13.16) differs from the exact potential vorticity (13.9) in that the geostrophic velocities replace \( u \) and \( v \).

Unfortunately, the momentum equations arising from (13.14) are more mathematically complex than either (13.8) or (13.12). This mathematical complexity arises from the relatively complex dependence of the geostrophic velocity \( u_G \) on the particle locations \( \mathbf{x}(a, \tau) \).

To simplify (13.14), we first recognize that it consists of two pieces: a piece

\[
\iint d\mathbf{a} \left\{ (u_G - R) \frac{\partial x}{\partial \tau} + (v_G + P) \frac{\partial y}{\partial \tau} \right\},
\]

(13.17)

corresponding to the Poisson bracket, and the Hamiltonian (13.15). From Section 9, we know that the Poisson bracket takes its simplest form in canonical coordinates. However, (13.17) is not in canonical coordinates, because it does not fit the form

\[
\sum_i p_i \frac{dq_i}{dt}.
\]

(13.18)

On the other hand, the Poisson-bracket part of the exact Lagrangian (13.2) does take the form

\[
\iint d\mathbf{a} \left\{ p_x \frac{\partial x}{\partial \tau} + p_y \frac{\partial y}{\partial \tau} \right\}
\]

(13.19)

of canonical coordinates with generalized momenta

\[
p_x = u - R(x, y), \quad p_y = v + P(x, y).
\]

(13.20)

(The Hamiltonian (13.3) is easily written as \( H[p_x, p_y, x, y] \).)

If \( f = 2\Omega \) (constant), then (13.11) is also in canonical form, because

\[
\iint d\mathbf{a} \left\{ -\Omega \frac{\partial y}{\partial \tau} + \Omega \frac{\partial x}{\partial \tau} \right\} = \iint d\mathbf{a} \left\{ 2\Omega x \frac{\partial y}{\partial \tau} \right\},
\]

(13.21)

to within an irrelevant total time-derivative — irrelevant because the variations allowed by Hamilton’s principle vanish at the endpoints in time. Thus, in the case of a constant Coriolis parameter, \( p = 2\Omega x \) and \( q = y \) are canonical coordinates. If \( f \) is nonconstant but takes a simple analytical form, it is easy to find canonical coordinates. For example, if \( f = f_0 + \beta y \), we may take \( R = f_0 + \gamma \beta y^2 \) and \( P = 0 \); then \( p = -R \) and \( q = x \) are canonical coordinates. However, the general form (13.11) is already nearly as simple as the canonical form.

The same cannot be said of the Lagrangian (13.14). The relatively complicated functional dependence of the geostrophic velocities \( u_G(x(a, \tau)) \) on the particle locations...
\( \mathbf{x}(a, \tau) \) gives the Poisson-bracket piece (13.17) of (13.14) a complicated, noncanonical form. Canonical coordinates exist, but (as is generally the case), they are very difficult to compute.

Fortunately, we have two powerful weapons to deploy. First, (13.17) differs from the (nearly) canonical form in (13.11) by terms that are relatively small; the \( u_G \)-terms in (13.17) are an order-Rossby number smaller than the \( R \)- and \( P \)-terms. This suggests a perturbative approach. Second, the approximate Lagrangian (13.14) already differs from the exact Lagrangian (13.2) at the second order in Rossby number. Hence, additional modifications to (13.14) are allowed, so long as they do not introduce errors of larger size.

With these points in mind, we let

\[
x_s = x + F(x, y), \quad y_s = y + G(x, y),
\]

(13.22)

be new coordinates, where \( F \) and \( G \) are functionals to be determined. We shall choose \( F \) and \( G \) such that (13.17) takes the form of (13.11) in the new coordinates, namely

\[
\iint \mathbf{a} \left\{ -R(x_s, y_s) \frac{\partial x_s}{\partial \tau} + P(x_s, y_s) \frac{\partial y_s}{\partial \tau} \right\},
\]

(13.23)

to within a consistent relative error of \( O(\text{Ro}^2) \), where \( \text{Ro} = \frac{U}{f_0 L} \) is the Rossby number.\(^{31}\) The functions \( R(x_s, y_s) \) and \( P(x_s, y_s) \) are the same functions that appear in the original shallow-water Lagrangian, but with the new coordinates in their argument-slots. That is, if \( R(x, y) = f_0 y + \mathcal{L} \beta y^2 \), then \( R(x_s, y_s) = f_0 y_s + \mathcal{L} \beta y_s^2 \).

To within the allowed error, (13.23) is

\[
\iint \mathbf{a} \left\{ -R \frac{\partial F}{\partial x} - \frac{\partial R}{\partial y} G \right\} \frac{\partial x}{\partial \tau} + \left( P + \frac{\partial P}{\partial x} F + \frac{\partial P}{\partial y} G \right) \frac{\partial y}{\partial \tau} \right\},
\]

(13.24)

In (13.24), all the terms are evaluated at \((x, y)\). Now we go to work on the \( \partial F/\partial \tau \) - and \( \partial G/\partial \tau \) -terms in (13.24). They are

\[
\iint \mathbf{a} \left\{ -R \frac{\partial F}{\partial \tau} + P \frac{\partial G}{\partial \tau} \right\} = \iint \mathbf{a} \left\{ -\frac{\partial}{\partial \tau} (RF) + F \frac{\partial R}{\partial \tau} + \frac{\partial}{\partial \tau} (PG) - G \frac{\partial P}{\partial \tau} \right\}.
\]

(13.25)

The total time-derivatives in (13.25) are irrelevant, because the variations allowed by Hamilton’s principle vanish at the endpoints in time. Applying the chain rule to the \( \partial R/\partial \tau \) - and \( \partial P/\partial \tau \) -terms in (13.25), and substituting back into (13.24), we find that
\[
(13.23) = \iint \mathrm{d}a \left\{ -fG - R \frac{\partial x}{\partial \tau} + (fF + P) \frac{\partial y}{\partial \tau} \right\}
\] 

(13.26)

after use of (13.4). Clearly (13.26) agrees with (13.17) to within a consistent error if we choose

\[
F = \frac{v_s}{f(x_s, y_s)}, \quad G = -\frac{u_s}{f(x_s, y_s)},
\]

(13.27)

where

\[
u_s = -\frac{g}{f(x_s, y_s)} \frac{\partial h}{\partial y}, \quad v_s = \frac{g}{f(x_s, y_s)} \frac{\partial h}{\partial x}.
\]

(13.28)

Once again, \(f(x_s, y_s)\) has the same dependence on its arguments as does \(f(x, y)\). That is, if \(f(x, y) = f_0 + \beta y\), then \(f(x_s, y_s) = f_0 + \beta y_s\). The denominators of (13.27) and (13.28) could just as well be \(f(x, y)\); the overall accuracy would be the same. However, our strategy is to write the dynamics as completely as possible in terms of the new variables \(x_s = (x_s, y_s)\).

With (13.27-28) and (13.22), the approximate Lagrangian becomes

\[
L_s[x, (a, \tau)] = \iint \mathrm{d}a \left\{ \frac{1}{2} \left( f_0 + \beta y \right)^2 + \frac{g}{f(x_s, y_s)} \frac{\partial h}{\partial x} \right\} - H_s,
\]

(13.29)

where

\[
H_s[x, (a, \tau)] = \frac{1}{2} \iint \mathrm{d}a \left\{ u_s^2 + v_s^2 + g \frac{\partial (a, b)}{\partial (x, y)} \right\}
\]

(13.30)

is the Hamiltonian, a consistent-order approximation to (13.3). We now consider both \(L_s\) and \(H_s\) to be functionals of the new variables

\[
x_s = x + \frac{v_s}{f_s}, \quad y_s = y - \frac{u_s}{f_s},
\]

(13.31)

where \(f_s = f(x_s, y_s)\). The approximate dynamics result from the requirement that

\[
\delta \int d\tau L_s = 0
\]

(13.32)

for arbitrary variations \(\delta x_s(a, \tau)\) in the new variables. These variations yield

\[
\delta x_s : \quad -f_s \frac{\partial y_s}{\partial \tau} = -\frac{\delta H_s}{\delta x_s}, \quad \delta y_s : \quad +f_s \frac{\partial x_s}{\partial \tau} = -\frac{\delta H_s}{\delta y_s}.
\]

(13.33)
Unfortunately, the functional derivatives of $H_s$ are somewhat awkward to compute, because $H_s$ cannot be written simply solely in terms of $x_s$. We assume for convenience that the fluid is unbounded and quiescent at infinity. By (13.30)

$$
\delta H_s = \iint da \left\{ u_s \delta u_s + v_s \delta v_s + \frac{1}{2} g \delta \frac{\partial (a,b)}{\partial (x,y)} \right\}
$$

$$
= \iint da \left\{ u_s \delta u_s + v_s \delta v_s + g \frac{\partial h}{\partial x} \delta x + g \frac{\partial h}{\partial y} \delta y \right\}.
$$

(13.34)

By (13.28) and (13.31), this is

$$
\delta H_s = \iint da \left\{ u_s \delta u_s + v_s \delta v_s + f_s v_s \delta x_s - f_s u_s \delta y_s \right\}
$$

$$
= \iint da \left\{ f_s v_s \delta x_s - f_s u_s \delta y_s + \frac{u_s^2 + v_s^2}{f_s} \delta f_s \right\}
$$

(13.35)

Thus

$$
\frac{\partial H_s}{\partial x_s} = f_s v_s + \frac{u_s^2 + v_s^2}{f_s} \frac{\partial f_s}{\partial x_s}, \quad \frac{\partial H_s}{\partial y_s} = -f_s u_s + \frac{u_s^2 + v_s^2}{f_s} \frac{\partial f_s}{\partial y_s}
$$

(13.36)

The dynamics (13.33) thus becomes

$$
\frac{\partial x_s}{\partial \tau} = u_s - \frac{u_s^2 + v_s^2}{f_s^2} \left( \frac{\partial f_s}{\partial y_s} \frac{\partial f_s}{\partial x_s} \right),
$$

(13.37)

where $u_s = (u_s, v_s)$. When $f$ is a constant, (13.37) reduce to a well-known form of the semigeostrophic equations. Thus (13.6) and (13.37) represent the generalization of semigeostrophic dynamics to arbitrary Coriolis parameter.

Unfortunately, the semigeostrophic equations are relatively difficult to solve either in $x$-space or in $x_s$-space. This awkwardness arises from the fact that we designed the transformation (13.31) to simplify the Poisson bracket to the maximum extent possible. It is no surprise, then, that the Hamiltonian $H_s$ is a more complicated functional of $x_s$ than it was of $x$. Is there a way to regain some of the former simplicity of the Hamiltonian without giving up the other advantages of $x_s$? Only if we are willing to make an additional restriction.

First consider the potential-energy term in (13.30). To the first two orders in Rossby number,
\[
\frac{\partial (x,y)}{\partial (a,b)} = \frac{\partial (x, -v_f/y_s + u_s/f_s)}{\partial (a,b)} \approx \frac{\partial (x,y_s)}{\partial (a,b)} \left( 1 - \frac{\partial}{\partial x_s} \left( \frac{v_s}{f_s} \right) + \frac{\partial}{\partial y_s} \left( \frac{u_s}{f_s} \right) \right). \tag{13.38}
\]

That is,
\[
h = h_s \left( 1 + \frac{\partial}{\partial x_s} \left( \frac{v_s}{f_s} \right) - \frac{\partial}{\partial y_s} \left( \frac{u_s}{f_s} \right) \right), \tag{13.39}
\]

where
\[
h_s \equiv \frac{\partial (a,b)}{\partial (x_s,y_s)} \tag{13.40}
\]

is the fluid depth in \( x_s \)-space. Thus we can consistently replace the Hamiltonian (13.30) by
\[
H_s[x_s, \mathbf{u}_s, \tau] = \frac{1}{2} \iint \! d\mathbf{a} \left\{ u_s^2 + v_s^2 + g \frac{\partial (a,b)}{\partial (x_s,y_s)} \left( 1 + \frac{\partial}{\partial x_s} \left( \frac{v_s}{f_s} \right) - \frac{\partial}{\partial y_s} \left( \frac{u_s}{f_s} \right) \right) \right\}. \tag{13.41}
\]

The only vestige of the old coordinates \( (x,y) \) in (13.29) and (13.41) are the \( \partial h / \partial x \)- and \( \partial h / \partial y \)-terms in the definitions (13.28) of \( \mathbf{u}_s \). Suppose then that we replace (13.28) by
\[
u_s^* \equiv -\frac{g}{f(x_s,y_s)} \frac{\partial h_s}{\partial y_s}, \quad v_s^* \equiv \frac{g}{f(x_s,y_s)} \frac{\partial h_s}{\partial x_s}, \tag{13.42}
\]

To maintain consistent accuracy, we must show that \( \mathbf{u}_s^* = \mathbf{u}_s \) at leading order. By (13.39),
\[
v_s = \frac{g}{f_s} \frac{\partial h_s}{\partial x_s} = \frac{g}{f_s} \frac{\partial h}{\partial x_s} \left[ h_s \left( 1 + \frac{\partial}{\partial x_s} \left( \frac{v_s}{f_s} \right) - \frac{\partial}{\partial y_s} \left( \frac{u_s}{f_s} \right) \right) \right] \tag{13.43}
\]
\[
= \frac{g}{f_s} \frac{\partial h_s}{\partial x_s} \left( 1 + \frac{\partial}{\partial x_s} \left( \frac{v_s}{f_s} \right) - \frac{\partial}{\partial y_s} \left( \frac{u_s}{f_s} \right) \right) + \frac{g h_s}{f_s} \frac{\partial}{\partial x_s} \left( \frac{\partial}{\partial y_s} \left( \frac{v_s}{f_s} \right) - \frac{\partial}{\partial y_s} \left( \frac{u_s}{f_s} \right) \right)
\]

(and similarly for \( u_s \)). Thus,
\[
v_s = v_s^* \left( 1 + O \left( \frac{U}{f_0 L} \right) \right) + O \left( \frac{gH}{f_0^2 L^2} U \right). \tag{13.44}
\]
Hence we can replace \( \mathbf{u}_s \) by \( \mathbf{u}_s^* \) if the Rossby number \( \frac{U}{f_0 L} \) is small and if the horizontal lengthscale \( L \) is large compared to the deformation radius \( \sqrt{gH / f_0} \). This extra assumption about the horizontal lengthscale is the aforementioned additional restriction.

If the flow is both nearly geostrophic and large-scale (in the sense that the lengthscale \( L \) is larger than the deformation radius), then we can replace \( \mathbf{u}_s \) by \( \mathbf{u}_s^* \) everywhere (including the definition (13.31) of the \( x_s \)). Then the old coordinates \( \mathbf{x} \) disappear entirely (except insofar as they are needed to transform the results back into \( \mathbf{x} \)-space). Hamilton’s principle leads to

\[
\delta x_s : - f_s \frac{\partial y_s}{\partial \tau} = - \frac{\partial \Phi}{\partial x_s}, \quad \delta y_s : + f_s \frac{\partial x_s}{\partial \tau} = - \frac{\partial \Phi}{\partial y_s},
\]

(13.45)

where

\[
\Phi = \left( \frac{g}{f_s} \right)^2 \nabla_s h_s \cdot \nabla_s h_s + gh_s \left( 1 + \nabla_s \cdot \left( \frac{g}{f_s^2} \nabla_s h_s \right) \right),
\]

(13.46)

and \( \nabla_s \) is the gradient operator in \( x_s \)-space. The \( \tau \)-derivative of (13.40) yields

\[
\frac{\partial h_s}{\partial \tau} + \frac{\partial}{\partial x_s} \left( h_s \frac{\partial x_s}{\partial \tau} \right) + \frac{\partial}{\partial y_s} \left( h_s \frac{\partial y_s}{\partial \tau} \right) = 0.
\]

(13.47)

Substituting (13.45) into (13.47) yields a closed evolution equation for \( h(x_s, y_s, t_s) \). The latter is equivalent to the potential vorticity equation,

\[
\left( \frac{\partial}{\partial x_s} + \frac{\partial x_s}{\partial \tau} \frac{\partial}{\partial x_s} + \frac{\partial y_s}{\partial \tau} \frac{\partial}{\partial y_s} \right) f_s \frac{h_s}{h_s} = 0.
\]

(13.48)

Once again, the conservation of potential vorticity, (13.48), is guaranteed by the particle-relabeling symmetry property of the approximation. Despite its appearance, the potential vorticity \( f_s/h_s \) includes a consistent-order approximation to the relative vorticity, hidden in the formula (13.31) for the transformation of coordinates.

In the ocean, the external deformation radius is about 2000 km; it makes no sense to assume that the lengthscale \( L \) is larger than that! However, the (first) internal deformation radius is much smaller (about 40 km), and even the largest numerical ocean circulation models barely resolve it. Thus, if we regard the shallow-water equations as a paradigm for an internal mode of the primitive equations, then the simplifications leading from semigeostrophy to (13.45) make sense.

To apply this strategy to the full primitive equations, it is necessary to divide the motion into a depth-averaged (barotropic) part and a (baroclinic) departure therefrom. The rigid-lid approximation filters out the barotropic inertia-gravity waves, and the baroclinic inertia-gravity waves are filtered by a geostrophic-balance constraint that applies to the baroclinic motion only. The details are numerous, but the key ideas —
that Hamiltonian methods promote approximations that maintain conservation laws, and coordinate transformations that dramatically simplify the equations — remain the same.

Notes for Chapter 7.
1. For reviews of Hamiltonian fluid dynamics see Morrison (1994), Shepherd (1990,1994) and Salmon (1988a).
2. In fact, we shall not use Noether’s theorem at all, preferring instead the modified form of the theory presented by Lanczos (1970, Appendix II). This modified theory can best be considered as a special case of Hamilton’s principle, in which the variations considered are inspired by the symmetries of the Hamiltonian alone. For a thorough discussion of Noether’s original theory, see for example Olver (1986).
3. See Whitham (1975) and references therein.
5. In seawater, the dependence of $E(\alpha,\eta,S)$ on the salinity $S$ also destroys the particle-relabeling symmetry (unless the buoyancy depends linearly on $\eta$ and $S$).
6. The term mean flow is appropriate because we eventually assume that the disturbance has a zero average (of some kind). In Section 12, we consider disturbances with nonzero (or undefined) averages. Sections 3, 4, and 5 are based upon the seminal paper by Bretherton (1971).
7. One must be careful not to overstate this conclusion. The circulation theorem (5.19) applies to a loop moving at the mean velocity of the fluid particles. If the waves were to produce a permanent change in the location of this loop, they could cause a permanent change in the mean flow despite the conservation law. See Bühler and McIntyre (1996).
8. This was the remarkable realization of Andrews and McIntyre (1978a,b). See also Hayes (1970) and the review paper by Grimshaw (1984).
9. This further approximation can also be looked upon as a constraint, that the vertical velocity vanish. However, this requires the extended form of Hamilton's principle (Section 7), in which the vertical velocity is varied independently.
10. See Green and Naghdi (1976) and Miles and Salmon (1985).
11. See Miles and Salmon (1985) and Section 9 of Salmon (1988a).
12. See Appendix B of Miles and Salmon (1985).
13. If we regard the functional $F[f(a)]$ as (the limit of) a sum over the gridpoints on a grid that is uniform in $a$, then the functional is an ordinary function of the grid-values, and the functional derivative is an ordinary derivative. This interpretation makes it easy to understand why $\partial F/\partial f(a)$ is not the same as $\partial F/\partial f(x)$: A grid that is uniform in $a$-space is not (generally) uniform in $x$-space.
15. If the fluid is nonhomentropic, then $C$ is not conserved following a fluid particle, and therefore cannot be regarded as one of the vorticity-labeling coordinates. However, in this case, vorticity is not conserved, even in the sense of Chapter 4, Section 3. In nonhomentropic flow, the right-hand side of (8.16b) gives rise to the pressure-torque term in the general vorticity equation (1.12) of Chapter 4.
16. For an introduction to tensor analysis, see Synge and Schild (1952) or (for a more modern approach) Schutz (1980). However, I believe that the discussion in Section 9 is self-contained and conveys the important concept of covariance.
17. See Littlejohn (1979) or Sudarshan and Mukunda (1983)

18. The analogy between the fluid variables and the discrete sets \(\{z^i\}\) and \(\{z^\alpha\}\) (where \(i\) is analogous to \(a\) and \(\alpha\) is analogous to \(x\)) is imprecise for two reasons. First, \(a\) and \(x\) vary continuously, whereas the superscripts take only integer values. Second, the transformation from Lagrangian to Eulerian fluid variables involves a change in both the dependent and the independent variables.

19. Actually, if one accepts (10.5), then (10.10) follows: In the coordinates for which (10.5) holds, the \(C^{(k)}\) are just the last \(K\) coordinates. Then, since (10.11) is covariant, these \(C^{(k)}\) satisfy (10.10) in any set of coordinates. In fact, the proofs of Darboux’s theorem and the existence of Casimirs are closely related; see Littlejohn (1982). However, it is still worthwhile to sketch the proof of (10.11) in order to appreciate the crucial importance of Jacobi’s identity.


22. The Casimirs for the full primitive equations obey functional differential equations that are analogous to (11.35) but much more complicated. I am unaware of any proof that (11.37) is their general solution.


25. See McIntyre and Shepherd (1987) and references therein.

26. For a more complete account, see the reviews by Abarbanel et al. (1986), Holm et al. (1985), and Shepherd (1990, 1994).

27. This is a form of Noether’s theorem; see Benjamin (1984).

28. Once again, Shepherd (1994) provides a good introduction.

29. This section is based upon the papers by Salmon (1983,1985,1988b), which give many further details.

30. For details, see Salmon (1985, Section 2 and Appendix A).

31. There is a completely different way to look at this, in which the Poisson bracket is held fixed, and the geostrophic constraints (13.13) are instead perturbed. See Salmon (1988b).

32. The pioneering papers on the semigeostrophic equations are by Hoskins and Bretherton (1972) and Hoskins (1975). See also Cullen and Purser (1989), Purser (1993) and references therein.

33. The details can be found in Salmon (1985).

34. See Salmon (1996).