Hamiltonian description of the ideal fluid*

P. J. Morrison[†]

Department of Physics and Institute for Fusion Studies, University of Texas, Austin, Texas 78712-1060

The Hamiltonian viewpoint of fluid mechanical systems with few and infinite number of degrees of freedom is described. Rudimentary concepts of finite-degree-of-freedom Hamiltonian dynamics are reviewed, in the context of the passive advection of a scalar or tracer field by a fluid. The notions of integrability, invariant-tori, chaos, overlap criteria, and invariant-tori breakup are described in this context. Preparatory to the introduction of field theories, systems with an infinite number of degrees of freedom, elements of functional calculus and action principles of mechanics are reviewed. The action principle for the ideal compressible fluid is described in terms of Lagrangian or material variables. Hamiltonian systems in terms of noncanonical variables are presented, including several examples of Eulerian or inviscid fluid dynamics. Lie group theory sufficient for the treatment of reduction is reviewed. The reduction from Lagrangian to Eulerian variables is treated along with Clebsch variable decompositions. Stability in the canonical and noncanonical Hamiltonian contexts is described. Sufficient conditions for stability, such as Rayleigh-like criteria, are seen to be only sufficient in the general case because of the existence of negative-energy modes, which are possessed by interesting fluid equilibria. Linearly stable equilibria with negative energy modes are argued to be unstable when nonlinearity or dissipation is added. The energy-Casimir method is discussed and a variant of it that depends upon the notion of dynamical accessibility is described. The energy content of a perturbation about a general fluid equilibrium is calculated using three methods. [S0034-6861(98)00102-0]

CONTENTS

I. INTRODUCTION

Why look at fluid mechanics from a Hamiltonian perspective? The simple answer is because it is there and it is beautiful. For ideal fluids the Hamiltonian form is not artificial or contrived, but something that is basic to the model. However, if you are a meteorologist or an oceanographer, perhaps what you consider to be beautiful is the ability to predict the weather next week or to understand transport caused by ocean currents. If this is the case, a more practical answer may be needed. Below, in the remainder of this Introduction, I shall give some arguments to this effect. However, I have observed that the Hamiltonian philosophy is like avocado: you either like it or you don't. In any event, since 1980 I have also observed a strong development in this field, and this is very likely to continue.

One practical reason for the Hamiltonian point of view is that it provides a unifying framework. In particular, when solving "real" problems one makes approximations about what the dominant physics is, considers different geometries, defines small parameters, expands, etc. In the course of doing this one performs various kinds of calculations again and again, for example, calculations regarding

(1) waves and instabilities by means of linear eigenanalyses;

(2) parameter dependency of eigenvalues as obtained by such eigenanalyses;

(3) stability that are based on arguments involving energy or other invariants;

(4) various kinds of perturbation theory;

(5) approximations that lead to low-degree-of-freedom dynamics.

After a while one discovers that certain things happen over and over again in the above calculations, for example,

(1) spectra whose nature is not arbitrary, but possesses limitations;

(2) certain types of bifurcations that occur upon collision of eigenvalues;

(3) Rayleigh-type stability criteria (these occur for a wide variety of fluid and plasma problems);

(4) simplifications based on common patterns;

(5) common methods for reducing the order of systems.

By understanding the Hamiltonian perspective, one knows in advance (within bounds) what answers to expect and what kinds of procedures can be performed.

In cases where dissipation is not important and approximations are going to be made, it is, in my opinion, desirable to have the approximate model retain the Hamiltonian structure of the primitive model. One may not want to introduce spurious unphysical dissipation or sources that destroy energy conservation or other conserved physical quantities. Understanding the Hamiltonian structure allows one to make Hamiltonian approximations. In physical situations where dissipation is important, I believe it is useful to see in which way the dynamics differ from what one expects for the ideal (dis-
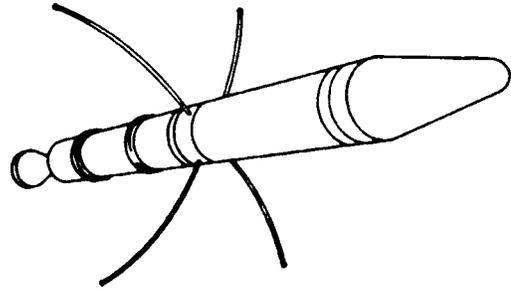
FIG. 1. Depiction of Explorer I satellite, which was destabilized by energy dissipation (after Likins, 1971).

sipationless) model. The Hamiltonian model thus serves as a sort of benchmark. Also, when approximating models with dissipation, we can isolate which part is dissipative and make sure that the Hamiltonian part retains its Hamiltonian structure and so on.

It is well known that Hamiltonian systems are not structurally stable in a strict mathematical sense (which I shall not define here). However, this obviously does not mean that Hamiltonian systems are not important; the physics point of view can differ from the mathematics. A simple linear oscillator with very small damping can behave over long periods of time like an undamped oscillator, even though the topology of its dynamics is quite different.

To say that a Hamiltonian system is structurally unstable is not enough. A favorite example of mine that illustrates this point concerns the first U.S. satellite, Explorer I, which was launched in 1958 (see Fig. 1). This spacecraft was designed so that its attitude would be stabilized by spin about its symmetry axis. However, the intended spin-stabilized state did not persist and the satellite began to tumble. This was attributed to energy dissipation in the small antennae shown in the figure. Thus, unlike the simple oscillator, in which the addition of dissipation has a small effect, here the addition of dissipation had a catastrophic effect. Indeed, this was a most expensive experiment on *negative-energy modes*, a universal phenomenon in fluids that I will discuss.

After Explorer I, in 1962 Alouette I was launched (see Fig. 2), which had an obvious design difference. This satellite behaved like a damped linear oscillator in
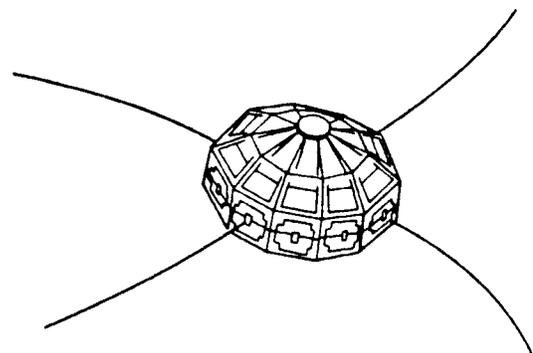
FIG. 2. Depiction of the satellite Alouette (after Likins, 1971), which was designed after Explorer I.

the sense that dissipation merely caused it to spin down. I should like to emphasize that the difference between the behavior of Explorer I and that of Alouette I lies in a mathematical property of the Hamiltonian dynamics of these spacecraft: it could have been predicted.

So, the purpose of my lectures is to describe the Hamiltonian point of view in fluid mechanics, and to do so in an accessible language. It is to give you some fairly general tools and tricks. I am not going to solve a single "real" problem; however, you will see specific examples of problems throughout the summer.[1] The first lecture (Sec. II) is somewhat different in flavor from the others. Imagine that you have succeeded in obtaining a finite Hamiltonian system out of some fluid model, the Kida vortex being a good example (see, for example, Meacham *et al.*, 1997). What should you expect of the dynamics? This first lecture, being a sketch of low-degree-of-freedom Hamiltonian dynamics, answers this to some degree. The next lectures (Secs. III–V) are concerned with the structure of infinite-degree-of-freedom Hamiltonian systems, although I shall often use finite systems for means of exposition. The last lecture (Sec. VI) is concerned with expressions for the energy of perturbations of an equilibrium state and their use in determining stability.

II. RUDIMENTS OF HAMILTONIAN SYSTEMS WITH FEW DEGREES OF FREEDOM, ILLUSTRATED BY PASSIVE ADVECTION IN TWO-DIMENSIONAL FLUIDS

In this introductory lecture we shall review some basic aspects of Hamiltonian systems with a finite number of degrees of freedom.[2] We illustrate, in particular, properties of systems with one, two, and three degrees of freedom by considering the passive advection of a tracer in two-dimensional incompressible fluid flow.[3] The tracer is something that moves with, but does not influence, the fluid flow; examples include neutrally buoyant particles and colored dye. The reason for mixing Hamiltonian system phenomenology with fluid advection is that the latter provides a nice framework for visualization, since, as we shall see, the phase space of the Hamiltonian system is in fact the physical space occupied by the fluid.

A point of view advocated in this lecture series is that an understanding of finite-dimensional Hamiltonian systems is useful for the eventual understanding of infinite-degree-of-freedom systems, such as the equations of various ideal fluid models. Such infinite systems are the main subject of these lectures. It is important to understand that the infinite systems are distinct from the pas-

sive advection problem that is treated in this lecture; the former are governed by partial differential equations, while the latter is governed by ordinary differential equations.

A. A model for two-dimensional fluid motion

In various situations fluids are described adequately by models in which motion occurs in only two spatial dimensions. An important example is that of rotating fluids in which the dominant physics is governed by geostrophic balance, where the pressure force is balanced by the Coriolis force. For these types of flows the well-known Taylor-Proudman theorem (see, for example, Pedlosky, 1987) states that the motion is predominantly two dimensional. A sort of general model that describes a variety of two-dimensional fluid motion is given by

$$\frac{\partial q}{\partial t} + [\psi, q] = \mathcal{S} + \mathcal{D}, \tag{1}$$

where $q(x, y, t)$ is a vorticity-like variable, $\psi(x, y, t)$ is a stream function, and both are functions of the spatial variable $(x, y) \in D$, where D is some spatial domain, and t is time. The quantities $\mathcal{S}$ and $\mathcal{D}$ denote sources and sinks, respectively. Examples of $\mathcal{S}$ include the input of vorticity by means of pumping or stirring, while examples of $\mathcal{D}$ include viscous dissipation and Ekman drag. Above, the Poisson bracket notation is used:

$$[f, g] := \frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x}, \tag{2}$$

[which is the Jacobian $\partial(f, g)/\partial(x, y)$] and we have assumed incompressible flow, which implies that the two components of the velocity field are given by

$$(u, v) = \left(-\frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x} \right). \tag{3}$$

In order to close the system, a "self-consistency" condition that relates q and ψ is required. We signify this by $q = \mathcal{L}\psi$. Examples include

(1) The two-dimensional Euler equation for which $q = \nabla^2 \psi$;

(2) The rotating fluid on the β-plane for which $q = \nabla^2 \psi + \beta y$.

In the former case q is the vorticity, while in the latter case q is the potential vorticity. Potential vorticity is a vorticity-like quantity that includes the "β-effect," an effect that arises in part from the deviation of the normal to the earth's surface with the rotation axes (see Pedlosky, 1987).

For convenience we shall suppose that the domain D is an annular region as depicted in Fig. 3. Many experiments have been performed in this geometry,[4] where the fluid swirls about in the θ and r directions and is pre-

---

[1] I have tried to retain language that preserves the flavor of the Woods Hole lecture series.

[2] For an introduction see Berry (1978) and the comprehensive article of Arnold (1963); for symplectic maps see Meiss (1992).

[3] Much has been written about passive advection; see, for example, Ottino (1990), del-Castillo-Negrete and Morrison (1993), and many references therein.

[4] Early experiments, along with theory, are described by Greenspan (1968); more recent work is discussed by Sommeria *et al.* (1991).
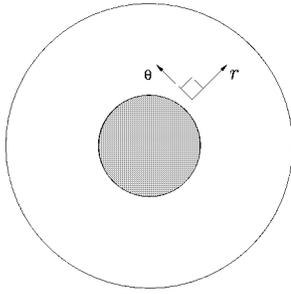
FIG. 3. Sketch of annular region with coordinates for rotating tank experiments.

dominantly two dimensional. The geometry of the annulus suggests the use of polar coordinates, which are given here by the formulas $x = r\sin\theta$ and $y = r\cos\theta$. In terms of r and θ the bracket of Eq. (2) becomes

$$[f,g] = \frac{1}{r}\left(\frac{\partial f}{\partial \theta}\frac{\partial g}{\partial r} - \frac{\partial f}{\partial r}\frac{\partial g}{\partial \theta}\right). \qquad (4)$$

The spatial variables (x,y) play the role below of canonical coordinates, with x being the configuration-space variable and y being the canonical momentum. The transformation from (x,y) to (r,θ) is a noncanonical transformation, and so the form of the Poisson bracket is altered as manifested by the factor of $1/r$. (In Sec. IV we shall discuss this in detail.) To preserve the canonical form we replace r by a new coordinate $J := r^2/2$ and the bracket becomes

$$[f,g] = \frac{\partial f}{\partial \theta}\frac{\partial g}{\partial J} - \frac{\partial f}{\partial J}\frac{\partial g}{\partial \theta}. \qquad (5)$$

These coordinates are convenient, since they can be action-angle variables, as we shall see.

A solution to Eq. (1) provides a stream function, $\psi(\theta,J,t)$. In this lecture we shall assume various forms for ψ, without going into detail as to whether or not these forms are solutions to Eq. (1) with particular choices of $\mathcal{L}$, $\mathcal{S}$, or $\mathcal{D}$. Here we shall just suppose that the tracers in the fluid, specks of dust if you like, follow particular assumed forms for the velocity field of the flow. The stream function gives a means for visualizing this. Setting ψ equal to a constant for some particular time defines an instantaneous streamline whose tangent is parallel to the velocity field. (See Fig. 4.)
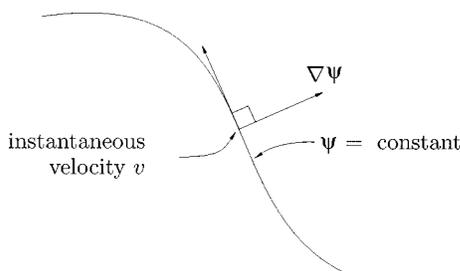
FIG. 4. A streamline defined in terms of ψ with instantaneous velocity.

B. Passive advection

Imagine that a tiny piece of the fluid is labeled, somehow, in such a way that it can be followed. As mentioned above, a small neutrally buoyant sphere or a small speck of dust might serve this purpose. Since such a tracer, the sphere or the speck, moves with the fluid, its dynamics is governed by

$$\dot{y} = v = \frac{\partial \psi}{\partial x} = [y, -\psi], \quad \dot{x} = u = -\frac{\partial \psi}{\partial y} = [x, -\psi], \quad (6)$$

or, in terms of the (θ,J) variables,

$$\dot{J} = \frac{\partial \psi}{\partial \theta}, \quad \dot{\theta} = -\frac{\partial \psi}{\partial J}. \qquad (7)$$

(Note: $\dot{} = d/dt$.) These equations are of the form of Hamilton's equations, which are usually written as

$$\dot{p}_i = [p_i, H] = -\frac{\partial H}{\partial q^i}, \quad \dot{q}^i = [q^i, H] = \frac{\partial H}{\partial p_i}, \qquad (8)$$

where $i = 1,2,\ldots N$, and the *Poisson bracket*, $[\ ,\ ]$, is defined by

$$[f,g] = \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i}. \qquad (9)$$

Here and henceforth we use repeated index sum notation. The quantities (q^i, p_i) constitute a set of canonically conjugate pairs with q^i being the canonical coordinate and p_i being the canonical momentum. Together they are coordinates for the $2N$-dimensional *phase space*. The function $H(q,p,t)$ is the Hamiltonian. Observe that y (or J), which physically is a coordinate, here plays the role of momentum, and $-\psi$ is the Hamiltonian.

We emphasize, once again, that the coordinates (x,y) are coordinates of a tracer, and the motion of the tracer is determined by a prescribed velocity field. This is to be distinguished from the Lagrangian variable description of the ideal fluid, which we treat in Sec. III, where the goal is to describe the velocity field as determined by the solution of a partial differential equation.

Before closing this subsection we give a bit of terminology. A single *degree of freedom* corresponds to each (q,p) pair. However, some account should be given of whether or not H depends explicitly upon time. It is well known that nonautonomous ordinary differential equations can be converted into autonomous ones by adding a dimension. Therefore researchers sometimes count a half of a degree of freedom for this. Thus Eq. (7) is a $1\frac{1}{2}$-degree-of-freedom system if ψ depends explicitly upon time; otherwise it is a one-degree-of-freedom system. This accounting is not very precise, since one might want to distinguish between different types of time dependency. We shall return to this point later.

C. Integrable systems: One degree of freedom

All one-degree-of-freedom systems are integrable. However, integrable systems of higher dimension are

rare in spite of the fact that some mechanics texts make them the centerpiece (if not the only piece). A theorem often credited to Siegel (see, for example, Moser, 1973) shows how integrable systems are of measure zero. What exactly it means to be integrable is an active area of research with a certain amount of subjectivity. For us, integrable systems will be those for which the motion is determined by the evaluation of N integrals. When this is the case, the motion is "simple" in the appropriate coordinates.

More formally, a system with a time-independent Hamiltonian $H(q,p)$ with N degrees of freedom is said to be *integrable* if there exist N independent, smooth constants of motion I_i, i.e.,

$$\dot{I}_i = [I_i, H] = 0, \tag{10}$$

that are in *involution*, i.e.,

$$[I_i, I_j] = 0. \tag{11}$$

The reason that the constants are required to be smooth and independent is that the equations $I_i = c_i$, where the c_i's are constants, must define N different surfaces of dimension $2N-1$ in the $2N$-dimensional phase space. The reason for the constants to be in involution is that one wants to use the I's (or combinations of them) as momenta, and momenta must pairwise commute. In coordinates of this type the motion is quite simple.

Sometimes additional requirements are added in definitions of integrability. For example, one can add the requirements that the surfaces $I_i = \text{const}$ for $i = 1, 2, \ldots, N$ be *compact* and *connected*.[5] If this is the case the motion takes place on an N-torus and there exist *action-angle variables* (J_i, θ^i) in terms of which Hamilton's equations have the form

$$\dot{J}_i = -\frac{\partial H}{\partial \theta^i} = 0, \quad \dot{\theta}^i = \frac{\partial H}{\partial J_i} = \Omega^i(J), \tag{12}$$

where $i, j = 1, 2, \ldots, N$. The first of Eqs. (12) implies that H is a function of J alone. When H does not depend upon a coordinate, the coordinate is said to be *ignorable* and its conjugate momentum is a constant of motion. In action-angle variables all coordinates are ignorable and the second of Eqs. (12) is easy to integrate, yielding

$$\theta^i = \theta_0^i + \Omega^i(J) \, t, \tag{13}$$

where θ_0^i is the integration constant, θ is defined *modulo* 2π, and $\Omega^i(J) := \partial H/\partial J_i$ are the frequencies of motion around the N-torus.

A good deal of the machinery of Hamiltonian mechanics was developed in the attempt to reduce equations to the action-angle form above. If one could find a coordinate transformation, in particular a canonical transformation (see Sec. IV), that took the system of

interest into the form of Eq. (12), then one could simply integrate and then map back to get the solution in closed form. The theory of canonical transformations, Hamilton-Jacobi theory, etc. sprang up because of this idea. However, it is now known that this procedure is not possible in general because generically Hamiltonian systems are not integrable. Typically systems are *chaotic*, i.e., trajectories wander in a seemingly random way in phase space rather than lying on an N-dimensional torus. A distinct feature of such trajectories is that they display *sensitive dependence* on initial conditions. We shall say a little about this below.

To conclude this subsection we return to our fluid mechanics example, in which context we show how all one-degree-of-freedom systems are integrable. In the case where ψ is time independent, we clearly have a single degree of freedom with one constant of motion, viz., ψ:

$$\dot{\psi} = \frac{\partial \psi}{\partial x} \dot{x} + \frac{\partial \psi}{\partial y} \dot{y} = 0, \tag{14}$$

which follows upon substitution of the equations of motion for the tracer, Eq. (6). To integrate the system one solves

$$\psi(x,y) = \psi^0 = \text{const.} \tag{15}$$

for $x = f(\psi_0, y)$, which is in principle (if not in practice) possible, and then inserts the result as follows:

$$\dot{y} = \frac{\partial \psi}{\partial x}(x,y) \bigg|_{x = f(\psi_0, y)} =: D(\psi^0, y). \tag{16}$$

Equation (16) is separable, which implies

$$\int_{y_0}^{y} \frac{dy'}{D(\psi^0, y)} = \int_{t_0}^{t} dt'. \tag{17}$$

Thus we have reduced the system to the evaluation of a single integral, a so-called quadrature. There are some sticky points, though, since $x = f(\psi^0, y)$ may not be single valued or explicitly invertible, and usually one cannot do the integral explicitly. Moreover, afterwards one must invert Eq. (17) to obtain the trajectory. These are only technical problems, ones that are easily surmounted with modern computers.

Generally equations of the form of Eq. (1) possess equilibrium or steady-state solutions when ψ and q depend upon only a single coordinate. The case of special interest here is that in which the domain is the annulus discussed above, polar coordinates are used, and ψ depends only upon r (or equivalently the canonical variable J). Physically this corresponds to a purely azimuthally symmetric, sheared fluid flow, where $v_\theta = v_\theta(r)$. In this case streamlines are "energy surfaces," which are merely concentric circles as depicted in Fig. 5. The counterpart of Eq. (13), the equations of motion for the speck of dust in the fluid, are

$$\theta = \theta^0 + \Omega(r)t, \quad r = r_0, \tag{18}$$

where $v_\theta = \Omega r$. Note that the speck goes round and round at a rate dependent upon its radius, but does not go in or out.

---

[5]Intuitively, one can think of the surface as contained within a hypersphere of some (noninfinite) radius, and between any two points of the surface a line can be drawn within the surface (see, for example, Arnold, 1963).
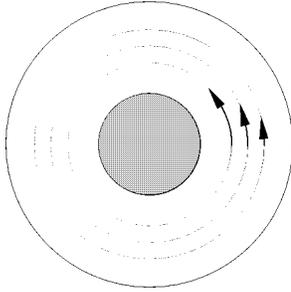
FIG. 5. Depiction of invariant circle streamlines for an azimuthally symmetric flow.

D. Chaotic dynamics: Two degrees of freedom

As noted, one-degree-of-freedom systems are always integrable, but two-degree-of-freedom systems typically are not. Nonintegrable systems exhibit chaos, which we briefly describe below.

Systems with two degrees of freedom have a four-dimensional phase space, which is difficult to visualize, so we do something else. A convenient artifice is the *surface of section* or as it is sometimes called the *Poincaré section*. Suppose the surface $H(q_1,q_2,p_1,p_2) = \text{const} =: E$ is compact (i.e., contained within a three-sphere). Since the motion is restricted to this surface, p_2 can be eliminated in favor of E, which we keep fixed. We could then plot the trajectory in the space with the coordinates (q_1,q_2,p_1), but simpler pictures are obtained if we instead plot a point in the (q_1,p_1) plane whenever q_2 returns to its initial value, say $q_2=0$.

We also require that the trajectory pierce this plane with the momentum p_2 having the same sign upon each piercing. This separates out the branches of the surface $H=E$. That q_2 will return is almost assured, since the Poincaré recurrence theorem[6] tells us that almost any orbit will return to within any ϵ-ball (points interior to a sphere of radius ϵ). It is unlikely it will traverse the ball without piercing $q_2=0$. [If there are no fixed points within the ball the vector field can be locally rectified, and unless there is no component normal to the (q_1,p_1) plane, which is unlikely, it will pierce.]

For integrable systems, an orbit either eventually returns to itself, in which case we have a *periodic orbit*, or it maps out a curve, which is an example of an *invariant set*. The latter case is typical, as illustrated in Fig. 6. In nonintegrable or chaotic systems this is not true, as is illustrated in Figs. 7(a) and 7(b), where it is seen that orbits make "erratic" patterns.

Now what about the fluid mechanics illustration? Can chaos exist? How can we have a two-degree-of-freedom system when we have only the two spatial coordinates, say (θ,J)? The answer is that explicit time dependence in ψ, the extra half degree of freedom, is enough for chaos. There is, in fact, a trick for *puffing up* a $1\frac{1}{2}$-degree-of-freedom system and making it look like a

---

[6]See Wintner (1947), which contains references to several original papers.
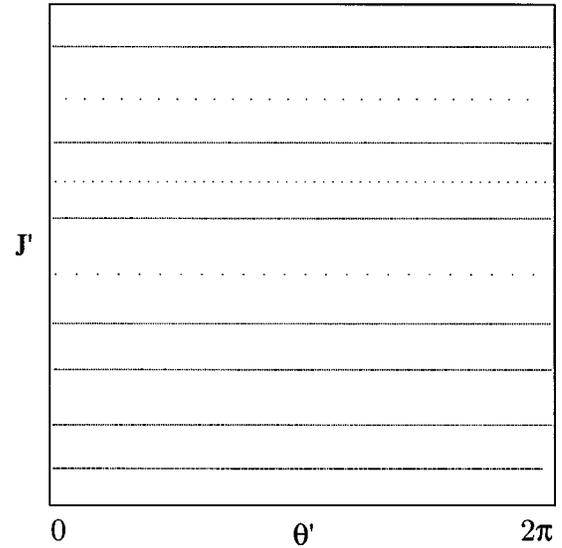
FIG. 6. Surface of section for an integrable system.

two-degree-of-freedom system, and vice versa.

Let s correspond to a fake time variable, set $t=\phi$, where ϕ is going to be a new canonical coordinate, and define a new Hamiltonian by

$$H(\theta,J,\phi,I) = -\psi(\theta,J,\phi) + I. \tag{19}$$

The equations of motion for this Hamiltonian are

$$\frac{d\theta}{ds} = \frac{\partial H}{\partial J} = -\frac{\partial \psi}{\partial J}, \quad \frac{dJ}{ds} = -\frac{\partial H}{\partial \theta} = \frac{\partial \psi}{\partial \theta}, \tag{20}$$

$$\frac{d\phi}{ds} = \frac{\partial H}{\partial I} = 1, \quad \frac{dI}{ds} = -\frac{\partial H}{\partial \phi} = \frac{\partial \psi}{\partial \phi}. \tag{21}$$

The first of Eqs. (21) tells us that $\phi=s+s_0=t$; we set $s_0=0$. Thus we obtain what we already knew, namely, that $\phi=t$ and that Eqs. (20) give the correct equations of motion. What is the role of the second of Eqs. (21)? This equation merely tells us that I has to change so as to make $H=\text{const}$.

The above trick becomes particularly useful when ψ is a periodic function of time: $\psi(\theta,J,t) = \psi(\theta,J,t+T)$. In this case it makes sense to identify $\phi+T$ with ϕ, because the velocity field is the same at these points. With this identification done, it is clear that a surface of section is obtained by plotting (θ,J) at intervals of T.

One can construct $1\frac{1}{2}$-degree-of-freedom Hamiltonian systems from two-degree-of-freedom Hamiltonian systems by using one of the configuration-space coordinates as a time variable. We leave the details of this calculation as an exercise.

Now suppose the stream function is composed of an azimuthal shear flow plus a propagating wave:

$$\psi(J,\theta,t) = \psi_0(J) + \psi_1(J)\cos[m_1(\theta-\omega_1 t)], \tag{22}$$

where $m_1 \in \mathbb{N}$ (i.e., m_1 is a natural number) and ψ_1 is assumed small in comparison to ψ_0. Here $\psi_0(J)$ represents the azimuthal background shear flow and the second term represents the wave, with ψ_1, m_1, and ω_1 being the radial eigenfunction, mode number, and frequency of the wave, respectively.

(a)

(b)

FIG. 7. Surface of section for a typical nonintegrable system, the standard map of Eq. (26) at (a) $k = 0.80$; (b) $k \approx 0.97$.

This system might look like a $1\frac{1}{2}$-degree-of-freedom system, but it is in fact integrable. The easiest way to see this is to boost it into the frame of reference rotating with the wave. In this frame the stream function becomes

$$\psi'(J',\theta',t) = \psi_0(J') + \psi_1(J')\cos(m_1\theta') - \omega_1 J', \quad (23)$$

where the transformation is $J' = J$, $\theta' = \theta - \omega_1 t$. This transformation is canonical and is derivable from the mixed-variable generating function $F(\theta,J') = J'(\theta - \omega_1 t)$. Note that the term $-\omega_1 J'$ accounts for the azimuthal rigid rotation generated from the frame shift.

In the absence of the wave it is clear that the trajectories in phase space are just circles, as shown in Fig. 5 (or straight lines, as plotted in Fig. 6). However, from the form of Eq. (22) it is clear that something interesting
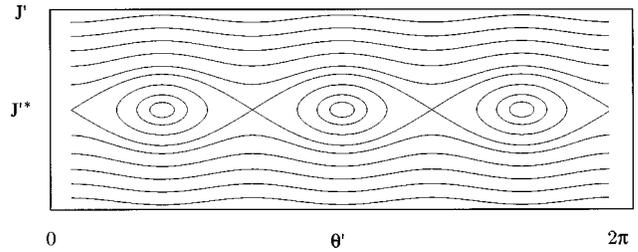
FIG. 8. Depiction of a single resonance in the surface of section.

is going to happen at stagnation points, that is, where

$$\frac{\partial \psi}{\partial \theta'} = \frac{\partial \psi}{\partial J'} = 0. \quad (24)$$

Stagnation points occur at places where the phase velocity of the wave matches the background azimuthal velocity. Here a *critical layer* opens up into an island chain. In the terminology of Hamiltonian dynamics this is called a *resonance* and looks as depicted in Fig. 8.

From the picture it is clear that orbits lie on surfaces and from the form of the stream function given by Eq. (23) it is clear that the motion can be solved by quadrature. The use of the coordinate $\theta' = \theta - \omega_1 t$ reduces this system to a single degree of freedom.

As noted above, the fact that we could reduce the $1\frac{1}{2}$-degree-of-freedom system to a single degree of freedom is the exception rather than the rule; generically it is not possible to get rid of time dependence by changing coordinate frames. This is the case, for example, for an azimuthal shear flow with the presence of two waves with *different* phase velocities, which has the stream function

$$\psi(J,\theta,t) = \psi_0(J) + \psi_1(J)\cos[m_1(\theta - \omega_1 t)]$$
$$+ \psi_2(J)\cos[m_2(\theta - \omega_2 t)]. \quad (25)$$

It is clear that in this case a frame no longer exists in which the flow is stationary. In general there will be chaotic motion of a tracer particle. In a frame moving at a phase velocity ω_1 a tracer particle wants to execute its integrable motion, as described above. However, it is perturbed by a time-dependent wave propagating by at a speed $|\omega_1 - \omega_2|$. In a frame moving at ω_2 the situation is reversed. A plot of both of the integrable motions, in their respective frames, is shown in Fig. 9. This is a plot of Eq. (23) for the first wave superimposed on a plot of the same function for the second wave, but with ψ_1, m_1, and ω_1 replaced by ψ_2, m_2, and ω_2. The form of ψ_1 and ψ_2 is chosen in this figure to be proportional to sech^2; the angle θ' is $\theta - \omega_1 t$ for the first wave and $\theta - \omega_2 t$ for the second. If the distance between the island chains is large, then this figure closely approximates the surface of section. The figure, in fact, suggests a basic mechanism of Hamiltonian chaos, the competition between resonances. If the resonances are close enough together a trajectory, in a sense, flips back and forth between the two integrable motions. When this happens a given trajectory may no longer map out a continuous curve. Gen-
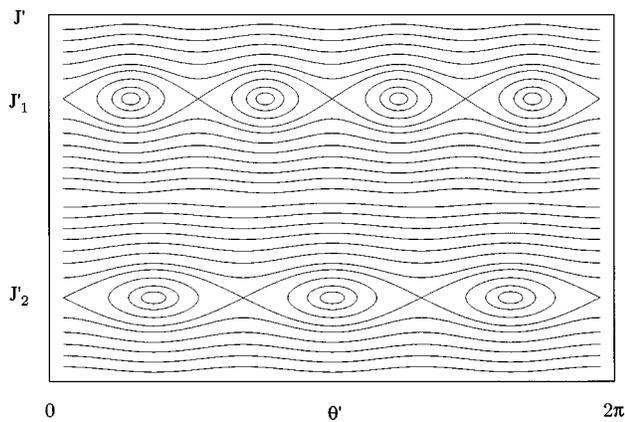
FIG. 9. Depiction of two resonances in the surface of section.

FIG. 10. Cartoon depicting the motion around invariant tori in systems with greater than two degrees of freedom.

erally separatrices become fuzzy, but some continuous curves still exist, as shown in Fig. 7.

As stated above, Fig. 9 is not a surface of section, because the resonances were plotted independently, but if they are far apart and the amplitudes of the resonances are both small it looks about right. To see the real surface of section one could integrate the differential equations numerically.[7] Instead of doing this you can consider the following toy, actually a serious toy, called the *standard map* (which is sometimes called the Chirikov-Taylor map):

$$\theta'_{n+1} = \theta'_n + J'_{n+1},$$
$$J'_{n+1} = J'_n - k \sin(\theta'_n), \qquad (26)$$

where J'_n and θ'_n are computed *modulo* 2π. This is an example of an area-preserving map; it was, in fact, used to obtain Figs. 6, 7(a), and 7(b). Area-preserving maps are nice because the surface of section can be obtained without having to iterate differential equations. Importantly, the standard map describes generic behavior of Hamiltonian systems near resonances—it is the prototype of area-preserving maps.

I recommend that you examine the standard map starting from $k=0$, gradually increasing k. The case in which $k=0$ was shown in Fig. 6, which clearly indicates integrable behavior. For $k \neq 0$ some of the invariant sets (continuous curves) are broken. As $k \to 0$ the measure of invariant sets approaches unity. This is in essence the celebrated Kolmogorov-Arnold-Moser (KAM) theorem. For larger k more and more curves are broken, but some still exist [see Fig. 7(a) where $k=0.80$ and Figs. 7(b) where $k=1.2$]. At a critical value of $k_c \approx 0.97$, curves that span $0 < \theta' \leq 2\pi$ no longer exist. The critical value k_c was calculated by Greene (1979) [see also Falcolini and de la Llave (1992)] to many decimal places.

The question of when the last continuous curve breaks is an important one in Hamiltonian dynamics
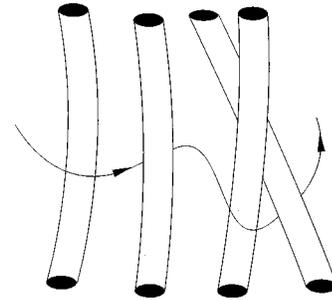
theory. In particular, it is of importance in the passive advection fluid mechanics problem since these curves are barriers to transport. One is interested in when these curves break as the sizes and positions of the resonances change. The method developed by Greene gives a precise answer to this question, but requires some effort. A simple but rough criterion that yields an estimate for when the continuous curves between two resonances cease to persist is given by the *Chirikov overlap criterion*. According to this criterion the last curve separating two resonances will be destroyed when the sum of the half-widths of the two resonances (calculated independently) equals the distance between the resonances; that is,

$$\frac{W_1}{2} + \frac{W_2}{2} = |J'_1 - J'_2|, \qquad (27)$$

where W_1 and W_2 denote the widths of the resonances while J'_1 and J'_2 denote their positions. This criterion is straightforward to apply and usually gives reasonable results. However, it must be borne in mind that it is only a rough estimate and as such has limitations. As noted above, more sophisticated criteria exist.

The study of two-degree-of-freedom Hamiltonian systems is a richly developed yet still open area of research. Unfortunately, in a single lecture it is only possible to scratch the surface and, it is hoped, whet your appetite. Conspicuously absent from this lecture is any discussion of the notions of universality and renormalization (see, for example, MacKay, 1982 or del-Castillo-Negrete *et al.*, 1997). There is much to be learned from the references we have cited.

E. ''Diffusion'': Three degrees of freedom

In closing we mention something about three-degree-of-freedom systems. For these systems the invariant sets that are remnants of the integrable N-tori do not divide the phase space. For three-degree-of-freedom systems the phase space is six dimensional and the corresponding three-dimensional invariant tori do not isolate regions. Because of this, trajectories are not confined and can wander around the tori. This phenomenon is generally called *Arnold diffusion*. A cartoon of this is shown in Fig. 10.

---

[7]To do this one can use standard Runge-Kutta packages. However, more sophisticated *symplectic* integration algorithms exist. See, for example, Sanz-Serna and Calvo (1994), Kueny (1993), and Kueny and Morrison (1995).

There is a great deal of literature dealing with the chaotic advection of a passive tracer in two-dimensional fluid systems. These studies typically involve model stream functions that are time periodic and hence are nonintegrable. For these systems the diffusion phenomenon mentioned above cannot occur. However, it is possible that the solution of Eq. (1) is not periodic, but quasiperiodic, a special case of which is represented by the following:

$$\psi(\theta,J,t) = -f(\theta,J,\omega_2 t,\omega_3 t), \tag{28}$$

where f is a function that satisfies

$$f(\theta,J,\omega_2 t,\omega_3 t) = f(\theta,J,\omega_2 t + 2\pi,\omega_3 t)$$
$$= f(\theta,J,\omega_2 t,\omega_3 t + 2\pi). \tag{29}$$

If ω_2/ω_3 is irrational, then ψ is not periodic.

One can puff up a system with a Hamiltonian of the form of Eq. (28) into a three-degree-of-freedom system by a technique similar to that described above. Let $\theta =: \theta_1$, $J =: J_1$, define

$$H(\theta_1,J_1,\theta_2,J_2,\theta_3,J_3) = f(\theta_1,J_1,\theta_2,\theta_3) + \omega_2 J_2$$
$$+ \omega_3 J_3, \tag{30}$$

and introduce the fake time s as before. Note that the last two terms of Eq. (30) are just the Hamiltonian for two linear oscillators in action-angle form, but here they are coupled to each other and to oscillator "1" through f. Hamilton's equations are

$$\frac{d\theta_1}{ds} = \frac{\partial f}{\partial J_1}, \quad \frac{d\theta_2}{ds} = \omega_2, \quad \frac{d\theta_3}{ds} = \omega_3, \tag{31}$$

$$\frac{dJ_1}{ds} = -\frac{\partial f}{\partial \theta_1}, \quad \frac{dJ_2}{ds} = -\frac{\partial f}{\partial \theta_2}, \quad \frac{dJ_3}{ds} = -\frac{\partial f}{\partial \theta_3}. \tag{32}$$

It is clear how the last two equations of (31) can be integrated and thus the system can be collapsed back down. The last two equations of (32) guarantee that J_2 and J_3 will vary so as to make H conserved.

The kind of quasiperiodic system treated in this subsection is undoubtedly relevant for the study of transport in two-dimensional fluids. Solutions of Eq. (1) are likely to be closer to quasiperiodic than periodic. A stream function that describes an azimuthally symmetric shear flow plus three waves with different speeds is quasiperiodic. Transport in such systems and their generalization to more frequencies is not well understood. (See, for example, Wiggins, 1992.)

III. FUNCTIONAL CALCULUS, TWO ACTION PRINCIPLES OF MECHANICS, AND THE ACTION PRINCIPLE AND CANONICAL HAMILTONIAN DESCRIPTION OF THE IDEAL FLUID

This lecture is devoted to developing techniques that are needed to describe infinite-dimensional Hamiltonian systems, and then to using these techniques to describe the canonical Hamiltonian description of the ideal fluid in terms of so-called Lagrangian or material variables. Specifically, in Sec. III.A techniques of functional calcu-

lus are presented, in Sec. III.B two traditional action principles of classical mechanics are reviewed along with their connection to Hamilton's equations, and in Sec. III.C the action principle and Hamiltonian description of the fluid are treated in detail.

A. Functional calculus

A *functional* is a map that takes functions into real numbers. Describing them correctly requires defining a function space, which is the domain of the functional, and the rule that assigns the real number. Like ordinary functions, functionals have notions of continuity, differentiability, the chain rule, etc. In this subsection we shall not be concerned with rigor, but with learning how to perform various formal manipulations.[8]

As an example of a functional consider the kinetic energy of a constant-density, one-dimensional, bounded fluid:

$$T[u] = \frac{1}{2} \int_{x_0}^{x_1} \rho_0 u^2 dx. \tag{33}$$

Here T is a functional of u which is indicated by the [ ] notation, a notation that we use in general to denote functionals. The function $u(x)$ is the fluid velocity, which is defined on $x \in [x_0,x_1]$, and ρ_0 is a constant fluid density. Given a function $u(x)$ we could put it into Eq. (33), do the integral, and get a number.

We would like to know in general how the value of a functional $K[u]$ changes as $u(x)$ changes a little, say $u(x) \to u(x) + \epsilon \delta u(x)$, where $u + \epsilon \delta u$ must still be in our domain. The first-order change in K induced by δu is called the *first variation*, δK, and is given by

$$\delta K[u;\delta u] := \lim_{\epsilon \to 0} \frac{K[u + \epsilon \delta u] - K[u]}{\epsilon}$$

$$= \frac{d}{d\epsilon} K[u + \epsilon \delta u] \bigg|_{\epsilon = 0}$$

$$=: \int_{x_0}^{x_1} \delta u \frac{\delta K}{\delta u(x)} dx =: \left\langle \frac{\delta K}{\delta u}, \delta u \right\rangle. \tag{34}$$

We shall assume that the limit exists and that there are no problems with the equalities above; later, however, we shall give an exercise in which something "interesting" happens.

The notation $\delta K[u;\delta u]$ is used because there is a difference in the behavior of the two arguments: generally δK is a linear functional in δu, but not so in u. The quantity $\delta K/\delta u(x)$ of Eq. (34) is the *functional derivative* of the functional K. This notation for the functional derivative is chosen since it emphasizes the fact that

---

[8]For further details, on a level roughly consistent with that given here, see Courant and Hilbert (1953), Chapter IV, and Gelfand and Fomin (1963).

$\delta K/\delta u$ is a gradient in function space. The reason why the arguments of u are sometimes displayed will become clear below.

For the example of Eq. (33) the first variation is given by

$$\delta T[u;\delta u]=\int_{x_0}^{x_1}\rho_0 u\,\delta u\,dx, \tag{35}$$

and hence the functional derivative is given by

$$\frac{\delta T}{\delta u}=\rho_0 u. \tag{36}$$

To see that the functional derivative is a gradient, let us take a sidetrack and consider the first variation of a function of n variables, $f(x_1,x_2,\ldots,x_n)=f(x)$:

$$\delta f(x;\delta x)=\sum_{i=1}^{n}\frac{\partial f(x)}{\partial x_i}\delta x_i=:\nabla f\cdot\delta x. \tag{37}$$

It is interesting to compare the definition of Eq. (37) with the last definition of Eq. (34). The $\cdot$ in Eq. (37) is analogous to the pairing $\langle\ ,\ \rangle$, while δx is analogous to δu. In fact, the index i is analogous to x, the argument of u. Finally, the gradient ∇f is analogous to $\delta K/\delta u$.

Consider now a more general functional, one of the form

$$F[u]=\int_{x_0}^{x_1}\mathcal{F}(x,u,u_x,u_{xx},\ldots)dx, \tag{38}$$

where $\mathcal{F}$ is an ordinary, sufficiently differentiable, function of its arguments. Note $u_x:=du/dx$, etc. The first variation of Eq. (38) yields

$$\delta F[u;\delta u]=\int_{x_0}^{x_1}\left(\frac{\partial\mathcal{F}}{\partial u}\delta u+\frac{\partial\mathcal{F}}{\partial u_x}\delta u_x\right.$$
$$\left.+\frac{\partial\mathcal{F}}{\partial u_{xx}}\delta u_{xx}+\cdots\right)dx, \tag{39}$$

which upon integration by parts becomes

$$\delta F[u;\delta u]=\int_{x_0}^{x_1}\delta u\left(\frac{\partial\mathcal{F}}{\partial u}-\frac{d}{dx}\frac{\partial\mathcal{F}}{\partial u_x}\right.$$
$$\left.+\frac{d^2}{dx^2}\frac{\partial\mathcal{F}}{\partial u_{xx}}-\cdots\right)dx$$
$$+\left.\left(\frac{\partial\mathcal{F}}{\partial u_x}\delta u+\cdots\right)\right|_{x_0}^{x_1}. \tag{40}$$

Usually the variations δu are chosen so that the last term, the boundary term, vanishes; e.g., $\delta u(x_0)=\delta u(x_1)=0$, $\delta u_x(x_0)=\delta u_x(x_1)=0$, etc. Sometimes the boundary term vanishes without a condition on δu because of the form of $\mathcal{F}$. When this happens the boundary conditions are called *natural*. Assuming, for one reason or the other, that the boundary term vanishes, Eq. (40) becomes

$$\delta F[u;\delta u]=\left\langle\frac{\delta F}{\delta u},\delta u\right\rangle, \tag{41}$$

where

$$\frac{\delta F}{\delta u}=\frac{\partial\mathcal{F}}{\partial u}-\frac{d}{dx}\frac{\partial\mathcal{F}}{\partial u_x}+\frac{d^2}{dx^2}\frac{\partial\mathcal{F}}{\partial u_{xx}}-\cdots. \tag{42}$$

The main objective of the calculus of variations is the extremization of functionals. A common terminology is to call a function $\hat{u}$, which is a point in the domain, an *extremal point* if $\delta F[u]/\delta u|_{u=\hat{u}}=0$. It could be a maximum, a minimum, or an inflection point. If the extremal point $\hat{u}$ is a minimum or maximum, then such a point is called an *extremum*.

The standard example of a functional that depends on the derivative of a function is the arc-length functional,

$$L[u]=\int_{x_0}^{x_1}\sqrt{1+u_x^2}\ dx. \tag{43}$$

We leave it to you to show that the shortest distance between two points is a straight line.

Another example is the functional defined by evaluating the function u at the point x'. This can be written as

$$u(x')=\int_{x_0}^{x_1}\delta(x-x')u(x)dx, \tag{44}$$

where $\delta(x-x')$ is the Dirac delta function and where we have departed from the [ ] notation. Applying the definition of Eq. (34) yields

$$\frac{\delta u(x')}{\delta u(x)}=\delta(x-x'). \tag{45}$$

This is the infinite-dimensional or continuum analog of $\partial x_i/\partial x_j=\delta_{ij}$, where δ_{ij} is the Kronecker delta function. Equation (45) shows why it is sometimes useful to display the argument of the function in the functional derivative.

The generalizations of the above ideas to functionals of more than one function and to more than a single spatial variable are straightforward. An example is given by the kinetic energy of a three-dimensional compressible fluid,

$$T[\rho,v]=\tfrac{1}{2}\int_D\rho v^2\ d^3x, \tag{46}$$

where the velocity has three rectangular components $v=(v_1,v_2,v_3)$ that depend upon $x=(x_1,x_2,x_3)\in D$ and $v^2=v\cdot v=v_1^2+v_2^2+v_3^2$. The functional derivatives are

$$\frac{\delta T}{\delta v_i}=\rho v_i,\qquad\frac{\delta T}{\delta\rho}=\frac{v^2}{2}. \tag{47}$$

We shall use these later.

For a more general functional $F[\psi]$, where $\psi(x)=(\psi^1,\psi^2,\ldots,\psi^\nu)$ and $x=(x_1,x_2,\ldots,x_n)$, the analog of Eq. (34) is

$$\delta F[\psi;\delta\psi]=\int_D\delta\psi_i\frac{\delta F}{\delta\psi_i(x)}\,d^nx=:\left\langle\frac{\delta F}{\delta\psi},\delta\psi\right\rangle. \tag{48}$$

As an exercise consider the pathological functional

$$P[\psi]=\int_{-1}^{1}\mathcal{P}(\psi_1,\psi_2)\ dx, \tag{49}$$

where

$$
\mathcal{P}=\begin{cases} \dfrac{\psi_1\psi_2^2}{\psi_1^2+\psi_2^2} & \text{if } \psi_{1,2}\neq 0, \\[2mm] 0 & \text{if } \psi_{1,2}=0. \end{cases} \tag{50}
$$

Calculate $\delta P[0,0;\delta\psi_1,\delta\psi_2]$. Part of this problem is to figure out what the problem is.

Next, we consider the important *functional chain rule*, which is a simple idea that underlies a great deal of literature relating to the Hamiltonian structure of fluids and plasmas.

Suppose we have a functional $F[u]$ and we know u is related to another function w by means of a linear operator

$$
u=\mathcal{O}w. \tag{51}
$$

As an example, u and w could be real-valued functions of a single variable x, and

$$
\mathcal{O}:=\sum_{k=0}^{n} a_k(x)\frac{d^k}{dx^k}, \tag{52}
$$

where, as usual, u, w, and a_k have as many derivatives as needed. We can define a functional of w by inserting Eq. (51) into $F[u]$:

$$
\bar{F}[w]:=F[u]=F[\mathcal{O}w]. \tag{53}
$$

Equating variations yields

$$
\left\langle \frac{\delta\bar{F}}{\delta w},\delta w \right\rangle = \left\langle \frac{\delta F}{\delta u},\delta u \right\rangle, \tag{54}
$$

where the equality makes sense if δu and δw are connected by Eq. (51), i.e.,

$$
\delta u=\mathcal{O}\delta w, \tag{55}
$$

where we assume an arbitrary δw induces a δu.

Inserting Eq. (55) into Eq. (54) yields

$$
\left\langle \frac{\delta\bar{F}}{\delta w},\delta w \right\rangle = \left\langle \frac{\delta F}{\delta u},\mathcal{O}\delta w \right\rangle =:\left\langle \mathcal{O}^\dagger\frac{\delta F}{\delta u},\delta w \right\rangle, \tag{56}
$$

where $\mathcal{O}^\dagger$ is the formal adjoint of $\mathcal{O}$. The DuBois-Reymond lemma of the calculus of variations states that if δw is arbitrary, then Eq. (56) implies

$$
\frac{\delta\bar{F}}{\delta w}=\mathcal{O}^\dagger\frac{\delta F}{\delta u}. \tag{57}
$$

This lemma is proven by assuming that Eq. (57) does not hold at some point x, selecting δw to be localized about the point x, and establishing a contradiction. A physicist would just set δw equal to the Dirac delta function to obtain the result.

Notice that nowhere did we assume that $\mathcal{O}$ was invertible—it need not be for the chain rule to work in one direction. This is because, in the sense displayed above, functionals transform in the other direction. Clearly the transformation of Eq. (51) is a special case, in that the two functions u and w are linearly related. However, if u depends nonlinearly upon w we can still obtain a relation of the form of Eq. (57). We shall demonstrate this in the more general case for a functional $F[\psi]$, where ψ is related to $\chi=(\chi^1,\chi^2,\ldots,\chi^\mu)$ in an arbitrary, possibly nonlinear and noninvertible, way:

$$
\psi^i=\psi^i[\chi], \quad i=1,2,\ldots,\nu. \tag{58}
$$

This [ ] notation could be confusing, since we have used it to denote functionals, but since we have already stated that ψ and χ are functions there should be no confusion. A variation of ψ induced by χ requires linearization of Eq. (58), which we write as

$$
\delta\psi^i=\frac{\delta\psi^i}{\delta\chi}[\chi;\delta\chi], \tag{59}
$$

or simply, since $\delta\psi/\delta\chi$ is a linear operator on $\delta\chi$, as

$$
\delta\psi^i=\frac{\delta\psi^i}{\delta\chi^j}\delta\chi^j, \tag{60}
$$

where $j=1,2,\ldots,\mu$. Inserting Eq. (60) into Eq. (54) implies

$$
\left\langle \frac{\delta F}{\delta\chi},\delta\chi \right\rangle = \left\langle \left(\frac{\delta\psi}{\delta\chi}\right)^\dagger\frac{\delta F}{\delta\psi},\delta\chi \right\rangle, \tag{61}
$$

whence it is seen that

$$
\frac{\delta F}{\delta\chi^j}=\left(\frac{\delta\psi^i}{\delta\chi^j}\right)^\dagger\frac{\delta F}{\delta\psi^i}. \tag{62}
$$

Here we have dropped the overbar on F, as is commonly done. In Eq. (62) it is important to remember that $\delta(\text{function})/\delta(\text{function})$ is a linear operator acting to its right, as opposed to $\delta(\text{functional})/\delta(\text{function})$, which is a gradient.

As an example consider functionals that depend upon the two components of the velocity field for an incompressible fluid in two dimensions, $u(x,y)$ and $v(x,y)$. These are linearly related to the stream function ψ by $u=-\partial\psi/\partial y$ and $v=\partial\psi/\partial x$. For this case Eq. (59) becomes

$$
\delta u=\frac{\delta u}{\delta\psi}\,\delta\psi=-\frac{\partial}{\partial y}\,\delta\psi,
$$

$$
\delta v=\frac{\delta v}{\delta\psi}\,\delta\psi=\frac{\partial}{\partial x}\,\delta\psi, \tag{63}
$$

and

$$
\frac{\delta F}{\delta\psi}=\frac{\partial}{\partial y}\frac{\delta F}{\delta u}-\frac{\partial}{\partial x}\frac{\delta F}{\delta v}. \tag{64}
$$

Now consider the *second variation*, $\delta^2 F$, and *second functional derivative*, $\delta^2 F/\delta\psi\delta\psi$. Since the first varia-

tion, $\delta F[\psi;\delta\psi]$, is a functional of ψ, a second variation can be made in this argument:

$$\delta^2 F[\psi;\delta\psi,\delta\hat{\psi}] = \frac{d}{d\eta}\delta F[\psi + \eta\delta\hat{\psi};\delta\psi]\Big|_{\eta=0}$$

$$=: \int_D \delta\psi^i \frac{\delta^2 F}{\delta\psi^i\delta\psi^j}\delta\hat{\psi}^j\, d^n x$$

$$=: \left\langle \delta\psi, \frac{\delta^2 F}{\delta\psi\delta\psi}\delta\hat{\psi} \right\rangle. \tag{65}$$

Observe that $\delta^2 F$ is a bilinear functional in $\delta\psi$ and $\delta\hat{\psi}$. If we set $\delta\hat{\psi}=\delta\psi$ we obtain a quadratic functional. Equation (65) defines $\delta^2 F/\delta\psi\delta\psi$, which is an operator that acts linearly on $\delta\hat{\psi}$ but depends nonlinearly on ψ. This operator possesses a symmetry analogous to the interchange of the order of second partial differentiation. To see this observe

$$\delta^2 F[\psi;\delta\psi,\delta\hat{\psi}] = \frac{\partial^2}{\partial\eta\partial\epsilon}F[\psi + \eta\delta\hat{\psi} + \epsilon\delta\psi]\Big|_{\epsilon=0,\eta=0}. \tag{66}$$

Since the order of differentiation in Eq. (66) is immaterial it follows that

$$\left(\frac{\delta^2 F}{\delta\psi^i\delta\psi^j}\right)^\dagger = \frac{\delta^2 F}{\delta\psi^j\delta\psi^i}. \tag{67}$$

This relation is necessary for establishing the Jacobi identity of noncanonical Poisson brackets.

As an example consider the second variation of the arc-length functional of Eq. (43). Performing the operations of Eq. (66) yields

$$\delta^2 L[u;\delta u,\delta\hat{u}] = \int_{x_0}^{x_1}\delta u_x\frac{1}{(1+u_x^2)^{3/2}}\,\delta\hat{u}_x dx. \tag{68}$$

Thus

$$\frac{\delta^2 L}{\delta u^2} = \frac{d}{dx}\frac{-1}{(1+u_x^2)^{3/2}}\frac{d}{dx}. \tag{69}$$

For an important class of function spaces, one can convert functionals into functions of a countably infinite number of arguments. This is a method for proving theorems concerning functionals and can also be useful for establishing formal identities. One way to do this would be to convert the integration of a functional into a sum by finite differencing. Another way to do this, for example for functionals of the form of Eq. (38), is to suppose $[x_0,x_1]=[-\pi,\pi]$ and expand in a Fourier series,

$$u(x) = \sum_{k=-\infty}^{\infty} u_k\, e^{ikx}. \tag{70}$$

Upon inserting Eq. (70) into Eq. (38) one obtains an expression for the integrand, which is, in principle, a known function of x. Integration then yields a function of the Fourier amplitudes, u_k. Thus we obtain

$$F[u] = F(u_0,u_1,u_{-1},\dots). \tag{71}$$

In closing this discussion of functional calculus we consider a functional, one expressed as a function of an infinite number of arguments, that demonstrates an "interesting" property. The functional is given by

$$F(x_1,x_2,\dots) = \sum_{k=1}^{\infty}(\tfrac{1}{2}a_k x_k^2 - \tfrac{1}{4}b_k x_k^4), \tag{72}$$

where the domain of F is composed of sequences $\{x_k\}$, and the coefficients are given by

$$a_k = \frac{1}{k^6},\quad b_k = \frac{1}{k^2}. \tag{73}$$

Assuming that Eq. (72) converges uniformly, the first variation yields

$$\delta F = \sum_{k=1}^{\infty}(a_k x_k - b_k x_k^3)\delta x_k, \tag{74}$$

which has three extremal points,

$$x_k^{(0)} = 0,\quad x_k^{(\pm)} = \pm(a_k/b_k)^{1/2}, \tag{75}$$

for all k. It is the first of these that will concern us. The second variation evaluated at $x_k^{(0)}$ is

$$\delta^2 F = \sum_{k=1}^{\infty}a_k(\delta x_k)^2, \tag{76}$$

where we assume Eq. (74) converges uniformly for x_k and δx_k. Since $a_k > 0$ for all k, Eq. (76) is positive definite; i.e.,

$$\delta^2 F > 0\quad \text{for}\quad \delta x_k \neq 0,\quad \text{for all } k. \tag{77}$$

However, consider ΔF defined by

$$\Delta F = F(x^{(0)} + \Delta x) - F(x^{(0)})$$

$$= \sum_{k=1}^{\infty}[\tfrac{1}{2}a_k(\Delta x_k)^2 - \tfrac{1}{4}b_k(\Delta x_k)^4], \tag{78}$$

which we evaluate at

$$\Delta x_k = \begin{cases} \dfrac{1}{m}, & k=m, \\[2mm] 0, & k\neq m, \end{cases} \tag{79}$$

and obtain

$$\Delta F < 0, \tag{80}$$

provided $m > 1$. Since m can be made as large as desired, we have shown that inside any neighborhood of $x^{(0)}$, no matter how small, $\Delta F < 0$. Therefore this extremal point is not a minimum — even though $\delta^2 F$ is positive definite.

A sufficient condition for proving that an extremal point is an extremum is afforded by a property known as *strong positivity*. If $\hat{\psi}$ is an extremal point and the quadratic functional $\delta^2 F[\hat{\psi};\delta\psi]$ satisfies

$$\delta^2 F[\hat{\psi};\delta\psi] \geq c\|\delta\psi\|^2,$$

where $c = \text{const} > 0$ and $\| \ \|$ is a norm defined on the domain of F, then $\delta^2 F[\hat{\psi}; \delta\psi]$ is strongly positive. This is sufficient for $\hat{\psi}$ to be a minimum. We leave it to you to explain why the functional $F(x_1, x_2, \dots)$ is not strongly positive. This example points to a mathematical technicality that is encountered when proving stability by Liapunov's method (see Sec. VI).

B. Two action principles of mechanics

Physicists have had a long-lasting love affair with the idea of generating physical laws by setting the derivative of some functional to zero. This is called an *action principle*. The most famous action principle is Hamilton's principle, which produces Lagrange's equations of mechanics upon variation. One reason action principles are appreciated is that they give a readily covariant theory, and means have been developed for building in symmetries. However, it should be pointed out that the use of continuous symmetry groups in this context is only a limited part of a deep and beautiful theory that was initiated by Sophus Lie and others. Perhaps the most convincing deep reason for the use of action principles is the cleanliness and utility of Feynman's path-integral formulation. The utility of action principles should not be understated. Indeed, they provide a good starting place for making approximations. However, a quote from Truesdell (1966) cannot be resisted:

"A fully conservative situation can be described by an action principle, which has the advantage of making the theory accessible also to physicists."

In any event, Hamilton's principle is an important prototype upon which modern theories are in part built. Shortly, we shall show how this story goes for the ideal fluid, but first we review some mechanics. [See Saletan and Cromer (1971) or Sudarshan and Mukunda (1974) for standard presentations.]

One approach to producing the equations of motion for a mechanics problem is first to identify the configuration space Q with coordinates $q = (q_1, q_2, \dots, q_N)$. Then, based on physical intuition, write down the kinetic and potential energies, T and V, respectively. The equations of motion then follow upon setting the functional derivative of the following action functional to zero:

$$S[q] = \int_{t_0}^{t_1} L(q, \dot{q}, t) \, dt, \tag{81}$$

where $L := T - V$ is the Lagrangian function. The functions $q(t)$ over which we are extremizing must satisfy the fixed end conditions $q(t_0) = q_0$ and $q(t_1) = q_1$. Thus $\delta q(t_0) = \delta q(t_1) = 0$. The functional derivative relations

$$\frac{\delta S[q]}{\delta q^i} = 0 \tag{82}$$

imply Lagrange's equations,

$$\frac{\partial L}{\partial q^i} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i}. \tag{83}$$

This is Hamilton's principle.

Since for particles in rectangular coordinates[9] usually

$$T = \tfrac{1}{2} \sum_{i=1}^{N} m_i (\dot{q}^i)^2, \quad V = V(q), \tag{84}$$

Eqs. (83) yield

$$m_i \ddot{q}_i = -\frac{\partial V}{\partial q^i}. \tag{85}$$

This is just Newton's second law with a conservative force. You will notice that Hamilton's principle does not yield Hamilton's equations—one way to get them is via the Legendre transformation.

The *Legendre transformation* is a trick for transferring functional dependence. Generally it is used in physics when one has a sort of "fundamental" function that describes a theory, whether it be a thermodynamic potential or, as is the case here, a Lagrangian. The Legendre transformation has a nice geometric interpretation, but we shall skip this. Here we shall use it to transform the N second-order differential equations of (85) into the $2N$ first-order equations of Hamilton.

Define a quantity $p_i := \partial L / \partial \dot{q}^i$, which is the canonical momentum, and consider

$$\bar{H}(q, p, \dot{q}, t) := p_i \dot{q}^i - L(q, \dot{q}, t). \tag{86}$$

Now we ask the question: How does $\bar{H}$ change if we independently change q, $\dot{q}$, p, and t a little? Evidently

$$\delta \bar{H} = \frac{\partial \bar{H}}{\partial q^i} \delta q^i + \frac{\partial \bar{H}}{\partial \dot{q}^i} \delta \dot{q}^i + \frac{\partial \bar{H}}{\partial p_i} \delta p_i + \frac{\partial \bar{H}}{\partial t} \delta t$$

$$= -\frac{\partial L}{\partial q^i} \delta q^i + \left(p_i - \frac{\partial L}{\partial \dot{q}^i} \right) \delta \dot{q}^i + \dot{q}^i \delta p_i$$

$$- \frac{\partial L}{\partial t} \delta t. \tag{87}$$

The first thing to notice is that if $\delta q = \delta p = \delta t = 0$, i.e., we only vary $\delta \dot{q}$, then $\delta \bar{H} = 0$, since $p_i = \partial L / \partial \dot{q}^i$. This means $\bar{H}$ is independent of $\dot{q}$, so we drop the overbar and write $H(q, p, t)$. Equating the remaining coefficients of the variations yields

$$\frac{\partial H}{\partial q^i} = -\frac{\partial L}{\partial q^i}, \quad \frac{\partial H}{\partial p_i} = \dot{q}^i, \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{88}$$

Lagrange's equations, (83), together with the definition of p_i and the middle equation of (88), give Hamilton's equations:

$$\dot{p}_i = -\frac{\partial H}{\partial q^i}, \quad \dot{q}^i = \frac{\partial H}{\partial p_i}. \tag{89}$$

---

[9]More precisely the kinetic energy contains $\eta_{ij} \dot{q}^i \dot{q}^j$, where the metric $\eta_{ij} := \delta_{ij}$ and its inverse can be used to raise and lower indices.
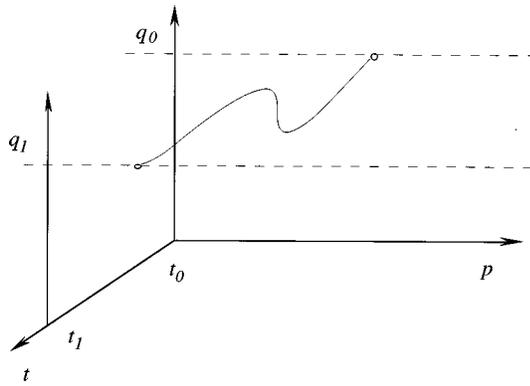
FIG. 11. Clothesline boundary conditions for the phase-space action principle.

In order to calculate $H(p,q,t)$ explicitly one uses $p_i = \partial L / \partial \dot{q}^i$ to solve for $\dot{q} = \dot{q}(p)$ and then inserts this into Eq. (86). This requires L to be convex in $\dot{q}$. Since there exist important physical cases in which L is not convex, Dirac and others developed a theory to handle this. An interesting application of Dirac's constraint theory for filtering out fast motion in geophysical fluid dynamical models has been developed by Salmon (1988a).

Now consider another action principle, which is sometimes called the *phase-space action.* This one, which directly yields Hamilton's equations, is given by

$$S[q,p] = \int_{t_0}^{t_1} [p_i \dot{q}^i - H(q,p,t)] \ dt, \qquad (90)$$

where S is a functional of q and p, independently. The end conditions are $q(t_0) = q_0$ and $q(t_1) = q_1$, i.e., q is fixed as before. However, the boundary condition on p is natural in that nothing is required of it at the ends. One has a sort of "clothesline" boundary condition as depicted in Fig. 11, where the curve is free to slide along the lines of constant q in the p direction.

Variation of S with respect to q and p yields, respectively,

$$\dot{p}_i = -\frac{\partial H}{\partial q^i}, \quad \dot{q}^i = \frac{\partial H}{\partial p_i}. \qquad (91)$$

Thus the phase-space action directly yields Hamilton's equations as the extremal condition.

C. Action principle and canonical Hamiltonian description of the ideal fluid in Lagrangian or material variables

Now we are in a position to talk about the action for fluid mechanics,[10] but we are going to do so in terms of

---

[10]An action principle for the ideal fluid dates back to Lagrange (1788), although of course he was unaware of the thermodynamics that we shall include. More modern references with generalizations include Serrin (1959), Eckart (1960), Newcomb (1962), and Salmon (1982, 1988b).

variables that might be new to you. Often, fluid mechanics is taught entirely in terms of Eulerian variables. In what follows, *Lagrangian variables*, or as they are sometimes called, material variables, will be central.

The idea we are going to pursue is a simple one. *If a fluid is described as a collection of fluid particles or elements, then both the Hamiltonian and the Lagrangian formalism that we have described above can be adapted to describe the ideal fluid.* The adaptation requires an extension to an infinite number of degrees of freedom in order to describe a continuum of fluid elements. This means that a fluid element is shrunk to zero size and that there is one for each point of the fluid. This is an idealization, since in reality fluid elements do not exist: if they were of macroscopic size, they would not maintain their integrity forever, and if they were of microscopic size, we would be outside the realm of fluid mechanics. However, there exists a precise Eulerian state corresponding to a Lagrangian state. It should be kept in mind that the above limitations apply to the fluid description in general, whether it be in Lagrangian or Eulerian variables.

Suppose the position of a fluid element, referred to a fixed rectangular coordinate system, is given by

$$q = q(a,t), \qquad (92)$$

where $q = (q_1, q_2, q_3)$. This is the material or Lagrangian variable. Here $a = (a_1, a_2, a_3)$ could be any label[11] that identifies a fluid particle, but below it will be taken to be the position of the fluid particle at time $t = 0$ in rectangular coordinates. The quantities $q^i(a,t)$ are coordinates for the configuration space $\mathcal{Q}$, which is in fact a function space because in addition to the three indices i there is the continuum label a. We assume that a varies over a fixed domain D, which is completely filled with fluid, and that the functions q map D onto itself. We shall assume that as many derivatives of q with respect to a as needed exist, but we shall not say more about $\mathcal{Q}$; in fact, not that much is known about the solution space for the 3D fluid equations in Lagrangian variables. At this stage we shall assume that the configuration space has been specified and proceed to discuss the potential energy of the fluid.

The fluid approximation assumes local thermodynamic equilibrium in spite of the fact that fluid motion is in general not quasistatic. Potential energy is stored in terms of pressure and temperature. More precisely, we adapt the energy representation of thermodynamics in which the extensive energy is treated as a function of the extensive variables, viz., the entropy and the volume. For a fluid it is convenient to consider the energy per unit mass, which we denote by U to be a function of the entropy per unit mass, s, and the mass density, ρ. The inverse of the latter quantity is a measure of the volume. The intensive quantities, pressure and temperature, are obtained as follows:

---

[11]Note that the freedom to relabel particles is associated with the Casimir invariants, which are discussed below. See Calkin (1963), Newcomb (1967), Bretherton (1970), Ripa (1981), Salmon (1982), and Padhye and Morrison (1996a, 1996b).

$$T = \frac{\partial U}{\partial s}(s,\rho), \quad p = \rho^2 \frac{\partial U}{\partial \rho}(s,\rho). \tag{93}$$

The second of equation (93) is a bit peculiar—it arises because the volume, the usual thermodynamic variable, is proportional to ρ^{-1}. Special choices for U produce specific thermodynamics for fluid flow.[12]

The quantities ρ and s are in fact Eulerian variables, which we must, in order to move ahead, describe in terms of Lagrangian variables. With this goal in mind, let us diverge for a moment and discuss the Lagrangian-Eulerian map. The difference between the two types of variables can be elucidated by describing two ways of watching fish. In the Eulerian picture one stays at a point and watches whatever fish happen by; in the Lagrangian picture one picks out a particular fish and keeps track of where it goes. Note that this analogy gets better if the fish are very small, neutrally buoyant, and dead.

Call r the *spatial* variable, i.e., the Eulerian point of observation. The Eulerian density is then related to the Lagrangian variable q as

$$\rho(r,t) = \int_D \delta(r - q(a,t))\rho_0(a)\, d^3a. \tag{94}$$

Here $\delta(r-q)$ is a three-dimensional Dirac delta function and $\rho_0(a)$ is an initial configuration of mass density ascribed to the particle labeled by a. It is akin to knowing the mass of the particle labeled by i in conventional particle mechanics.

Equation (94) embodies mass conservation. This can be seen by using a property of the δ function, $\delta(f(x)) = \delta(x - x_0)/|f'(x_0)|$, where x_0 is the only place where $f(x_0) = 0$. In three dimensions this yields

$$\rho(r,t) = \frac{\rho_0(a)}{\mathcal{J}(a,t)}\bigg|_{a = q^{-1}(r,t)}, \tag{95}$$

where the Jacobian $\mathcal{J} = \det(\partial q^i/\partial a^j)$. That this is local mass conservation follows from

$$\rho\, d^3q = \rho_0\, d^3a, \tag{96}$$

where d^3a is an initial volume element that maps into d^3q at time t, and $d^3q = \mathcal{J}\, d^3a$. (When integrating over D we shall replace d^3q by d^3r.)

In addition to the mass ascribed to a fluid particle, one could ascribe other quantities, e.g., color, smell, or what have you. In the ideal fluid, the entropy per unit mass s is such a quantity. We suppose that initially $s = s_0(a)$ and that it remains so. A form similar to Eq. (94) corresponding to this statement is

$$\sigma(r,t) = \int_D \sigma_0(a)\delta(r - q(a,t))d^3a, \tag{97}$$

where $\sigma(r,t) = \rho(r,t)s(r,t)$ is the entropy per unit volume and $\sigma_0 = \rho_0(a)s_0(a)$. Thus the counterpart of Eq. (95) is

$$s(r,t) = s_0(a)\big|_{a = q^{-1}(r,t)}. \tag{98}$$

This is merely the statement that the quantity s stays put on a fluid particle.

Completing the Lagrange-Euler map requires the specification of the Eulerian velocity field, something that is not needed now, but that we record here for later reference. By now you will have noticed that the Euler-Lagrange map naturally takes the Lagrangian variables into Eulerian densities. Thus we consider the momentum density $M := \rho v$. A form for M similar to Eqs. (94) and (97) is

$$M(r,t) = \int_D \dot{q}(a,t)\delta(r - q(a,t))\rho_0(a)d^3a, \tag{99}$$

where the $\cdot$ notation now means differentiation with respect to time at fixed label a. Performing the integration produces the counterpart of Eqs. (95) and (96), viz.,

$$v(r,t) = \dot{q}(a,t)\big|_{a = q^{-1}(r,t)}, \tag{100}$$

which is the usual relation between the Lagrangian variable and the Eulerian velocity field.

Now we can return to our quest for the potential energy. Since the energy per unit volume is given by ρU, the total potential-energy function is evidently

$$V[q] = \int_D \rho_0 U(s_0, \rho_0/\mathcal{J})d^3a. \tag{101}$$

Observe that Eq. (101) is a functional of q that depends only upon $\mathcal{J}$ and hence only upon $\partial q/\partial a$.

The next step required for constructing Hamilton's principle is to obtain an expression for the kinetic-energy functional. This is clearly given by

$$T[q] = \tfrac{1}{2}\int_D \rho_0 \dot{q}^2 d^3a, \tag{102}$$

where we use the shorthand $\dot{q}^2 := \eta_{ij}\dot{q}^i\dot{q}^j$. Observe that Eq. (102) is a functional of q that depends only upon $\dot{q}$.

From Eqs. (101) and (102) the Lagrangian functional is obtained,

$$L[q,\dot{q}] = \int_D [\tfrac{1}{2}\rho_0 \dot{q}^2 - \rho_0 U(s_0, \rho_0/\mathcal{J})]d^3a$$

$$=: \int_D \mathcal{L}(q,\dot{q},\partial q/\partial a,t)d^3a, \tag{103}$$

where $\mathcal{L}(q,\dot{q},\partial q/\partial a,t)$ is the Lagrangian density. Thus the action functional is given by

$$S[q] = \int_{t_0}^{t_1} L[q,\dot{q}]dt = \int_{t_0}^{t_1} dt \int_D [\tfrac{1}{2}\rho_0 \dot{q}^2 - \rho_0 U]d^3a. \tag{104}$$

Observe that this action functional is like that for finite-degree-of-freedom systems, as treated above, except that the sum over particles is replaced by integration over D, i.e.,

---

[12]For barotropic or isentropic flow, U depends only on ρ. For an ideal monoatomic gas $U(\rho,s) = c\rho^{\gamma-1}\exp(\alpha s)$, where c, γ, and α are constants. See, for example, Serrin (1959) for a discussion of equations of state.

$$\int_D d^3a \leftrightarrow \sum_i \ . \tag{105}$$

The mass of each "particle" of the continuum corresponds to $\rho_0 d^3a$.

The end conditions for Hamilton's principle for the fluid are the same as before,

$$\delta q(a,t_0) = \delta q(a,t_1) = 0. \tag{106}$$

However, in addition, boundary conditions are needed because there is now going to be integration by parts with respect to a. It is assumed that these are such that all surface terms vanish. Shortly we shall see what this implies.

In order to apply Hamilton's principle, we must functionally differentiate Eq. (104); thus it is necessary to know something about differentiating determinants. Recall

$$\frac{\partial q^k}{\partial a^j} \frac{A_k^i}{\mathcal{J}} = \delta_j^i \tag{107}$$

where A_k^i is the cofactor of $\partial q^k / \partial a^i =: q_{,i}^k$. (Remember repeated indices are to be summed.) A convenient expression for A_k^i is given by

$$A_k^i = \tfrac{1}{2} \epsilon_{kjl} \epsilon^{imn} \frac{\partial q^j}{\partial a^m} \frac{\partial q^l}{\partial a^n}, \tag{108}$$

where $\epsilon_{ijk} (= \epsilon^{ijk})$ is the skew-symmetric tensor (density), which vanishes if any two of i,j,k are equal; it is equal to 1 if i,j,k are unequal and a cyclic permutation of 1,2,3, and is otherwise equal to -1. In functionally differentiating Eq. (104) we shall require the following relation:

$$\frac{\partial \mathcal{J}}{\partial q_{,j}^i} = A_i^j, \tag{109}$$

which follows from Eq. (107).

For Lagrangian density functionals of the form $\mathcal{L}(q, \dot{q}, \partial q / \partial a, t)$, the functional derivative $\delta S / \delta q(a,t) = 0$ implies

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^i}\right) + \frac{\partial}{\partial a^j}\left(\frac{\partial \mathcal{L}}{\partial q_{,j}^i}\right) - \frac{\partial \mathcal{L}}{\partial q^i} = 0, \tag{110}$$

provided the surface integral vanishes:

$$\int_{t_0}^{t_1} \int_{\partial D} p \, \delta q^i A_i^j n_j d^2 a = \int_{t_0}^{t_1} \int_{\partial D} p \, \delta q \cdot \hat{n} d^2 q. \tag{111}$$

The equality above follows upon changing from integration over a to integration over q. Clearly the surface term vanishes if any of the following are true on ∂D:

(i) $\delta q^i = 0$,

(ii) $p = (\rho_0^2 / \mathcal{J}^2)(\partial U / \partial \rho) = 0$,

(iii) $\delta q \cdot \hat{n} = 0$,

where p is the pressure and $\hat{n}$ is a unit normal vector to ∂D. While all of these possibilities result in the vanishing

of the surface term, (i) is clearly more than is necessary, in light of (iii), which merely states that fluid particles are not forced through the boundary. When D is a box and periodic boundary conditions are imposed, the vanishing of the surface term is automatic. When D is "all space" one has the option (ii), which asserts that the pressure vanishes at infinity. Condition (iii) is the common physical condition.

From Eq. (110) the equation of motion is obtained,

$$\rho_0 \ddot{q}_i + A_i^j \frac{\partial}{\partial a^j}\left(\frac{\rho_0^2}{\mathcal{J}^2} \frac{\partial U}{\partial \rho}\right) = 0. \tag{112}$$

Here we have used $\partial A_i^j / \partial a^j = 0$, which you can work out using Eq. (108). Alternatively, upon using Eq. (107), we find that the equation of motion can be written in the form

$$\rho_0 \ddot{q}_j \frac{\partial q^j}{\partial a^i} + \mathcal{J} \frac{\partial}{\partial a^i}\left(\frac{\rho_0^2}{\mathcal{J}^2} \frac{\partial U}{\partial \rho}\right) = 0. \tag{113}$$

We leave it to you to show that Eq. (112) can be transformed into Eulerian form:

$$\rho\left(\frac{\partial v}{\partial t} + v \cdot \nabla v\right) = -\nabla p, \tag{114}$$

where $v = v(r,t)$. A useful identity in this regard is

$$\frac{\partial}{\partial q^k} = \frac{1}{\mathcal{J}} A_k^i \frac{\partial}{\partial a^i}. \tag{115}$$

With Eq. (115) it is clear that Eq. (112) is of the form of Newton's second law. The Legendre transform follows easily: the canonical momentum density is

$$\pi_i(a,t) := \frac{\delta L}{\delta \dot{q}^i(a)} = \rho_0 \dot{q}_i, \tag{116}$$

and

$$H[\pi,q] = \int d^3a[\pi \cdot \dot{q} - \mathcal{L}] = \int d^3a\left[\frac{\pi^2}{2\rho_0} + \rho_0 U\right]. \tag{117}$$

Hamilton's equations are then

$$\dot{\pi}_i = -\frac{\delta H}{\delta q^i}, \quad \dot{q}^i = \frac{\delta H}{\delta \pi_i}. \tag{118}$$

These equations can also be written in terms of the Poisson bracket,

$$\{F,G\} = \int \left[\frac{\delta F}{\delta q} \cdot \frac{\delta G}{\delta \pi} - \frac{\delta G}{\delta q} \cdot \frac{\delta F}{\delta \pi}\right] d^3a, \tag{119}$$

viz.,

$$\dot{\pi}_i = \{\pi_i, H\}, \quad \dot{q}^i = \{q^i, H\}. \tag{120}$$

Here $\delta q^i(a) / \delta q^j(a') = \delta_j^i \delta(a - a')$ has been used, a relation analogous to $\partial q^j / \partial q^i = \delta_i^j$ for finite systems [recall Eq. (45)].

In conclusion we point out that variational principles similar to that given above exist for essentially all ideal-

fluid models, including incompressible flow, magnetohydrodynamics, etc. One can even obtain directly two-dimensional scalar vortex dynamics by considering constrained variations, but we shall not pursue this here.

IV. NONCANONICAL HAMILTONIAN DYNAMICS AND EXAMPLES

Preparatory to discussing the Hamiltonian structure of the ideal fluid in terms of Eulerian variables, we now consider systems that are Hamiltonian, but written in terms of variables (or coordinates) that are not canonical. In Sec. IV.A we describe noncanonical Hamiltonian dynamics for systems with both finite and infinite degrees of freedom, and in Sec. IV.B we give many examples, which have been culled from the literature. In Sec. V we shall show how to derive the noncanonical Hamiltonian description for the ideal fluid from the canonical Lagrangian variable Hamiltonian description of Sec. III.C.

A. Noncanonical Hamiltonian dynamics

Let us start out by playing a sort of game. Suppose we have a system of ordinary differential equations:

$$\dot{z}^i = V^i(z), \quad i=1,2,\ldots,M. \tag{121}$$

How would you know if this system is a Hamiltonian system? If you came upon the equations during research you might have some idea based upon the physics, but assume that this is not the case here. What would you do?

One thing you might try is to check *Liouville's theorem*. Hamilton's equations have the property

$$\frac{\partial \dot{q}^i}{\partial q^i} + \frac{\partial \dot{p}_i}{\partial p_i} = \frac{\partial^2 H}{\partial q^i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q^i} = 0, \tag{122}$$

from which one can show that phase-space volume is conserved; i.e., if

$$\mathcal{V}(t) = \int \prod_{i=1}^{N} dp_i\, dq^i, \tag{123}$$

where the integration is over a volume interior to an arbitrary comoving surface, then

$$\frac{d\mathcal{V}}{dt} = 0. \tag{124}$$

The surface may distort, and in general it will do so in a major way, but the volume inside remains constant. The analogous statement for the system of Eq. (121) is incompressibility of the vector field; i.e.,

$$\frac{\partial V^i}{\partial z^i} = 0, \tag{125}$$

whence it follows that

$$\mathcal{V}(t) = \int \prod_{i=1}^{N} dz_i \tag{126}$$

is constant in time.

Suppose Eq. (125) is not true, as is the case for the following example:

$$\dot{z}_1 = -\frac{z_2^3}{3} - z_2,$$

$$\dot{z}_2 = \frac{z_1}{z_2^2 + 1}. \tag{127}$$

For this system

$$\frac{\partial \dot{z}_1}{\partial z_1} + \frac{\partial \dot{z}_2}{\partial z_2} = -\frac{2 z_1 z_2}{(z_2^2 + 1)^2} \neq 0. \tag{128}$$

You would be mistaken if, based on Eq. (128), you concluded that Eq. (127) was not Hamiltonian. In fact this system is a disguised simple harmonic oscillator. It has been disguised by making a noncanonical coordinate change, something that we shall discuss below.

So, is there a general method for determining whether or not a system is Hamiltonian? Probably the answer is no, since one must first find a Hamiltonian, and this requires a technique for finding constants of motion. There is no completely general way of doing this.[13] Nevertheless we can say some things, but to do so we must investigate Hamiltonian systems in arbitrary coordinates.

You might wonder, why would equations ever arise in noncanonical variables? Surely the physics would make things come out right. To the contrary, variables that are the most physically compelling need not be canonical variables. *The Eulerian variables that describe ideal continuous media are in general noncanonical.* Examples of systems that are typically written in terms of noncanonical variables are Liouville's equation for the dynamics of the phase-space density of a collection of particles, the BBGKY hierarchy of kinetic theory, the Vlasov equation of plasma physics, ideal fluid dynamics and various approximations thereof, magnetized fluids, ... : essentially every fundamental equation that describes classical media is of this type.

So with the above motivation, let us turn to discussing noncanonical Hamiltonian dynamics for systems with a finite number of degrees of freedom, using ideas that extend back (at least) to Sophus Lie. The first step is to write Hamilton's equations in covariant form. Thus we define

$$z^i = \begin{cases} q^i & \text{for } i=1,2,\ldots,N, \\ p_{i-N} & \text{for } i=N+1,\ldots,2N. \end{cases} \tag{129}$$

The z^i are coordinates on phase space which we denote by $\mathcal{Z}$. In terms of the z's Hamilton's equations take the compact form

---

[13]Techniques for finding constants of motion do exist, but necessarily possess limitations. See, for example, Ramani *et al.* (1989) and references therein.

$$\dot{z}^i = J_c^{ij} \frac{\partial H}{\partial z^j} = [z^i, H], \tag{130}$$

where the Poisson bracket is given by

$$[f,g] = \frac{\partial f}{\partial z^i} J_c^{ij} \frac{\partial g}{\partial z^j}, \tag{131}$$

with

$$(J_c^{ij}) = \begin{pmatrix} 0_N & I_N \\ -I_N & 0_N \end{pmatrix}. \tag{132}$$

Above, the repeated indices are to be summed over $1, 2, \ldots, 2N$. In Eq. (132), 0_N is an $N \times N$ matrix of zeros and I_N is the $N \times N$ unit matrix. The subscript c of J_c indicates that the system is written in terms of canonical coordinates. It is important to realize that we have only rewritten Hamilton's equations in new notation, albeit in a form that is suggestive.

Now consider a general, time-independent change of coordinates

$$\bar{z}^i = \bar{z}^i(z). \tag{133}$$

The Hamiltonian H transforms as a scalar:

$$H(z) = \bar{H}(\bar{z}). \tag{134}$$

Taking time derivatives of Eq. (133) yields

$$\dot{\bar{z}}^l = \frac{\partial \bar{z}^l}{\partial z^i} \dot{z}^i = \frac{\partial \bar{z}^l}{\partial z^i} J_c^{ij} \frac{\partial H}{\partial z^j} = \left[\frac{\partial \bar{z}^l}{\partial z^i} J_c^{ij} \frac{\partial \bar{z}^m}{\partial z^j} \right] \frac{\partial \bar{H}}{\partial \bar{z}^m}. \tag{135}$$

Upon defining

$$J^{lm} := \frac{\partial \bar{z}^l}{\partial z^i} J_c^{ij} \frac{\partial \bar{z}^m}{\partial z^j}, \tag{136}$$

we see that Hamilton's equations are covariant and that J^{lm}, which is called the *cosymplectic form*, transforms as a contravariant tensor of second rank. In the new variables, Hamilton's equations become

$$\dot{\bar{z}}^l = J^{lm}(\bar{z}) \frac{\partial \bar{H}}{\partial \bar{z}^m} = [\bar{z}^l, H], \tag{137}$$

where the Poisson bracket is now given by

$$[f,g] = \frac{\partial f}{\partial \bar{z}^l} J^{lm} \frac{\partial g}{\partial \bar{z}^m}. \tag{138}$$

Notice that in Eq. (137) we have displayed the explicit $\bar{z}$ dependence in J^{lm}. This was done to emphasize an important distinction—that between *covariance* and *form invariance*. Equation (136) is a statement of covariance, while a statement of form invariance is given by

$$J_c^{lm} = \frac{\partial \bar{z}^l}{\partial z^i} J_c^{ij} \frac{\partial \bar{z}^m}{\partial z^j}. \tag{139}$$

This is, in fact, the (most important) definition of a canonical transformation. Form invariance here means that the form of the J^{ij} and hence Hamilton's equations

remains the same. Evidently, the first N of $\bar{z}^l$ are coordinates, while the second N are momenta, so it is a simple matter to revert to the usual form of Hamilton's equations in the new canonical variables $\bar{z}^l$.

Let us now return to Liouville's theorem. Taking the divergence of Eq. (137) yields

$$\frac{\partial \dot{\bar{z}}^l}{\partial \bar{z}^l} = \frac{\partial J^{lm}}{\partial \bar{z}^l} \frac{\partial \bar{H}}{\partial \bar{z}^m} + J^{lm} \frac{\partial^2 \bar{H}}{\partial \bar{z}^l \partial \bar{z}^m}. \tag{140}$$

The second term vanishes because J^{lm} is antisymmetric and $\partial^2 \bar{H} / \partial \bar{z}^l \partial \bar{z}^m$ is symmetric. This is all there is for the usual Liouville's theorem, since in the canonical case J^{lm} is constant, so the first term vanishes. However, for Hamilton's equations written in noncanonical coordinates the following is typically (but not necessarily) true:

$$\frac{\partial \dot{\bar{z}}^l}{\partial \bar{z}^l} = \frac{\partial J^{lm}}{\partial \bar{z}^l} \frac{\partial \bar{H}}{\partial \bar{z}^m} \neq 0. \tag{141}$$

This was the situation for our little example of Eq. (127).

It might have occurred to you that changing the coordinates may hide but cannot destroy volume preservation. This is clear if we simply change coordinates in Eq. (123):

$$\mathcal{V}(t) = \int \mathcal{J} \prod_{l=1}^{2N} d\bar{z}^l. \tag{142}$$

If we include the Jacobian $\mathcal{J} := \det(\partial z^i / \partial \bar{z}^j)$ in the integrand, then Liouville's theorem is still satisfied. There is a nice formula relating $\mathcal{J}$ and J, which is obtained by taking the determinant of Eq. (136) and using the determinant product rule:

$$\mathcal{J} = \frac{1}{\sqrt{\det J}}. \tag{143}$$

Observe that there are many J's with the same $\mathcal{J}$.

Before leaving this discussion of Liouville's theorem we mention that even though J is a function of $\bar{z}$, it is still possible for $\partial \dot{\bar{z}}^l / \partial \bar{z}^l = 0$. This can happen because $\bar{H}$ is such that the two vectors of Eq. (141) are perpendicular or it may happen that $\partial J^{lm} / \partial \bar{z}^l = 0$, even though J is a function of $\bar{z}$. The latter case occurs for fluid models and underlies attempts to describe turbulence by using statistical mechanics concepts (see, for example, Montgomery and Kraichnan, 1979).

Now it is clear that the essence of what it means to be Hamiltonian does not lie in the canonical form. Where does it lie? It lies in some coordinate invariant properties of J. To illustrate this we shall play another sort of game. Suppose you have a system of the form of Eq. (121) and you want to know if it is Hamiltonian. Moreover, suppose you are clever enough to realize that Liouville is not the answer, because you know that Hamiltonian systems in noncanonical coordinates look like Eq. (137) with Poisson brackets that look like Eq. (138). Finally, suppose somehow you have found a constant of motion, call it H, and you think this is the en-

ergy and therefore a good bet for the Hamiltonian. Clearly if it is Hamiltonian, then

$$V^i(z) = J^{ij} \frac{\partial H}{\partial z^j}, \quad i = 1,2,\ldots,M. \tag{144}$$

Now, everything in Eq. (144) is known except J^{ij}, which is required to be antisymmetric because of definition (136). The antisymmetry automatically makes $dH/dt = 0$ and leaves M equations for $(M^2 - M)/2$ unknown quantities in J^{ij}. Suppose that with some fiddling around you have found a candidate J. [Try this for the simple example of Eq. (128).] Does a transformation exist such that you can transform the candidate J back to J_c?

The answer to this question is given by an old theorem that is credited to Darboux. If the J^{ij} you have found makes a good Poisson bracket, that is, when Eq. (138) is assembled, it satisfies

$$[f,g] = -[g,f] \quad \forall \ f,g, \tag{145}$$

$$[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = 0 \quad \forall \ f,g, \tag{146}$$

and moreover if $\det J \neq 0$, then Darboux says there exists a transformation (at least locally) where $J \to J_c$. Note, a requirement for $\det J \neq 0$ is that $M = 2N$, since odd-dimensional antisymmetric matrices have zero determinant. We shall not prove Darboux's theorem, but will mention that Eq. (146) is the important ingredient. This is an integrability condition known as the Jacobi identity; it is the central identity of a Lie algebra—a nonassociative algebra—which has a product with the properties (145) and (146). Above, the elements of the Lie algebra are functions defined on the phase space. We shall say more about this later in Sec. V.

The bracket properties, (145) and (146), can be translated into properties required of the cosymplectic form. The first is clearly

$$J^{ij} = -J^{ji}. \tag{147}$$

The second, with a little work, can be shown to be equivalent to

$$S^{ijk} := J^{il} \frac{\partial J^{jk}}{\partial z^l} + J^{jl} \frac{\partial J^{ki}}{\partial z^l} + J^{kl} \frac{\partial J^{ij}}{\partial z^l} = 0. \tag{148}$$

In going from Eq. (146) to Eq. (148) it is observed that all the terms involving second derivatives that arise upon calculating $[f,[g,h]] + [g,[h,f]] + [h,[f,g]]$ cancel; the only terms to survive are those where the derivative of the outer bracket acts upon the J of the inner bracket. This fact makes life much easier when verifying the Jacobi identity.

Now suppose everything worked out right except that the J you found had $\det J = 0$, with some rank $2N < M$. What then? A generalization of the Darboux theorem, which was proven by Lie long ago (see, for example, Eisenhart, 1961 and Littlejohn, 1982), says that J can be transformed (locally) into the following form:

$$(J_c) = \begin{pmatrix} 0_N & I_N & 0 \\ -I_N & 0_N & 0 \\ 0 & 0 & 0_{M-2N} \end{pmatrix}. \tag{149}$$
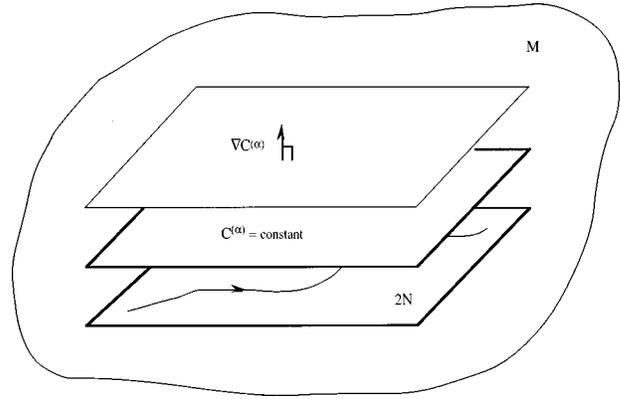
FIG. 12. The local structure of phase space in a noncanonical Hamiltonian system.

Interesting things happen in places where the rank of J changes. Later we shall say something about this, too.

From Eq. (149) it is clear that in the right coordinates the system is an N-degree-of-freedom system with some extraneous coordinates, $M - 2N$ in fact. The geometrical picture is as depicted in Fig. 12.

Through any point of the M-dimensional phase space $\mathcal{Z}$ there exists a regular Hamiltonian phase space $\mathcal{P}$ of dimension $2N$. These surfaces are called *symplectic leaves*. A consequence of the degeneracy is that there exists a special class of invariants that is built into the phase space. They are called *Casimir invariants*, a name which derives from the Lie algebra for angular momentum. Since the rank of J is $2N$, there exist possibly $M - 2N$ independent null eigenvectors. A consequence of the generalization of Darboux's theorem is that the independent null eigenvectors exist and, moreover, the null space can in fact be spanned by the gradients of the Casimir invariants, which satisfy

$$J^{ij} \frac{\partial C^\alpha}{\partial z^j} = 0, \tag{150}$$

where $\alpha = 1,2,3,\ldots,M-2N$. That the Casimir invariants are constants of motion follows from

$$\dot{C}^\alpha = \frac{\partial C^\alpha}{\partial z^i} J^{ij} \frac{\partial H}{\partial z^j} = 0. \tag{151}$$

Note that they are constants of motion for *any* Hamiltonian; they are, as noted above, built into the phase space and are in this sense kinematic constants of motion. The dynamics is determined by the Hamiltonian H. Note that the surfaces $\mathcal{P}$ of dimension $2N$ in the figure are the intersections of the $M - 2N$ surfaces defined by $C^\alpha = \text{const}$. Dynamics generated by any H that begins on a particular $\mathcal{P}$ surface remain there.

The picture we have described above is the finite-dimensional analog of the Hamiltonian form possessed by Eulerian continuous-media theories. We shall describe the Poisson brackets for some of them soon, but now we mention that if these media theories are written in terms of the density variables (for example, the ρ, σ, and M of Sec. III.C), then the cosymplectic operator will be seen to be linear in the variables. The finite-

dimensional analog of this is that J^{ij} has a special form that is linear in the z^k:

$$J^{ij} = c_k^{ij} z^k, \tag{152}$$

where the c_k^{ij} are constants—in fact, structure constants for a Lie algebra (see Sec. V). In light of Eqs. (147) and (148) they must satisfy

$$c_k^{ij} = -c_k^{ji} \tag{153}$$

and

$$c_m^{ij} c_l^{mk} + c_m^{jk} c_l^{mi} + c_m^{ki} c_l^{mj} = 0. \tag{154}$$

Brackets with J's of the form of Eq. (152) are called *Lie-Poisson brackets*.

It is interesting to reexamine the condition for Liouville's theorem (141) for J's of the above form,

$$\frac{\partial J^{lm}}{\partial z^l} \frac{\partial H}{\partial z^m} = c_l^{lm} \frac{\partial H}{\partial z^m} = 0. \tag{155}$$

In general, structure constants do not possess antisymmetry or symmetry upon interchange of an up with a down index. However, sometimes they do, as in the case of so(3) (see Sec. III.B). In general semisimple Lie algebras can, by a coordinate change, be brought into a form in which the structure constants are completely antisymmetric (Jacobson, 1962). In these coordinates there is Liouville's theorem without the need for inserting a Jacobian as in Eq. (143). This, as noted above, is typically the case for fluid theories in Eulerian variables.

In infinite dimensions the analog of Eq. (138) is given by

$$\{F,G\} = \int_D \frac{\delta F}{\delta \psi^i} \mathfrak{J}^{ij} \frac{\delta G}{\delta \psi^j} d\mu =: \left\langle \frac{\delta F}{\delta \psi}, \mathfrak{J} \frac{\delta G}{\delta \psi} \right\rangle, \tag{156}$$

where $\psi^i(\mu,t)$ are functions, $\mu = (\mu_1, \ldots, \mu_n)$ is a "spatial" or Eulerian observation variable, and ψ^i, $i = 1, \ldots, n$ are n components of the field. Now $\mathfrak{J}$ is an operator, and we require

$$\{F,G\} = -\{G,F\}, \tag{157}$$

$$\{F,\{G,H\}\} + \{G,\{H,F\}\} + \{H,\{F,G\}\} = 0 \tag{158}$$

for all *functionals* F, G, and H. Analogous to Eq. (147) the antisymmetry condition of Eq. (157) requires $\mathfrak{J}$ to be *skew symmetric*, i.e.,

$$\langle f, \mathfrak{J} g \rangle = \langle \mathfrak{J}^\dagger f, g \rangle = -\langle g, \mathfrak{J} f \rangle. \tag{159}$$

The Jacobi identity (158) for infinite-dimensional systems has a condition analogous to Eq. (148); one need only consider variations of $\mathfrak{J}$ when calculating $\{F,\{G,H\}\} + \{G,\{H,F\}\} + \{H,\{F,G\}\}$.[14] For Eulerian media, as noted above, the *cosymplectic operator* typically has the Lie-Poisson form

$$\mathfrak{J}^{ij} = \mathcal{C}_k^{ij} \psi^k, \tag{160}$$

where $\mathcal{C}_k^{ij}$ are *structure operators*. We shall clarify the meaning of these structure operators by examples; a bit more will be said in Sec. V.

B. Examples

Below the noncanonical Poisson brackets for several systems are presented. The first, the free rigid body, is a finite-dimensional system, while the others are infinite dimensional. We present the brackets here and refer the reader to the references for treatment of the Jacobi identity.

1. Free rigid body

The equations that govern the motion of the free rigid body are Euler's equations, the following three-dimensional system:

$$\ell_1 = \ell_2 \ell_3 \left(\frac{1}{I_3} - \frac{1}{I_2} \right),$$

$$\ell_2 = \ell_3 \ell_1 \left(\frac{1}{I_1} - \frac{1}{I_3} \right),$$

$$\ell_3 = \ell_1 \ell_2 \left(\frac{1}{I_2} - \frac{1}{I_1} \right), \tag{161}$$

which correspond to the statement of torque-free motion in a frame frozen into the body with axes aligned with the principal axes. (See Sec. V for more details.) The energy is purely rotational kinetic energy; since the coordinate axes are principal axes it takes the form

$$H = \frac{1}{2} \sum_{i=1}^3 \frac{\ell_i^2}{I_i}. \tag{162}$$

The function H is easily seen to be a constant of motion upon differentiating with respect to time and making use of Eq. (161). The Poisson bracket for this system is of Lie-Poisson type,

$$[f,g] = -\epsilon_{ijk} \ell_k \frac{\partial f}{\partial \ell_i} \frac{\partial g}{\partial \ell_j}. \tag{163}$$

The structure constants are ϵ_{ijk}, which are those of SO(3), that is, the group of rotations. The Jacobi identity is assured since the ϵ_{ijk}, being structure constants, satisfy condition (154)—something that is not difficult to verify directly.[15] It is evident upon substituting Eq. (162) into Eq. (163) that

$$\ell_i = [\ell_i, H], \tag{164}$$

which is equivalent to Eq. (161). This system possesses the Casimir invariant

---

[14]This is the infinite-dimensional analog of Eq. (148), which is described further by Lax (1975) and Morrison (1981a, 1982).

[15]The structure constants for any Lie algebra will satisfy condition (154). The algebra SO(2,1) is appropriate for describing the dynamics of the Kida vortex, as shown in Meacham *et al.* (1997).

$$C = \tfrac{1}{2} \sum_{i=1}^{3} \ell_i^2, \tag{165}$$

which satisfies

$$[C, f] = 0, \quad \forall \ f. \tag{166}$$

Thus the global picture of the phase space $\mathcal{Z}$, which here corresponds to Fig. 12, is one where the symplectic leaves are nested two-dimensional spheres in the three-dimensional space with coordinates (ℓ_1, ℓ_2, ℓ_3).

2. Korteweg–de Vries equation

We write the famous Korteweg–de Vries (KdV) equation (see, for example, Whitham, 1974), which describes long-wavelength water waves or ion-acoustic waves in plasmas, in the following form:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0. \tag{167}$$

Here $x \in D$, which can be (and typically is) chosen to be $(-\infty, \infty)$ or $[-\pi, \pi]$. In the former case the appropriate boundary condition is $u(\pm \infty) = 0$, while in the latter case periodic boundary conditions are appropriate. The KdV equation possesses a countable infinity of constants of motion, but the one that is of interest now is the following:

$$H = \int_D \left[\tfrac{1}{6} u^3 - \tfrac{1}{2} \left(\frac{\partial u}{\partial x} \right)^2 \right] dx. \tag{168}$$

The noncanonical Poisson bracket, due to Gardner (1971), is given by

$$\{F, G\} = -\int_D \frac{\delta F}{\delta u} \frac{\partial}{\partial x} \frac{\delta G}{\delta u} dx, \tag{169}$$

from which it is seen that the cosymplectic operator is

$$\mathfrak{J} = -\frac{\partial}{\partial x}. \tag{170}$$

The skew symmetry of Eq. (169) follows upon integration by parts; the Jacobi identity can be shown to be automatic since the cosymplectic operator is independent of u. Inserting the functional derivative of Eq. (168),

$$\frac{\delta H}{\delta u} = \left(\tfrac{1}{2} u^2 + \frac{\partial^2 u}{\partial x^2} \right), \tag{171}$$

into Eq. (169) yields

$$\frac{\partial u}{\partial t} = \{u, H\} = -\frac{\partial}{\partial x} \left(\tfrac{1}{2} u^2 + \frac{\partial^2 u}{\partial x^2} \right) = -u \frac{\partial u}{\partial x} - \frac{\partial^3 u}{\partial x^3}. \tag{172}$$

This bracket possesses one Casimir invariant,

$$C[u] = \int_D u \, dx. \tag{173}$$

It is easily verified that $\{C, F\} = 0$ for all functionals F. The phase space $\mathcal{Z}$ in this case is infinite dimensional—it being a function space composed of all admissible functions u. The symplectic leaves are of one fewer dimension, but they are also infinite dimensional.

Note that the bracket above is not linear in u and is therefore not of Lie-Poisson form, in spite of the fact that we have claimed that the standard Hamiltonian form for theories of media is of this type. You may know that the KdV equation is special—it being integrable by the inverse scattering method—so it is not too surprising that it has a Hamiltonian structure that is inconsistent with the credo. Although the basic equations that describe media in terms of Eulerian variables have the Lie-Poisson form, when approximations are made this form can change.

3. One-dimensional pressureless fluid

Now we consider an equation even simpler than the KdV equation, that of a one-dimensional pressureless fluid,

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0. \tag{174}$$

This equation has, in jest, been referred to as both the dispersionless KdV equation and the inviscid Burger's equation. That it models a fluid suggests that the Hamiltonian ought to be just the kinetic-energy functional,

$$H[u] = \int_D \tfrac{1}{2} u^2 \, dx, \tag{175}$$

there being no internal energy. The following bracket, with the above Hamiltonian, produces Eq. (174):

$$\{F, G\} = -\tfrac{1}{3} \int_D u \left[\frac{\delta F}{\delta u} \frac{\partial}{\partial x} \frac{\delta G}{\delta u} - \frac{\delta G}{\delta u} \frac{\partial}{\partial x} \frac{\delta F}{\delta u} \right] dx, \tag{176}$$

that is,

$$\frac{\partial u}{\partial t} = \{u, H\} = -\tfrac{1}{3} \left(u \frac{\partial u}{\partial x} + \frac{\partial (u^2)}{\partial x} \right) = -u \frac{\partial u}{\partial x}. \tag{177}$$

The cosymplectic operator is clearly given by

$$\mathfrak{J} = -\tfrac{1}{3} \left(u \frac{\partial}{\partial x} + \frac{\partial}{\partial x} u \right) = -\tfrac{1}{3} \left(2u \frac{\partial}{\partial x} + \frac{\partial u}{\partial x} \right). \tag{178}$$

The following Casimir invariant is easily obtained by solving $\{C, F\} = 0$ for all functionals F, i.e., by searching for null eigenvectors of Eq. (178) and undoing the functional derivative:

$$C = \int_D |u|^{1/2} \ dx. \tag{179}$$

It is evident that the Hamiltonian

$$H[u] = \tfrac{1}{6} \int_D u^3 \, dx, \tag{180}$$

together with the bracket (169), will also produce Eq. (174). Thus it is possible for a system to have two Hamil-

tonian structures: two functionally independent Hamiltonians with two distinct Poisson brackets. This rarity occurs for the above system, the KdV equation, and other systems. It is a symptom of integrability (see Magri, 1978).

4. One-dimensional compressible fluid

Now we consider a somewhat more complicated model, the one-dimensional compressible fluid with a pressure that depends only upon the density. The equations of motion for this system are the following:

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} - \frac{1}{\rho}\frac{\partial p}{\partial x},$$

$$\frac{\partial \rho}{\partial t} = -\frac{\partial(\rho u)}{\partial x}. \tag{181}$$

The Hamiltonian has a kinetic-energy part plus an internal-energy part,

$$H[\rho,u] = \int_D [\tfrac{1}{2}\rho u^2 + \rho U(\rho)]dx, \tag{182}$$

and the Poisson bracket[16] is given by

$$\{F,G\} = -\int_D \left[\frac{\delta F}{\delta \rho}\frac{\partial}{\partial x}\frac{\delta G}{\delta u} - \frac{\delta G}{\delta \rho}\frac{\partial}{\partial x}\frac{\delta F}{\delta u}\right]dx. \tag{183}$$

The cosymplectic operator,

$$(\mathfrak{J}^{ij}) = \begin{pmatrix} 0 & \dfrac{\partial}{\partial x} \\[2mm] \dfrac{\partial}{\partial x} & 0 \end{pmatrix}, \tag{184}$$

is seen to be skew symmetric upon integration by parts and systematic neglect of the surface terms. The Jacobi identity follows since the cosymplectic operator is independent of the dynamical variables.

Observe that this bracket, like the two above, is not Lie-Poisson. However, upon transforming from the dependent variables (u,ρ) to (M,ρ), where $M = \rho u$, it obtains the Lie-Poisson form. We shall not do this transformation here but consider this below when we treat the ideal fluid in three spatial dimensions.

Setting $\{F,C\} = 0$ for all F yields two equations,

$$\frac{\partial}{\partial x}\frac{\delta C}{\delta \rho} = 0, \quad \frac{\partial}{\partial x}\frac{\delta C}{\delta u} = 0, \tag{185}$$

from which we obtain the following Casimir invariants:

---

[16]A bracket of this type was given by Morrison and Greene (1980), Morrison (1982), and Dubrovin and Novikov (1984). In the last reference such brackets were termed brackets of *hydrodynamic type* and an interesting connection between the cosymplectic operator and a metric on function space was obtained: the Jacobi identity was shown to corrrespond to flatness.

$$C_1[u] = \int_D u\,dx, \quad C_2[\rho] = \int_D \rho\,dx. \tag{186}$$

Using

$$\frac{\delta H}{\delta u} = \rho u,$$

$$\frac{\delta H}{\delta \rho} = \tfrac{1}{2}u^2 + h(\rho), \tag{187}$$

where $h(\rho) := \rho U_\rho + U$ is the enthalpy (note that $\delta H/\delta\rho = $ constant is Bernoulli's law), in Eq. (183) produces

$$\frac{\partial u}{\partial t} = \{u,H\} = -\frac{\partial}{\partial x}(\tfrac{1}{2}u^2 + U + \rho U_\rho),$$

$$\frac{\partial \rho}{\partial t} = \{\rho,H\} = -\frac{\partial}{\partial x}(\rho u). \tag{188}$$

These equations can be seen to be equivalent to Eqs. (181) upon making use of $h_x = p_x/\rho$ (recall $p = \rho^2 U_\rho$).

5. Two-dimensional Euler scalar vortex dynamics

The vortex dynamics we consider here, unlike the examples above, have two spatial variables, $r := (x,y) \in D$, in addition to time; that is, this is a $2+1$ theory. The noncanonical Poisson bracket possessed by this system (Morrison, 1981a, 1982; Olver, 1982; Marsden and Weinstein, 1983) is the prototype of $2+1$ theories, it being shared by the 1D Vlasov-Poisson equation (Morrison, 1980, 1981b), quasigeostrophy or the Hasegawa-Mima equation (Weinstein, 1983a; Zakharov and Piterbarg, 1988), and others.

The single dynamical variable for the 2D Euler equation is the scalar vorticity, defined by

$$\omega(r,t) := \hat{z}\cdot\nabla\times v, \tag{189}$$

where v is the Eulerian velocity field and $\hat{z}$ is the ignored coordinate. The velocity field is assumed to be nondivergent, $\nabla\cdot v = 0$, and hence the stream function ψ is introduced:

$$v = \left(-\frac{\partial\psi}{\partial y}, \frac{\partial\psi}{\partial x}\right), \tag{190}$$

which is related to the vorticity through

$$\omega = \nabla^2\psi. \tag{191}$$

The equation of motion for this system is

$$\frac{\partial\omega}{\partial t} = -v\cdot\nabla\omega = -[\psi,\omega], \tag{192}$$

where

$$[f,g] = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}. \tag{193}$$

There is some subtlety with the boundary conditions. The physical boundary condition for the ideal fluid is that no flow penetrates the boundary, i.e., the normal

component of v vanishes. This amounts to ψ being constant on ∂D. Since ω is the dynamical variable, one might expect, from a mathematics point of view, the boundary condition to be ω constant on ∂D. Then it is natural to set variations of ω to zero on the boundary to eliminate surface terms obtained upon integration by parts. Although this boundary condition may be correct for the Vlasov-Poisson equation, it is unphysical for the ideal fluid where the vorticity at a point on the boundary is generally not constant. When boundary terms do not vanish with physical boundary conditions, generally the mathematics is signalling something physical. In this case it is signalling the fact that surfaces of constant vorticity possess dynamics, an idea that is the basis of the "contour dynamics" approximation technique. To describe this is beyond the scope of these notes (see Lewis *et al.*, 1986). However, all these complications can be avoided by choosing the domain D to be a finite box and imposing periodic boundary conditions. Alternatively, D can be chosen to be $\mathbb{R}^2$ with vanishing vorticity at infinity; however, as is well known in electrostatics, this requires a potential that diverges logarithmically.

The energy in this model is purely kinetic. Thus the Hamiltonian is given by

$$H[\omega] = \tfrac{1}{2} \int_D v^2 d^2r = \tfrac{1}{2} \int_D |\nabla \psi|^2 d^2r$$

$$= \tfrac{1}{2} \int_D \int_D \omega(r) K(r|r') \omega(r') d^2r d^2r'$$

$$= -\tfrac{1}{2} \int_D \omega \psi d^2r, \tag{194}$$

where K is defined by

$$\psi(r) = - \int_D K(r|r') \omega(r') d^2r'. \tag{195}$$

Observe that in the case where $D = \mathbb{R}^2$ the last equality of Eq. (194) requires the elimination of the logarithmic singularity that comes from integration by parts. The noncanonical Poisson bracket for this system is given by

$$\{F, G\} = \int_D \omega \left[\frac{\delta F}{\delta \omega}, \frac{\delta G}{\delta \omega} \right] d^2r, \tag{196}$$

which is of the Lie-Poisson form. The cosymplectic operator in this case is

$$\mathfrak{J} = -[\omega, \cdot]. \tag{197}$$

Skew symmetry follows from

$$\int_D f[g, h] d^2r = - \int_D g[f, h] d^2r, \tag{198}$$

which is obtained upon integration by parts and neglect of the boundary terms. The Jacobi identity for $\mathfrak{J}$ is inherited from that for [,], as is the case for Lie-Poisson brackets. The Casimir invariant for the bracket of Eq. (196) is given by

$$C[\omega] = \int \mathcal{C}(\omega) d^2r, \tag{199}$$

where $\mathcal{C}$ is an arbitrary function. Since $\mathcal{C}$ is arbitrary, C in fact constitutes an infinity of constants of motion. These arise from the incompressibility of phase space.[17] We mention that even though there are an infinity of constants this is insufficient for the 2D Euler equations to be integrable. In order to obtain the equations of motion we require

$$\frac{\delta H}{\delta \omega} = -\psi. \tag{200}$$

Evidently,

$$\frac{\partial \omega}{\partial t} = \{\omega, H\} = -\left[\omega, \frac{\delta H}{\delta \omega}\right] = [\omega, \psi], \tag{201}$$

which is equivalent to Eq. (192).

6. Three-dimensional ideal fluid

For this last example we consider the ideal fluid in three dimensions, our first example of a $3+1$ theory where the spatial variables are the rectangular coordinates $r := (x, y, z) = :(x_1, x_2, x_3) \in D$. The dynamical variables that we use were introduced in Sec. III: the three components of the Eulerian velocity field v, the density ρ, and the entropy per unit mass s. We use s rather than the pressure p, but it is a simple matter to alter this. The equations of motion are

$$\frac{\partial v}{\partial t} = -v \cdot \nabla v - \frac{1}{\rho} \nabla p, \tag{202}$$

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho v), \tag{203}$$

$$\frac{\partial s}{\partial t} = -v \cdot \nabla s. \tag{204}$$

Recall that the thermodynamics is embodied in an internal-energy function $U(\rho, s)$, from which, in addition to the pressure $p = \rho^2 U_\rho$, the temperature is given by $T = U_s$.

The Hamiltonian functional is given by

$$H[v, \rho, s] = \int_D [\tfrac{1}{2} \rho v^2 + \rho U(\rho, s)] d^3r, \tag{205}$$

and the noncanonical Poisson bracket (Morrison and Greene, 1980) is

---

[17] If one divides D up into small cells and attaches a value of ω to each cell, then these constants state that the number of cells with a given value of attached ω remains constant in time. Thus the dynamics is restricted to be a rearrangement. See Gardner (1963) and Morrison (1987).

$$\{F,G\}=-\int_D\left[\left(\frac{\delta F}{\delta\rho}\nabla\cdot\frac{\delta G}{\delta v}-\frac{\delta G}{\delta\rho}\nabla\cdot\frac{\delta F}{\delta v}\right)\right.$$

$$+\left(\frac{\nabla\times v}{\rho}\cdot\frac{\delta G}{\delta v}\times\frac{\delta F}{\delta v}\right)$$

$$\left.+\frac{\nabla s}{\rho}\cdot\left(\frac{\delta F}{\delta s}\frac{\delta G}{\delta v}-\frac{\delta G}{\delta s}\frac{\delta F}{\delta v}\right)\right]d^3r. \qquad (206)$$

This bracket is familiar in that the first term is the generalization to three dimensions of that for the $1+1$ compressible fluid given above. Similarly, recognizing that via the chain rule $\delta F/\delta v=\nabla\times\delta F/\delta\omega$, we can see that the second term is a $3+1$ generalization of that for the $2+1$ scalar vortex dynamics given above.[18] The third term is not familiar, but had we included entropy in our $1+1$ fluid theory its one-dimensional counterpart would have been present.

Using

$$\frac{\delta H}{\delta v}=\rho v, \quad \frac{\delta H}{\delta\rho}=\tfrac{1}{2}v^2+(\rho U)_\rho, \quad \frac{\delta H}{\delta s}=\rho U_s, \quad (207)$$

we find that Eqs. (202)–(204) are equivalent to

$$\frac{\partial v}{\partial t}=\{v,H\}, \quad \frac{\partial\rho}{\partial t}=\{\rho,H\}, \quad \frac{\partial s}{\partial t}=\{s,H\}. \qquad (208)$$

In order to obtain the equations of motion from the above and in order to prove the Jacobi identity, integrations by parts must be performed and surface terms involving functionals must be neglected. The boundary condition appropriate for the ideal fluid, as noted above, is $\hat{n}\cdot v=0$ on ∂D, but this is a boundary condition on v, not on the functionals directly. The function space of functionals must be such that these terms vanish for all functionals. When D is a finite domain, there is a complication with the vanishing of these terms, as in the case for the 2D Euler equations. This problem is not an issue when periodic boundary conditions are used or when $D=\mathbb{R}^3$, for in these cases the space of functionals can be defined readily. When D is a finite domain one might try to eliminate the surface terms by requiring all functionals to satisfy $\hat{n}\cdot\delta F/\delta v=0$, but it must be verified that this space of functionals is closed, i.e., the bracket of two functionals with this property produces a functional with this property. An alternative method that circumvents this complication is to build the boundary condition into the Hamiltonian by a suitable potential-energy functional.

It is evident that the Poisson bracket of Eq. (206) is not of Lie-Poisson form. However, if a transformation from the variables v, ρ, and s to the conserved variables $M:=\rho v$, ρ, and $\sigma:=\rho s$, which were introduced in Sec. III (and alluded to above), is made, then the bracket (Morrison and Greene, 1980; a commutator description is given in Dzyaloshinskii and Volovick, 1980) becomes

$$\{F,G\}=-\int_D\left[M_i\left(\frac{\delta F}{\delta M_j}\frac{\partial}{\partial x^j}\frac{\delta G}{\delta M_i}-\frac{\delta G}{\delta M_j}\frac{\partial}{\partial x^j}\frac{\delta F}{\delta M_i}\right)\right.$$

$$+\rho\left(\frac{\delta F}{\delta M}\cdot\nabla\frac{\delta G}{\delta\rho}-\frac{\delta G}{\delta M}\cdot\nabla\frac{\delta F}{\delta\rho}\right)$$

$$\left.+\sigma\left(\frac{\delta F}{\delta M}\cdot\nabla\frac{\delta G}{\delta\sigma}-\frac{\delta G}{\delta M}\cdot\nabla\frac{\delta F}{\delta\sigma}\right)\right]d^3r. \qquad (209)$$

This transformation requires the use of the chain rule for functional derivatives, which gives formulas such as the following:

$$\left.\frac{\delta F}{\delta\rho}\right|_{v,s}=\left.\frac{\delta F}{\delta\rho}\right|_{M,s}+\frac{M}{\rho}\cdot\frac{\delta F}{\delta M}+\frac{\sigma}{\rho}\frac{\delta F}{\delta\sigma}. \qquad (210)$$

It is straightforward to show that the bracket (209) together with the Hamiltonian

$$H[M,\rho,\sigma]=\int_D\left[\frac{M^2}{2\rho}+\rho U(\rho,\sigma/\rho)\right]d^3r \qquad (211)$$

produces the fluid equations of motion in conservation form as

$$\frac{\partial M}{\partial t}=\{M,H\}, \quad \frac{\partial\rho}{\partial t}=\{\rho,H\}, \quad \frac{\partial\sigma}{\partial t}=\{\sigma,H\}. \qquad (212)$$

Now consider the condition for the Casimir invariants, $\{F,C\}=0$ for all F. From Eq. (206) it can be seen that this implies

$$\nabla\cdot\frac{\delta C}{\delta v}=0, \quad \frac{1}{\rho}\nabla s\cdot\frac{\delta C}{\delta v}=0,$$

$$\nabla\frac{\delta C}{\delta\rho}+\frac{(\nabla\times v)}{\rho}\times\frac{\delta C}{\delta v}-\frac{\nabla s}{\rho}\frac{\delta C}{\delta s}=0. \qquad (213)$$

One solution of these equations is

$$C_1[\rho,s]=\int_D\rho f(s)d^3r, \qquad (214)$$

where f is an arbitrary function. If we eliminate the entropy variable s from the theory, then another solution is the helicity

$$C_2[v]=\int_D v\cdot\nabla\times v\,d^3r. \qquad (215)$$

One can also show that there is a potential vorticity Casimir invariant (Padhye and Morrison, 1996a; 1996b), but it will be left to you to investigate the general solution of Eqs. (213).

7. General comments

Above we have presented a variety of noncanonical Poisson brackets, of one, two, and three spatial dimensions and of one or more field variables, culminating in that of the three-dimensional fluid with the field vari-

---

[18]See also, for three-dimensional vortex dynamics, Kuznetsov and Mikhailov (1980).

ables (v, ρ, s) or (M, ρ, σ).[19] In closing this section we make some brief comments about the classification of the various brackets.

Consider the cases in which there is only a single field variable. We presented two such $1+1$ theories, that for the KdV equation and that for the pressureless fluid. It is natural to ask whether or not these brackets are in some sense equivalent. Is it possible by a coordinate change to map one into the other? A simple scaling analysis suggests that a quadratic transformation might do this. Indeed the transformation

$$\widetilde{u} = \tfrac{1}{6} u^2, \tag{216}$$

with the inverse[20]

$$u = \sqrt{6\widetilde{u}}, \tag{217}$$

demonstrates this equivalence. Inserting

$$\frac{\delta F}{\delta u} = \sqrt{\frac{2\widetilde{u}}{3}} \frac{\delta F}{\delta \widetilde{u}} \tag{218}$$

into the KdV bracket yields

$$\{F, G\} = -\int_D \frac{\delta F}{\delta u} \frac{\partial}{\partial x} \frac{\delta G}{\delta u} \, dx$$

$$= -\tfrac{1}{3} \int_D \widetilde{u} \left[\frac{\delta F}{\delta \widetilde{u}} \frac{\partial}{\partial x} \frac{\delta G}{\delta \widetilde{u}} - \frac{\delta G}{\delta \widetilde{u}} \frac{\partial}{\partial x} \frac{\delta F}{\delta \widetilde{u}} \right] dx. \tag{219}$$

Now it is evident from Eq. (219) and from above, where we changed fluid variables from (v, ρ, s) to (M, ρ, σ), that sometimes brackets can be mapped into the Lie-Poisson form by an invertible transformation. The study of when this can be done is an interesting area that we shall not address here. However, since typically for fluid theories this transformation can be done, this suggests a classification of such theories by their Lie-Poisson brackets, which in turn are classified by the Lie groups (see Sec. V) corresponding to their structure operators. Thus theories can be classified by a Lie group[21] and the corresponding Casimir invariants are determined. In the case of $1+1$ theories discussed above, the group is that of coordinate changes and the algebra is in essence the infinitesimal generator $\partial/\partial x$. In the case of the $2+1$

theory of Euler's fluid equations the group is the group of canonical transformations of the plane, or equivalently area-preserving transformations. When one increases the number of spatial dimensions the possibilities increase. When more than one field variable is considered the groups become more complicated: they are groups by extension, such as the direct product or the semidirect product. Treatment of this area is beyond the scope of these lectures, although we shall comment briefly on this in the context of Clebsch variables in Sec. V.

V. TUTORIAL ON LIE GROUPS AND ALGEBRAS, REDUCTION, AND CLEBSCH VARIABLES

This section begins with Sec. V.A, a tutorial on Lie groups and Lie algebras, which was in fact a lecture that was given after the others. It is placed here for continuity, and it is intended to require little background and to be a bare-bones minimum needed for understanding Secs. V.B and V.C. It can be skipped by the cognoscenti, who may proceed directly to Sec. V.B, where reduction, a procedure for obtaining smaller Hamiltonian systems from bigger ones, is described. Reduction is employed in Sec. V.B.4 to derive the noncanonical Hamiltonian description of the ideal fluid in terms of Eulerian variables from the canonical Hamiltonian description in terms of Lagrangian variables. Section V is concluded with Sec. V.C, where an inflation procedure is described for "canonizing" noncanonical Lie-Poisson brackets by transforming to Clebsch variables. The less mathematically inclined may skip Sec. V altogether and proceed, with little loss of continuity, directly to Sec. VI, where stability is discussed.

A. Tutorial on Lie groups and Lie algebras

The study of Lie groups can be demanding because it mixes both algebra and geometry: a Lie group $\mathfrak{G}$ is both a group and a differentiable manifold. To simplify matters, our discussion will be a local[22] one; i.e., calculations will be done in terms of coordinates, in a style reminiscent of Lie's original work.[23] The elements of a Lie group, which are uncountably infinite in number, correspond to points of the manifold, which for our purpose can be thought of as Euclidean space. To be concrete we shall consider a realization in which elements of $\mathfrak{G}$ correspond to functions that define transformations (coor-

---

[19]Additional examples relevant to geophysical fluid dynamics can be found in Shepherd (1990).

[20]This transformation is invertible if u and $\bar{u}$ are positive, a property that in some cases is preserved by the dynamics. One can extend this by using $\bar{u} = \mathrm{sgn}(u) u^2/6$, which is one-to-one. Formal manipulation with this transformation produces the same bracket as (219).

[21]This classification idea is an old one. It appears in Sudarshan (1963) and Sudarshan and Mukunda (1974) in the context of describing dynamics by canonical realizations of Lie groups (cf. Sec. V). Development in the geometrical setting was given by Arnold (1966a, 1969), who, in contrast to our approach, emphasized the Lagrange bracket. See also Weinstein (1983b) and Arnold et al. (1990) for finite systems.

[22]The term local is used in contradistinction to *global*, by which is meant the study of the topological nature of surfaces in the large. Such a study was begun by Riemann and Poincaré; results are predominately a product of the present century.

[23]For background and further reading the reader is directed to the following references, which are at roughly the same level and in the same language as that given here: Eisenhart (1961), Hammermesh (1962), Gürsey (1963), and Loewner (1971).

dinate changes) on a space $\mathcal{Z}$, which can also be viewed locally as Euclidean space. We describe this in detail in Sec. V.A.1, following with a discussion of Lie algebras in Sec. V.A.2, and concluding this tutorial with definitions of the terms realization and representation in Sec. V.A.3.

1. Lie groups

Suppose the space $\mathcal{Z}$ has coordinates z^i, $i=1,2,\ldots,M$, and a family of transformations is given by

$$z'^i = f^i(z,a), \quad i=1,2,\ldots,M, \tag{220}$$

where $z=(z^1,z^2,\ldots,z^M)\in\mathcal{Z}$ and $a=(a^1,a^2,\ldots,a^N)\in\mathfrak{G}$ denotes a parametrization of the family. For each value of a the functions f constitute a one-to-one transformation of $\mathcal{Z}$ onto itself. For convenience we denote this by T_a. Thus $T_a:\mathcal{Z}\to\mathcal{Z}$ and $z'=T_a z$. The set of T_a's forms a group under composition of functions.

It is important to distinguish between the M-dimensional space $\mathcal{Z}$ and the N-dimensional group manifold $\mathfrak{G}$. The latter is called either the *parameter space*, *group space*, or the *group manifold*. We are introducing $\mathcal{Z}$ now so that you have something concrete to visualize, but this is really unnecessary—it could be done completely in the abstract.

Another distinction to be made is between the *passive* and *active* viewpoints of the transformation T_a. In the passive viewpoint (adopted above) the point of $\mathcal{Z}$ remains fixed and T_a represents a change in the coordinates used to identify the point. In the active viewpoint there are dynamics of a sort; a point of $\mathcal{Z}$ is mapped onto a new point. Below you are, for the most part, free to think in terms of either viewpoint.

The group product, as noted above, is composition. *Closure* requires the existence of a group element T_c such that

$$T_c z = T_b T_a z \tag{221}$$

for all T_b and T_c. Hence there must be a function $\phi(b,a)=c$. It is this function that really defines the group. If one assumes that ϕ possesses three derivatives in each of its arguments, it is a wonderful thing that this guarantees the existence of all derivatives. We shall see how this goes, but not work it out in detail. In terms of the functions of Eqs. (220), closure can be stated as follows:

$$f(z,c)=f(f(z,a),b)=f(z,\phi(b,a)). \tag{222}$$

A simple example of a Lie group is that of SO(2), rotations of the plane. These are linear transformations given by

$$\begin{bmatrix} z'^1 \\ z'^2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} z^1 \\ z^2 \end{bmatrix}, \tag{223}$$

which we write as

$$z' = T_\theta z. \tag{224}$$

This is a one-parameter group with $\theta\in[0,2\pi)$. Closure requires that a rotation through an angle θ followed by a rotation through an angle ψ be equivalent to a rotation through some angle χ:

$$T_\psi T_\theta = \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\psi+\theta) & \sin(\psi+\theta) \\ -\sin(\psi+\theta) & \cos(\psi+\theta) \end{bmatrix} =: T_\chi. \tag{225}$$

Clearly, the analog of $c=\phi(b,a)$ is $\chi=\phi(\psi,\theta)=\psi+\theta$, mod 2π.

You may know that, in addition to closure, groups have three other properties: associativity, the existence of an identity, and the existence of an inverse. These properties are natural if you think about elements of the group corresponding to coordinate changes.

Associativity requires

$$T_a(T_b T_c) = (T_a T_b)T_c. \tag{226}$$

Since $T_a T_b z = f(f(z,b),a)$, the right-hand side is

$$(T_a T_b)T_c z = f(f(f(z,c),b),a) = f(f(z,c),\phi(a,b))$$

$$= f(z,\phi(\phi(a,b),c)) = T_{\phi(\phi(a,b),c)} z. \tag{227}$$

Similarly $T_b T_c = T_{\phi(b,c)}$, and the left-hand side is

$$T_a(T_b T_c) = T_{\phi(a,\phi(b,c))}. \tag{228}$$

Upon comparing Eqs. (227) and (228) we see that associativity implies

$$\phi(a,\phi(b,c)) = \phi(\phi(a,b),c). \tag{229}$$

This relation is clearly not satisfied for all functions ϕ; it in fact places a strict restriction on the functions that may define a group product, as we shall see.

The *identity* element of the group is denoted by T_0. It must satisfy

$$T_0 T_a = T_a T_0 = T_a \tag{230}$$

or

$$f(f(z,a),0) = f(f(z,0),a) = f(z,a). \tag{231}$$

Therefore

$$\phi(0,a) = \phi(a,0) = a. \tag{232}$$

For every element a of a group $\mathfrak{G}$ there must exist an *inverse*, which we denote by a^{-1}, such that

$$T_a T_{a^{-1}} = T_{a^{-1}} T_a = T_0. \tag{233}$$

Evidently,

$$\phi(a,a^{-1}) = \phi(a^{-1},a) = 0. \tag{234}$$

In order for these equations to have a unique solution for a^{-1}, given a,

$$\det\left(\frac{\partial\phi(a,b)}{\partial a}\right) \neq 0; \quad \det\left(\frac{\partial\phi(a,b)}{\partial b}\right) \neq 0. \tag{235}$$

It is easy to verify the above properties for the example of SO(2); it is recommended that you do this.

Now let us return to the quest of determining what the group properties say about ϕ. Taylor-expanding ϕ about $a=b=0$ through third order yields

$$\phi^\nu(a,b) = \phi^\nu(0,0) + \frac{\partial \phi^\nu(0,0)}{\partial a^\kappa} a^\kappa + \frac{\partial \phi^\nu(0,0)}{\partial b^\kappa} b^\kappa + \frac{1}{2} \frac{\partial^2 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda} a^\kappa a^\lambda + \frac{\partial^2 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda} a^\kappa b^\lambda + \frac{1}{2} \frac{\partial^2 \phi^\nu(0,0)}{\partial b^\kappa \partial b^\lambda} b^\kappa b^\lambda$$

$$+ \frac{1}{3!} \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda \partial a^\mu} a^\kappa a^\lambda a^\mu + \frac{1}{2} \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda \partial b^\mu} a^\kappa a^\lambda b^\mu + \frac{1}{2} \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda \partial b^\mu} a^\kappa b^\lambda b^\mu$$

$$+ \frac{1}{3!} \frac{\partial^3 \phi^\nu(0,0)}{\partial b^\kappa \partial b^\lambda \partial b^\mu} b^\kappa b^\lambda b^\mu + \mathcal{O}(4), \tag{236}$$

where derivatives with respect to a are taken in the first slot of ϕ^ν and those with respect to b in the second. Since

$$\phi(a,0) = \phi(0,a) = a \tag{237}$$

for all a, it is clear that $\phi(0,0) = 0$, and upon differentiating Eq. (237) we obtain

$$\frac{\partial \phi^\nu(0,0)}{\partial a^\kappa} = \delta^\nu_\kappa,$$

$$\frac{\partial \phi^\nu(0,0)}{\partial b^\kappa} = \delta^\nu_\kappa. \tag{238}$$

Differentiating Eq. (237) twice and then thrice in the nonzero argument implies

$$\frac{\partial^2 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda} = \frac{\partial^2 \phi^\nu(0,0)}{\partial b^\kappa \partial b^\lambda} = \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda \partial a^\mu} = \frac{\partial^3 \phi^\nu(0,0)}{\partial b^\kappa \partial b^\lambda \partial b^\mu}$$

$$= 0. \tag{239}$$

However, Eq. (237) does not contain information about mixed derivatives; viz.,

$$\frac{\partial^2 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda}, \quad \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda \partial b^\mu}, \quad \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda \partial b^\mu}. \tag{240}$$

So, thus far we have reduced Eq. (236) to

$$\phi^\nu(a,b) = a^\nu + b^\nu + \frac{\partial^2 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda} a^\kappa b^\lambda$$

$$+ \frac{1}{2} \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda \partial b^\mu} a^\kappa a^\lambda b^\mu$$

$$+ \frac{1}{2} \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda \partial b^\mu} a^\kappa b^\lambda b^\mu + \mathcal{O}(4). \tag{241}$$

To go farther the associativity condition (229) is imposed. If you expand this condition through second order, in anticipation of a result, you will be disappointed. Associativity places no constraint to this order. If you attempt to expand through third order you will also be disappointed because you will generate a tedious mess. Nevertheless, perseverance and a tad of cleverness results in a condition on ϕ. If we define

$$c^\nu_{\kappa\lambda} := \frac{\partial^2 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda} - \frac{\partial^2 \phi^\nu(0,0)}{\partial a^\lambda \partial b^\kappa}, \tag{242}$$

which obviously satisfies

$$c^\nu_{\kappa\lambda} = -c^\nu_{\lambda\kappa}, \tag{243}$$

the condition obtained is

$$c^\nu_{\kappa\lambda} c^\lambda_{\delta\gamma} + c^\nu_{\gamma\lambda} c^\lambda_{\kappa\delta} + c^\nu_{\delta\lambda} c^\lambda_{\gamma\kappa} = 0. \tag{244}$$

The numbers $c^\nu_{\kappa\lambda}$ were called *structure constants* by Sophus Lie. They are the heart of the matter.

You might wonder what happens to next order. It turns out that Eqs. (243) and (244) are enough to determine ϕ, the group product rule, and thus the essence of the group. (Technically, only the part that is connected to the identity is determined.)

2. Lie algebras

Studying the group manifold in a neighborhood of the identity leads to a study of Lie algebras. Also, it leads to differential equations for ϕ, equations with important consequences that we shall use in Sec. V.B where we discuss reduction. To this end, suppose δa is small, that is, near to the identity, and consider

$$T_{\tilde{a}} := T_a T_{\delta a} \tag{245}$$

or

$$\tilde{a} = \phi(a, \delta a). \tag{246}$$

Since ϕ was assumed to be continuous, $\tilde{a}$ must be near a, and so we write

$$\tilde{a} = a + da = \phi(a, \delta a) \tag{247}$$

or, in terms of the transformations,

$$T_{a+da} z = T_a T_{\delta a} z. \tag{248}$$

This is depicted in Fig. 13.
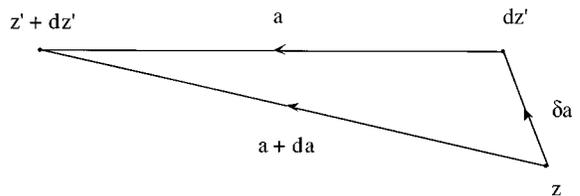
FIG. 13. Depiction of the Lie group transformation $T_{a+da} z = T_a T_{\delta a} z$, a relation used for deriving Lie group generators.

Taylor-expanding Eq. (246) about a and $b = \delta a = 0$ yields

$$a^\alpha + da^\alpha = \phi^\alpha(a,0) + \frac{\partial \phi^\alpha(a,b)}{\partial b^\beta}\bigg|_{b=0} \delta a^\beta + \cdots, \quad (249)$$

where the greek indices α, β, etc., which we shall use to denote coordinates of the group manifold, run over $1, 2, \ldots, N$. From Eq. (249),

$$da^\alpha = L^\alpha_\beta(a) \delta a^\beta \tag{250}$$

where

$$L^\alpha_\beta(a) := \frac{\partial \phi^\alpha(a,b)}{\partial b^\beta}\bigg|_{b=0}. \tag{251}$$

Consider now a function defined on the group manifold $F : \mathfrak{G} \to \mathbb{R}$. How does $F(a)$ differ from $F(\tilde{a})$?

$$dF(a) := F(\tilde{a}) - F(a) = F(a + da) - F(a)$$

$$\approx \frac{\partial F}{\partial a^\alpha} da^\alpha = \frac{\partial F}{\partial a^\alpha} L^\alpha_\beta(a) \delta a^\beta$$

$$=: \delta a^\beta X_\beta F(a). \tag{252}$$

Clearly, if we choose $F(a) = a$, Eq. (252) implies

$$da^\gamma = \delta a^\beta X_\beta a^\gamma = L^\gamma_\beta \delta a^\beta. \tag{253}$$

We shall use this later.

The quantities X_β defined by

$$X_\beta := L^\alpha_\beta(a) \frac{\partial}{\partial a^\alpha}, \tag{254}$$

are called the *infinitesimal generators* of the Lie group. They are in fact elements of the *Lie algebra*, $\mathfrak{g}$, associated with $\mathfrak{G}$. The elements of the Lie algebra X_β are first-order linear differential operators that add like ordinary vectors. In fact, the X_β are to be thought of as vectors with components L^α_β and basis vectors $\partial/\partial a^\alpha$. These quantities also possess a natural product rule, $[X_\alpha, X_\beta]$, which is the ordinary commutator. Plainly the commutator of two first-order linear differential operators is another such operator, so the product rule is closed.

For SO(2) the group product rule is defined by $\chi = \phi(\psi, \theta) = \psi + \theta$, whence we obtain for the generators

$$X_\psi = \frac{\partial}{\partial \psi}, \quad X_\theta = \frac{\partial}{\partial \theta}.$$

The Lie algebra $\mathfrak{g}$ in this case is composed of linear combinations of the above quantities and the product is simply $[X_\psi, X_\theta] = 0$. This algebra is particularly simple in that the product always vanishes.

Now we shall obtain a differential equation for the group, and then discuss briefly some important theorems proven by Lie. Recall Eq. (253), which we derived by expanding $T_{a+da} = T_a T_{\delta a}$. Since $\phi(a, a^{-1}) = 0$ must be solvable for a^{-1} for any a, this implies $L^\alpha_\beta(a)$ must have an inverse for all a. We call this $L^{-1\beta}_\gamma$, i.e.,
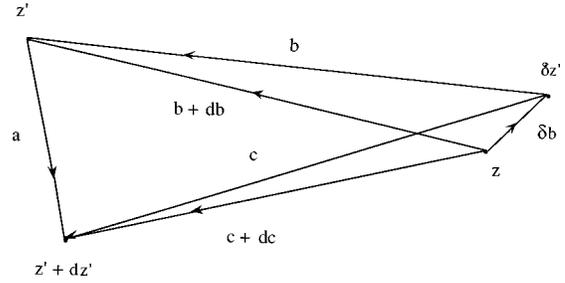
FIG. 14. Depiction of the Lie group transformation $T_{c+dc} = T_a T_{b+db} = T_a T_b T_{\delta b}$, a relation used for deriving the differential equations for a Lie group.

$$L^\alpha_\beta L^{-1\beta}_\gamma = \delta^\alpha_\gamma \tag{255}$$

and Eq. (253) can be inverted,

$$\delta a^\beta = L^{-1\beta}_\alpha da^\alpha. \tag{256}$$

Now suppose

$$T_{c+dc} = T_a T_{b+db} = T_a T_b T_{\delta b}, \tag{257}$$

which is depicted in Fig. 14.

Equation (257) implies

$$c + dc = \phi(a, b + db) = \phi(a, \phi(b, \delta b)). \tag{258}$$

If $dc = db = \delta b = 0$, then $c = \phi(a,b)$; thus Eq. (258) becomes, by associativity,

$$c + dc = \phi(\phi(a,b), \delta b) = \phi(c, \delta b). \tag{259}$$

Therefore

$$c^\alpha + dc^\alpha = \phi^\alpha(c, \delta b) = \phi^\alpha(c,0) + L^\alpha_\beta(c) \delta b^\beta + \cdots, \tag{260}$$

and

$$dc^\alpha = L^\alpha_\beta(c) \delta b^\beta = L^\alpha_\beta(c) L^{-1\beta}_\gamma(b) db^\gamma, \tag{261}$$

where the second equality follows from Eq. (256). We can see that

$$\frac{\partial c^\alpha}{\partial b^\gamma} = L^\alpha_\beta(c) L^{-1\beta}_\gamma(b), \tag{262}$$

but since $c = \phi(a,b)$

$$\frac{\partial \phi^\alpha(a,b)}{\partial b^\gamma} = L^\alpha_\beta(\phi(a,b)) L^{-1\beta}_\gamma(b). \tag{263}$$

Equation (263) is a system of partial differential equations of a form known as the Mayer-Lie type. Here $\phi(a,b)$ is the unknown and a is a fixed parameter. It is a system that determines the b dependence of ϕ with the a dependence coming from the initial condition $\phi(a,0) = a$. A similar equation holds where the roles of a and b are reversed. In order for a system of equations of this type to possess a solution, it must satisfy an integrability condition, viz.,

$$\frac{\partial^2 \phi^\alpha(a,b)}{\partial b^\mu \partial b^\gamma} = \frac{\partial^2 \phi^\alpha(a,b)}{\partial b^\gamma \partial b^\mu}, \tag{264}$$

which implies

$$\frac{\partial}{\partial b^\mu}[L_\beta^\alpha(\phi(a,b))L_\gamma^{-1\beta}(b)]$$

$$=\frac{\partial}{\partial b^\gamma}[L_\beta^\alpha(\phi(a,b))L_\mu^{-1\beta}(b)]. \qquad (265)$$

Performing the differentiation in Eq. (265),

$$\frac{\partial L_\beta^\alpha(c)}{\partial c^\nu}\frac{\partial \phi^\nu(a,b)}{\partial b^\mu}L_\gamma^{-1\beta}(b)+L_\beta^\alpha(c)\frac{\partial L_\gamma^{-1\beta}(b)}{\partial b^\mu}$$

$$=\frac{\partial L_\beta^\alpha(c)}{\partial c^\nu}\frac{\partial \phi^\nu(a,b)}{\partial b^\gamma}L_\mu^{-1\beta}(b)+L_\beta^\alpha(c)\frac{\partial L_\mu^{-1\beta}(b)}{\partial b^\gamma},$$

$$(266)$$

and then using Eq. (263) yields

$$\frac{\partial L_\beta^\alpha(c)}{\partial c^\nu}[L_\delta^\nu(c)L_\mu^{-1\delta}(b)L_\gamma^{-1\beta}(b)$$

$$-L_\delta^\nu(c)L_\gamma^{-1\delta}(b)L_\mu^{-1\beta}(b)]$$

$$=L_\beta^\alpha(c)\left[\frac{\partial L_\mu^{-1\beta}(b)}{\partial b^\gamma}-\frac{\partial L_\gamma^{-1\beta}(b)}{\partial b^\mu}\right]. \qquad (267)$$

Now the left-hand side can be made a function of c alone and the right-hand side can be made a function of b alone, by multiplying by "$L(b)L(b)L^{-1}(c)$" with the appropriate indices. We obtain

$$L_\alpha^{-1\gamma}(c)\left[\frac{\partial L_\beta^\alpha(c)}{\partial c^\nu}L_\delta^\nu(c)-\frac{\partial L_\delta^\alpha(c)}{\partial c^\nu}L_\beta^\nu(c)\right]$$

$$=L_\delta^\nu(b)L_\beta^\alpha(b)\left[\frac{\partial L_\nu^{-1\gamma}(b)}{\partial b^\alpha}-\frac{\partial L_\alpha^{-1\gamma}(b)}{\partial b^\nu}\right]. \qquad (268)$$

Since the points b and c were arbitrary, the two sides of Eq. (268) must equal the same constants. Upon setting $c=0$ these constants are determined to be the structure constants, $c_{\beta\delta}^\gamma$, defined by Eq. (242). Using

$$L_\delta^\nu(0)=L_\delta^{-1\nu}(0)=\delta_\delta^\nu \qquad (269)$$

and

$$\frac{\partial L_\beta^\alpha(0)}{\partial c^\nu}=\frac{\partial^2\phi^\alpha(0,0)}{\partial c^\nu\partial b^\beta} \qquad (270)$$

yields, upon setting the two sides of Eq. (268) equal to the constants, the following two equations:

$$\frac{\partial L_\beta^\alpha(c)}{\partial c^\nu}L_\gamma^\nu(c)-\frac{\partial L_\gamma^\alpha(c)}{\partial c^\nu}L_\beta^\nu(c)=c_{\gamma\beta}^\delta L_\delta^\alpha(c), \qquad (271)$$

$$\frac{\partial L_\nu^{-1\delta}(b)}{\partial b^\gamma}-\frac{\partial L_\gamma^{-1\delta}(b)}{\partial b^\nu}=c_{\alpha\beta}^\delta L_\nu^{-1\alpha}(b)L_\gamma^{-1\beta}(b), \qquad (272)$$

which can be shown directly to be equivalent.

Equation (272) is an important equation known as the Maurer-Cartan equation. Since its left-hand side is a

"curl," the divergence of its right-hand side must vanish. This is true provided the structure constants satisfy Eq. (244). Therefore Eq. (272) can be solved for $L^{-1}(b)$. With this value of L^{-1}, Eq. (263) is solved for ϕ.

Above we have described the connection between Lie groups and the Lie algebra of generators. It needs to be emphasized that Lie proved the remarkable theorem that makes this connection: given the Lie algebra of generators[24]

$$[X_\alpha,X_\beta]=c_{\alpha\beta}^\gamma X_\gamma, \qquad (273)$$

where the structure constants $c_{\alpha\beta}^\gamma$ satisfy Eqs. (243) and (244), or equivalently

$$[X_\alpha,X_\beta]=-[X_\beta,X_\alpha], \qquad (274)$$

$$[X_\alpha,[X_\beta,X_\gamma]]+[X_\beta,[X_\gamma,X_\alpha]]+[X_\gamma,[X_\alpha,X_\beta]]=0, \qquad (275)$$

there exists some Lie group for which the c's are the structure constants. Moreover, in the vicinity of the identity this group is unique. The proof of this theorem in the general case is difficult. It requires a deep understanding of the structure of Lie algebras; namely, that any Lie algebra can be decomposed into the sum of two kinds of algebras—a semisimple algebra and a solvable algebra. It is not possible to pursue this within the confines of a single lecture like this (see, for example, Jacobson, 1962).

3. Realization and representation

At the beginning of this tutorial we used the term realization to describe the group $\mathfrak{G}$ that was composed of transformations on the space $\mathcal{Z}$. In closing this tutorial we define more carefully, but briefly, what is meant by the term realization, and in addition we define what is meant by the companion term representation.

Let us begin by supposing that an abstract group has been defined. To be specific we shall assume that a and b are elements of the abstract group and that the product element of the group is given by $\phi(a,b)$. A *realization* of the abstract group is a specific group for which there is a mapping from the abstract group to the specific group that preserves the group product. For example, the group of transformations, $T_a,T_b,\ldots$ of Sec. V is a realization of the abstract group with the product rule $\phi(a,b)$. Preservation of the product rule amounts to

$$T_aT_b=T_{\phi(a,b)}.$$

An important realization possessed by all abstract groups is the group of transformations on itself. This is defined by $T_a\cdot:=\phi(a,\cdot)$, which is easily shown to preserve the product rule.

A *representation* of a group is a realization in terms of transformations that are linear. A simple example of such a group is the Lie group SO(2) as defined above.

---

[24]Note for the example of SO(2) that the structure constants $c_{\alpha\beta}^\gamma$ are all zero. Such algebras are called Abelian.

The notions of realization and representation do not apply to groups alone, but to any algebraic structure. For example, a Lie algebra could be realized by a set of matrices with a product defined by the ordinary matrix commutator. Another realization of a Lie algebra, the *canonical realization*, is composed of functions on phase space together with a product defined by the Poisson bracket. Canonical realizations arise naturally by the procedure of reduction, to which we now turn.

B. Reduction

Reduction is a procedure for obtaining from a given Hamiltonian system one of smaller dimension. The idea dates back to Lie, Poincaré, and Cartan. Reduction produces examples of generating dynamics via a canonical realization of a Lie group, which is a subgroup associated with a Lie algebra composed of the ordinary Poisson bracket and a selected set of functions defined on phase space.[25]

There are two parts to reduction: kinematic and dynamic. The kinematic part is concerned with the use of special variables that have a certain closure property, while the dynamic part refers to a type of symmetry of the Hamiltonian, viz., that the Hamiltonian be expressible in terms of the special variables. The symmetry of the Hamiltonian can motivate the choice of the reduced variables. For example, for the ideal fluid the form of the Hamiltonian suggests the use of Eulerian variables as a reduced set (see Sec. V.B.4).

The symmetry of the Hamiltonian gives rise to one or more constants of motion (Casimirs) that can, in principle, be used to reduce the order of the system. However, the term reduction is, in a sense, a misnomer since in actuality the procedure does not reduce the order of a system, but splits the system in such a way that it can be integrated in stages.

In this subsection we discuss reduction in general terms for finite systems and then consider a reduction that we term *standard reduction*, in which the new variables are linear in the momenta. This is followed by two examples: the free rigid body and the ideal fluid. In the last subsection we discuss Clebsch variables, a reduction that is bilinear in canonical coordinates and momenta.

1. Reduction of finite-dimensional systems

In the first part of reduction, that which pertains to kinematics, the system is transformed into a useful set of (generally) noncanonical coordinates. To see how this goes, we begin with the canonical Poisson bracket

$$[f,g] = \frac{\partial f}{\partial z^i} J_c^{ij} \frac{\partial g}{\partial z^j}, \tag{276}$$

where $i,j = 1,2,...,2N$, and recall

[25]See, for example, Sudarshan (1963), Sudarshan and Mukunda (1974), Marsden and Weinstein (1974), and Arnold *et al.* (1990).

$$(J_c^{ij}) = \begin{bmatrix} 0_N & I_N \\ -I_N & 0_N \end{bmatrix}, \tag{277}$$

and

$$z = (q,p), \tag{278}$$

and then suppose we have a set of functions $w^\alpha(z)$, with $\alpha = 1,...,M$, where in general, these functions are noninvertible functions of z and $M < 2N$. Let us also suppose that f and g obtain their z dependence through the functions w, i.e.,

$$f(z) = \bar{f}(w(z)). \tag{279}$$

Differentiation of Eq. (279) yields

$$\frac{\partial f}{\partial z^i} = \frac{\partial \bar{f}}{\partial w^\alpha} \frac{\partial w^\alpha}{\partial z^i}, \tag{280}$$

which upon insertion into Eq. (276) gives

$$[f,g] = \frac{\partial f}{\partial w^\alpha} \frac{\partial g}{\partial w^\beta} \left(\frac{\partial w^\alpha}{\partial z^i} J_c^{ij} \frac{\partial w^\beta}{\partial z^j} \right), \tag{281}$$

where we have dropped the overbar. The quantity

$$J^{\alpha\beta} := \frac{\partial w^\alpha}{\partial z^i} J_c^{ij} \frac{\partial w^\beta}{\partial z^j} \tag{282}$$

is in general a function of z. However, it is possible that $J^{\alpha\beta}$ can be written as a function of w only. When this closure condition occurs, we have a reduction. Said another way, we have reduction if we obtain a Lie algebra realization composed of the functions w and the Poisson bracket.

In order for functions of w together with the bracket

$$[f,g] = \frac{\partial f}{\partial w^\alpha} \frac{\partial g}{\partial w^\beta} J^{\alpha\beta}(w) \tag{283}$$

to be a Lie algebra, it is necessary for $[\ ,\ ]$ to satisfy the Jacobi identity for all such functions. This is equivalent to

$$S^{\alpha\beta\gamma}(w) := J^{\alpha\delta} \frac{\partial J^{\beta\gamma}}{\partial w^\delta} + J^{\gamma\delta} \frac{\partial J^{\alpha\beta}}{\partial w^\delta} + J^{\beta\delta} \frac{\partial J^{\gamma\alpha}}{\partial w^\delta} = 0. \tag{284}$$

(Recall Sec. IV.) Substituting Eq. (282) into Eq. (284) gives

$$S^{\alpha\beta\gamma}(w) = \frac{\partial w^\alpha}{\partial z^i} J^{ij} \frac{\partial w^\delta}{\partial z^j} \frac{\partial}{\partial w^\delta} \left(\frac{\partial w^\beta}{\partial z^k} J^{kl} \frac{\partial w^\gamma}{\partial z^l} \right) + \cdots \tag{285}$$

$$= \frac{\partial w^\alpha}{\partial z^i} J^{ij} \frac{\partial}{\partial z^j} \left(\frac{\partial w^\beta}{\partial z^k} J^{kl} \frac{\partial w^\gamma}{\partial z^l} \right) + \cdots \tag{286}$$

$$= [w^\alpha, [w^\beta, w^\gamma]] + [w^\beta, [w^\gamma, w^\alpha]]$$
$$+ [w^\gamma, [w^\alpha, w^\beta]] = 0, \tag{287}$$

where the last equality follows from the original Jacobi identity applied to the functions w^α. Thus any reduction produces a bracket that satisfies the Jacobi identity.

Now consider briefly the second part of reduction, that which concerns the symmetry property of the Hamiltonian. In order to have a reduced description of the dynamics, i.e., one entirely in terms of the w's, the original Hamiltonian $H(z)$ must be expressible solely in terms of these variables, i.e., there must exist a function $\bar{H}(w)$ such that

$$H(z)=\bar{H}(w). \tag{288}$$

Equation (288) is in fact a statement of symmetry, since the function $H(z)$ in reality depends on a fewer number of variables, the w's. This is a condition that must be verified case by case, but it is not difficult if one knows the generators of the symmetry.

2. Standard reduction

For standard reduction the functions w have the special form

$$w_k=L_k^i(q)p_i, \tag{289}$$

where $i,k=1,2,\ldots,N$. Note that the greek index α of the previous subsection has been changed to k and it now sums to $M=N$. The index has also been lowered for reasons that will become clear in a moment. Writing out Eq. (282),

$$J_{kl}=\frac{\partial w_k}{\partial p_i}\frac{\partial w_l}{\partial q^i}-\frac{\partial w_l}{\partial p_i}\frac{\partial w_k}{\partial q^i}, \tag{290}$$

and inserting Eq. (289) into Eq. (290) yields

$$J_{kl}=L_k^i\frac{\partial L_l^j}{\partial q^i}p_j-L_l^i\frac{\partial L_k^j}{\partial q^i}p_j=\left(L_k^i\frac{\partial L_l^j}{\partial q^i}-L_l^i\frac{\partial L_k^j}{\partial q^i}\right)p_j. \tag{291}$$

Closure occurs if constant numbers c_{kl}^m can be found such that

$$\left(L_k^i\frac{\partial L_l^j}{\partial q^i}-L_l^i\frac{\partial L_k^j}{\partial q^i}\right)p_j=c_{kl}^mL_m^jp_j=c_{kl}^mw_m. \tag{292}$$

The form of Eq. (292) may ring a bell. Recall the discussion in Sec. V.A where we talked about integrability and obtained the Maurer-Cartan equation. From Eq. (271) it is clear that if the L's are chosen to be the components of the infinitesimal generators of some Lie algebra, then Eq. (292) holds, if the constant numbers c_{kl}^m are the structure constants of the Lie algebra.

Since we lowered the indices above, the structure constants have one contravariant index and two covariant indices, which is the opposite of that of Sec. IV. We changed them here so that the identification above with

the Maurer-Cartan equation would come out right.[26] We emphasize, though, that the important thing about this construction is that there exist L's that make Eq. (292) an identity. These L's will then be suitable for defining a transformation of the form of Eq. (289) that yields a reduced Poisson bracket. The reduced bracket is of the Lie-Poisson form with the indices given as follows:

$$J_{kl}=c_{kl}^mw_m. \tag{293}$$

Sometimes a minus sign is needed in the cosymplectic form in order for the equations to come out right; for example, this was the case for the rigid-body bracket of Sec. IV.B. It is evident that if the Jacobi identity (154) is satisfied for c_{kl}^m, then it will be satisfied for $-c_{kl}^m$. Also, it is clear that there exist reductions to brackets with both signs, since condition (292) will be satisfied by $L_k^i \to -L_k^i$ if $c_{kl}^m \to -c_{kl}^m$. Thus Lie-Poisson brackets come in pairs.

Let us now turn to the question of whether or not the Hamiltonian can be written in terms of the reduced variables $H(z)=\bar{H}(w)$. First consider the important case in which the Hamiltonian has the following quadratic form:

$$H(q,p)=\frac{1}{2}\mu^{ij}(q)p_ip_j, \tag{294}$$

where the "metric" satisfies $\mu^{ij}(q)=\mu^{ji}(q)$. Since we have supposed that L_k^i originates from a Lie group, it must, as described in Sec. V.A.2, have an inverse. Therefore inserting

$$p_i=L_i^{-1k}w_k \tag{295}$$

into Eq. (294) yields

$$H(q,p)=\frac{1}{2}\mu^{ij}(q)L_i^{-1k}L_j^{-1l}w_kw_l. \tag{296}$$

Thus we obtain the following condition[27] for reduction:

$$\mu^{ij}(q)L_i^{-1k}L_j^{-1l}=\mu_0^{kl},$$

where μ_0 is independent of q.

Now consider the case of a general Hamiltonian. We will have reduction if

$$H(q,p)=H(q,L^{-1}(q)w)=\bar{H}(w), \tag{297}$$

which upon differentiation with respect to q yields

$$\left(\frac{\partial H}{\partial q^l}+\frac{\partial H}{\partial p_i}\frac{\partial L_i^{-1k}}{\partial q^l}w_k\right)\Bigg|_{p=L^{-1}(q)w}\equiv0, \tag{298}$$

---

[26]Technically, Eq. (289) defines a transformation from phase space with coordinates (q,p), i.e., $T^*\mathfrak{G}$, to the algebra $\mathfrak{g}^*$, which is the dual of the Lie algebra $\mathfrak{g}$. The dual algebra $\mathfrak{g}^*$ is the space of real linear functionals on $\mathfrak{g}$ and thus has the indices reversed. Clearly there is more to this story: see Souriau (1970) or Marsden and Weinstein (1974) for a discussion in terms of the current terminology of geometry.

[27]This condition is automatically satisfied if the configuration space is a Lie group and $\mu^{ij}(q)$ is a left-invariant metric. Such metrics are easily constructed by translating a μ_0 defined at the identity. See Appendix 2 of Arnold (1978).

for $l = 1, 2, \ldots, N$. Equation (298) is a general, necessary and sufficient condition for reduction of the Hamiltonian. Clearly this condition is satisfied by any Hamiltonian that is obtained from an $\bar{H}$ according to $H(q,p) = \bar{H}(L(q)p)$.

Since reduction involves the symmetry of Eq. (297) and symmetries are related to constants of motion, it should come as no surprise that a general expression for constants of motion, which are the Casimir invariants, comes along with the reduction framework. A clean way of seeing this is afforded by *triple-bracket* dynamics (Bialynicki-Birula and Morrison, 1991), which is a generalization of a formalism due to Nambu (1973).

This triple-bracket construction begins by considering a semisimple Lie algebra with structure constants c_{ij}^{k} and metric tensor[28] g_{ij} which is given by

$$g_{ij} = c_{il}^{k} c_{jk}^{l}. \tag{299}$$

This quantity can be used to raise and lower indices.

The fact that the structure constants have three indices hints at the existence of a geometric bracket operation on three functions, and it would be appealing if all three functions appeared on equal footing. This can be achieved by using the fully antisymmetric form of the structure constants,

$$c_{ijk} = g_{il} c_{jk}^{l}, \tag{300}$$

from which the following triple bracket is constructed:

$$[f, g, h] = c_{ijk} \frac{\partial f}{\partial w_i} \frac{\partial g}{\partial w_j} \frac{\partial h}{\partial w_k}. \tag{301}$$

A simple relationship exists between $[f,g,h]$ and the Lie-Poisson bracket $[f,g]$. This is made manifest by inserting the Casimir of the Lie algebra, as given by

$$C := \frac{1}{2} g^{ij} w_i w_j, \tag{302}$$

into one of the slots of the triple bracket:

$$[f,g] = [f,g,C], \tag{303}$$

where $[f,g]$ has the cosymplectic form of Eq. (293). Due to this relationship, time evolution can be represented as follows:

$$\frac{df}{dt} = [f, H, C], \tag{304}$$

where f is an arbitrary dynamical variable. In this formulation the dynamics are determined by two generating functions, the Hamiltonian H and the Casimir C; because of the complete antisymmetry, the Casimir is necessarily conserved.

3. Reduction of the free rigid body

The free rigid body, which is a sort of prototype for reduction, is a good example because it is finite dimen-
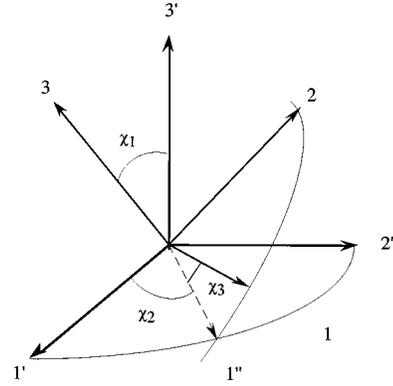
FIG. 15. The Euler angles, χ_1, χ_2, and χ_3 that are used to describe the rigid body.

sional and the computations are relatively easy. A free rigid body is a rigid body that is subject to no external forces, and thus a frame of reference can be assumed in which the center of mass is fixed. It takes three numbers to specify the state of the body: if a mark is placed on (or in) the body as a reference point, then two angles specify the orientation of the line from the center of mass to the mark, while another angle is needed to specify the orientation relative to the line; i.e., the location of another mark (not along the line) is determined by a rotation about the line. Thus the dimension of the configuration space $\mathcal{Q}$, for the free rigid body, is three. A traditional set of coordinates is provided by the Euler angles[29] $\chi = (\chi_1, \chi_2, \chi_3)$, which are defined by Fig. 15.

Clearly, the rotation matrix $O(\chi)$ that takes the primed into the unprimed axes is the product of three rotations through the three Euler angles.

By imagining infinitesimal rotations $\delta\chi$, or by consulting a mechanics book, you can obtain the following formulas relating the angular velocities, relative to a set of rectangular coordinates fixed in the body, to the time rate of change of the Euler angles:

$$\omega_1 = \dot{\chi}_1 \cos\chi_3 + \dot{\chi}_2 \sin\chi_1 \sin\chi_3,$$

$$\omega_2 = -\dot{\chi}_1 \sin\chi_3 + \dot{\chi}_2 \sin\chi_1 \cos\chi_3,$$

$$\omega_3 = \dot{\chi}_3 + \dot{\chi}_2 \cos\chi_1. \tag{305}$$

This formula can be derived from the rotation matrix: $\omega_i = \frac{1}{2} \epsilon_{ijk} O_{lk} \dot{O}_{lj} = \frac{1}{2} \epsilon_{ijk} O_{lk} O_{lj,m} \dot{\chi}_m =: D_{im} \dot{\chi}_m$. The body axes are convenient since in these axes the moment-of-inertia tensor is constant in time and one can choose them so that the moment-of-inertia tensor is diagonal, the so-called principal axes. In these coordinates the Lagrangian is deceptively simple,

$$L(\chi, \dot{\chi}) = \frac{1}{2}(I_1 \omega_1^2 + I_2 \omega_2^2 + I_3 \omega_3^2), \tag{306}$$

it being merely the kinetic energy since there are no external forces. Note, however, that upon insertion of

---

[28]This metric is also called either the trace form or the Killing form. A semisimple Lie algebra is one for which g_{ij} has no zero eigenvalues; hence g_{ij} has an inverse, which we denote by g^{ij} (see, for example, Jacobson, 1962).

[29]Since the structure constants for the rotation group are purely antisymmetric, it is customary to use all lowered indices. We follow this custom in this subsection.

Eq. (305) the Lagrangian becomes a complicated function of χ and $\dot{\chi}$. I will leave it as an exercise for you to calculate the equations of motion.

Since the Lagrangian is convex in $\dot{\chi}$ we can effect the Legendre transformation. The canonical momenta are given by

$$p_i = \frac{\partial L}{\partial \dot{\chi}_i} = \frac{\partial L}{\partial \omega_j} \frac{\partial \omega_j}{\partial \dot{\chi}_i} = \ell_j \frac{\partial \omega_j}{\partial \dot{\chi}_i}, \tag{307}$$

i.e.,

$$p_i = L_{ij}^{-1}(\chi)\ell_j, \tag{308}$$

where

$$(L^{-1}) = \begin{pmatrix} \cos\chi_3 & -\sin\chi_3 & 0 \\ \sin\chi_1\sin\chi_3 & \sin\chi_1\cos\chi_3 & \cos\chi_1 \\ 0 & 0 & 1 \end{pmatrix}, \tag{309}$$

and the angular momenta $\ell_i := I_i\omega_i$ (not summed). Comparing Eq. (305) with Eq. (308) reveals that $D^T = L^{-1}$ with T indicating transpose.

The inverse of Eq. (308) is given by

$$\ell_j = L_{ji}(\chi)p_i, \tag{310}$$

where

$$(L) = \frac{1}{\sin\chi_1} \begin{pmatrix} \sin\chi_1\cos\chi_3 & \sin\chi_3 & -\sin\chi_3\cos\chi_1 \\ -\sin\chi_1\sin\chi_3 & \cos\chi_3 & -\cos\chi_3\cos\chi_1 \\ 0 & 0 & \sin\chi_1 \end{pmatrix}. \tag{311}$$

Observe that Eq. (310) is of the form of the standard reduction formula (289).

Upon Legendre transformation, the Hamiltonian is obtained:

$$H(p,\chi) = p_i\dot{\chi}_i - L = \frac{1}{2}\ell_k\omega_k = \frac{1}{2}\sum_k \frac{\ell_k^2}{I_k}$$

$$= \frac{1}{2}\sum_k \frac{1}{I_k} L_{ki}L_{kj}p_ip_j, \tag{312}$$

which obviously possesses the necessary symmetry of Eq. (288).

It remains to show explicitly that the variables ℓ_i allow a reduction. To see this consider $[\ell_i, \ell_j]$, which upon insertion of Eq. (310) becomes

$$[\ell_i, \ell_j] = L_{rs}^{-1}\ell_s\left(\frac{\partial L_{ir}}{\partial \chi_k}L_{jk} - \frac{\partial L_{jr}}{\partial \chi_k}L_{ik}\right), \tag{313}$$

as expected from the results of the previous subsection. Since the right-hand side of Eq. (313) is difficult to evaluate, we make use of

$$\frac{\partial L_{ir}}{\partial \chi_k}L_{rs}^{-1} = -L_{ir}\frac{\partial L_{rs}^{-1}}{\partial \chi_k}, \tag{314}$$

which follows upon differentiating

$$L_{ir}L_{rs}^{-1} = \delta_{is}, \tag{315}$$

to obtain

$$[\ell_i, \ell_j] = \left(\ell_s\frac{\partial L_{rs}^{-1}}{\partial \chi_k}\right)(L_{jr}L_{ik} - L_{ir}L_{jk}). \tag{316}$$

The matrix in the first parentheses is not too difficult to calculate. The evaluation of the second parentheses amounts to the determination of three matrices; since $[\ell_i, \ell_j]$ is antisymmetric only $[\ell_1, \ell_2]$, $[\ell_1, \ell_3]$, and $[\ell_2, \ell_3]$ must be obtained. Multiplying out the matrices of the two parentheses (three times) yields the following compact result:

$$[\ell_i, \ell_j] = -\epsilon_{ijk}\ell_k. \tag{317}$$

This result is to be expected: since Eq. (310) is of the form of Eq. (289), Eq. (317) should be of the form of Eq. (293).

This example demonstrates that reduction allows us to solve for the motion of the rigid body in two stages. First, Euler's equations can be solved to obtain the ℓ's, which do not contain information about the configuration of the rigid body. Second, the known ℓ's can be inserted into the left-hand side of Eq. (305) to yield three more equations to be solved for the Euler angles.

4. Reduction for the ideal fluid: Lagrangian to Eulerian variables

Now consider reduction for the ideal fluid, which amounts to the transformation from Lagrangian to Eulerian variables (Littlejohn, 1981; Morrison, 1981a; Marsden *et al.*, 1983). In the Lagrangian variable description of Sec. III we showed that the Hamiltonian

$$H[\pi,q] = \int_D \left[\frac{\pi^2}{2\rho_0} + \rho_0 U(s_0, \rho_0/\mathcal{J})\right]d^3a, \tag{318}$$

together with the canonical Poisson bracket

$$[F,G] = \int_D \left[\frac{\delta F}{\delta q}\cdot\frac{\delta G}{\delta \pi} - \frac{\delta G}{\delta q}\cdot\frac{\delta F}{\delta \pi}\right]d^3a, \tag{319}$$

produces the ideal fluid equations of motion. We first forget about the Hamiltonian and concentrate on what happens to the Poisson bracket when we change from (q,π), the Lagrangian canonically conjugate pair of variables, to (ρ,σ,M), the Eulerian noncanonical variables.

Recall from Sec. III that

$$\rho(r,t) = \int_D \rho_0(a)\delta(r - q(a,t))d^3a,$$

$$\sigma(r,t) = \int_D \sigma_0(a)\delta(r - q(a,t))d^3a,$$

$$M(r,t) = \int_D \pi(a,t)\delta(r - q(a,t))d^3a. \tag{320}$$

Clearly, from the above three relations we can calculate (ρ,σ,M) for a given displacement field q and a given momentum field π. The chain rule thus goes the way that is needed in order to calculate variations of

$$F[q,\pi] = \bar{F}[\rho,\sigma,M]. \tag{321}$$

In Eq. (321) we are supposing that F obtains its q and π dependence through some functional $\bar{F}$ of (ρ, σ, M). The functionals F and $\bar{F}$ are defined on different functions, which are themselves defined on different (spacelike) domains, a and r, respectively.

Consider the variation of F,

$$\delta F = \int_D \left[\frac{\delta F}{\delta q} \cdot \delta q + \frac{\delta F}{\delta \pi} \cdot \delta \pi \right] d^3 a$$

$$= \int_D \left[\frac{\delta F}{\delta \rho} \delta \rho + \frac{\delta F}{\delta \sigma} \delta \sigma + \frac{\delta F}{\delta M} \cdot \delta M \right] d^3 r. \qquad (322)$$

Note that the two domains of integration coincide, although the variables of integration have different names. Now we represent the functional derivatives with respect to the Lagrangian fields in terms of functional derivatives with respect to the Eulerian fields. This will allow us to express the bracket *entirely* in terms of the Eulerian fields. The variations of the Eulerian fields induced by a variation of the Lagrangian fields are

$$\delta \rho = - \int_D \rho_0(a) \nabla \delta(r - q) \cdot \delta q \; d^3 a,$$

$$\delta \sigma = - \int_D \sigma_0(a) \nabla \delta(r - q) \cdot \delta q \; d^3 a,$$

$$\delta M = \int_D [\delta \pi \, \delta(r - q) - \pi \nabla \delta(r - q) \cdot \delta q] \; d^3 a. \qquad (323)$$

Above (and below) the ∇ operator operates on the r dependence. Inserting Eq. (323) into the second equation of (322), interchanging the order of integration, and equating the coefficients of δq and $\delta \pi$ implies

$$\frac{\delta F}{\delta q} = - \int_D \left[\rho_0 \frac{\delta \bar{F}}{\delta \rho} + \sigma_0 \frac{\delta \bar{F}}{\delta \sigma} + \pi \cdot \frac{\delta \bar{F}}{\delta M} \right] \nabla \delta(r - q) d^3 r$$

$$= \int_D \left[\rho_0 \nabla \frac{\delta \bar{F}}{\delta \rho} + \sigma_0 \nabla \frac{\delta \bar{F}}{\delta \sigma} + \pi_i \nabla \frac{\delta \bar{F}}{\delta M_i} \right] \delta(r - q) d^3 r$$

$$\qquad (324)$$

and

$$\frac{\delta F}{\delta \pi} = \int_D \frac{\delta \bar{F}}{\delta M} \delta(r - q) d^3 r = \frac{\delta \bar{F}}{\delta M} \bigg|_{r = q} =: \frac{\delta \bar{F}}{\delta M'}, \qquad (325)$$

where the second equality of Eq. (324) is obtained upon integrating by parts and assuming that the boundary terms vanish. In the second equality of Eq. (325) we have added the prime to remind us that this quantity depends only upon the integration variable a. Inserting Eqs. (324) and (325), for both F and G, into Eq. (319) yields

$$\{\bar{F}, \bar{G}\} = \int_D \int_D \delta(r - q) \left\{ \rho_0 \nabla \frac{\delta \bar{F}}{\delta \rho} \cdot \frac{\delta \bar{G}}{\delta M'} \right.$$

$$+ \sigma_0 \nabla \frac{\delta \bar{F}}{\delta \sigma} \cdot \frac{\delta \bar{G}}{\delta M'} + \pi_i \frac{\partial}{\partial x^j} \frac{\delta \bar{F}}{\delta M_i} \frac{\delta \bar{G}}{\delta M'_j}$$

$$- \rho_0 \nabla \frac{\delta \bar{G}}{\delta \rho} \cdot \frac{\delta \bar{F}}{\delta M'} - \sigma_0 \nabla \frac{\delta \bar{G}}{\delta \sigma} \cdot \frac{\delta \bar{F}}{\delta M'}$$

$$\left. - \pi_i \frac{\partial}{\partial x^j} \frac{\delta \bar{G}}{\delta M_i} \frac{\delta \bar{F}}{\delta M'_j} \right\} d^3 r \, d^3 a. \qquad (326)$$

After interchanging the order of integration, we can carry out the integral over $d^3 a$:

$$\{F, G\} = - \int_D \left[M_i \left(\frac{\delta F}{\delta M_j} \frac{\partial}{\partial x^j} \frac{\delta G}{\delta M_i} - \frac{\delta G}{\delta M_j} \frac{\partial}{\partial x^j} \frac{\delta F}{\delta M_i} \right) \right.$$

$$+ \rho \left(\frac{\delta F}{\delta M} \cdot \nabla \frac{\delta G}{\delta \rho} - \frac{\delta G}{\delta M} \cdot \nabla \frac{\delta F}{\delta \rho} \right)$$

$$\left. + \sigma \left(\frac{\delta F}{\delta M} \cdot \nabla \frac{\delta G}{\delta \sigma} - \frac{\delta G}{\delta M} \cdot \nabla \frac{\delta F}{\delta \sigma} \right) \right] d^3 r. \qquad (327)$$

Equation (327) is the noncanonical bracket that was given in Sec. IV. It is a bracket expression written entirely in terms of Eulerian field variables and in terms of Eulerian functionals, that is, functionals that depend on Eulerian fields integrated over the Eulerian spatial domain. Thus we have effected the bracket or kinematic part of the reduction.

To complete the reduction we must write the Hamiltonian in terms of ρ, σ, and M. Reduction is not possible unless there exists an Eulerian Hamiltonian $\bar{H}$ that satisfies

$$H[q, \pi] = \bar{H}[\rho, \sigma, M], \qquad (328)$$

upon substitution of Eqs. (320). In general, functionals of the Lagrangian variables cannot be written entirely in terms of the Eulerian variables as in Eq. (328); however, for the ideal fluid the Hamiltonian functional that does the trick is of course

$$\bar{H}[\rho, \sigma, M] = \int_D \left[\frac{M^2}{2\rho} + \rho U(\rho, \sigma/\rho) \right] d^3 r. \qquad (329)$$

Clearly Eq. (329) becomes equal to Eq. (318) upon substitution of Eqs. (320). Therefore we have completed the second step needed for reduction.

In summary, by reduction we have obtained a Hamiltonian description of the ideal fluid entirely in terms of the noncanonical Eulerian variables (ρ, σ, M): the noncanonical Poisson bracket of Eq. (327) together with the Hamiltonian of Eq. (329) produces the fluid equations of motion in the Eulerian variables.

C. Clebsch variables

In this section we consider *Clebsch variables*. These are canonical variables that reduce to noncanonical vari-

ables where, as mentioned above, the noncanonical variables are bilinear in the momenta and configuration space coordinates. We shall use the term Clebsch to describe all such bilinear transformations for which there is a reduction. However, particular forms are of special interest. Below we consider finite systems, infinite systems, the semidirect product, and several examples of each, notably the Clebsch representation for the ideal fluid, whence the name Clebsch originates (Clebsch, 1857, 1859; see also Lamb, 1945, art. 167).

1. Clebsch variables for finite systems

It is well known that the three components of the angular momentum, $q \times p$, form a canonical realization; if one restricts phase-space functions to be functions of only these three variables, then the canonical Poisson bracket of two such functions produces another such function. This is just the closure condition discussed in the previous section. The resulting noncanonical Poisson bracket in this case, like that for the free rigid body, is that corresponding to SO(3).

We shall present the Clebsch reduction from an historical, if not logical, point of view. Suppose we have a noncanonical Lie-Poisson bracket of the form

$$[f,g] = w_k c_{ij}^k \frac{\partial f}{\partial w_i} \frac{\partial g}{\partial w_j}, \tag{330}$$

where c_{ij}^k are the structure constants for an arbitrary Lie algebra. We know from the previous section that a canonical Poisson bracket, with a transformation of the form of Eq. (289), reduces to this form. Now we turn things around and ask the question, can we inflate Eq. (330) and obtain other canonical descriptions? Here we have used the word inflation, since we are not talking about the canonical description of Sec. IV on the symplectic leaves, which would be a further "reduction." This inflation is in essence what Clebsch did for the ideal fluid: he found a set of variables that uniquely determines the usual physical fluid variables, but the inverse of his transformation does not exist. For this reason we say there are "gauge" conditions analogous to those for the vector potential in electromagnetism.

The following transformation, which is motivated by the angular momentum reduction described above, is a finite-dimensional generalization of Clebsch's transformation:

$$w_i = c_{ij}^k p_k q^j, \tag{331}$$

where all indices are summed on $1,2,\ldots,N$. The quantities w^i could be thought of as components of a generalized angular momentum. Given a canonical description in terms of the q^i and p_i,

$$\{f,g\} = \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i}, \tag{332}$$

the bracket in terms of w is obtained by a reduction. This can be seen upon substituting

$$\frac{\partial f}{\partial p_i} = \frac{\partial f}{\partial w_j} c_{jk}^i q^k,$$

$$\frac{\partial f}{\partial q^i} = \frac{\partial f}{\partial w_j} c_{ji}^k p_k \tag{333}$$

into Eq. (332)

$$\frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} = p_r q^t (c_{jt}^k c_{ik}^r - c_{it}^k c_{jk}^r) \frac{\partial f}{\partial w_i} \frac{\partial g}{\partial w_j}$$

$$= w_k c_{ij}^k \frac{\partial f}{\partial w_i} \frac{\partial g}{\partial w_j}, \tag{334}$$

where the last equality follows upon making use of the Jacobi identity for the structure constants, Eq. (244).

Given any noncanonical Lie-Poisson system in terms of w, one can obtain an inflated canonical system of equations in terms of the Clebsch q and p; if these equations are solved for $q(t)$ and $p(t)$, then the w constructed according to Eq. (331) solves the noncanonical system.

2. Clebsch variables for infinite systems

Here we shall be a bit formal and define things in somewhat general terms. To begin with, we shall denote by $\langle \, , \, \rangle$ a pairing between a vector space and its dual. We shall, for now, leave the particular form of this unspecified, but we have in the back of our mind an integration like that in Eq. (156). The second slot of $\langle \, , \, \rangle$ can be thought of as an infinite-dimensional analog of the finite-dimensional "up" indices, while the first slot is the analog of the "down" indices. We shall refer to elements of the second slot as belonging to a Lie algebra $\mathfrak{g}$ and those of the first slot, its dual, as belonging to $\mathfrak{g}^*$. Thus $\langle \, , \, \rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$. In general the pairing is not symmetric.

In terms of the pairing, noncanonical Lie-Poisson brackets have the compact form

$$\{F,G\} = \langle \chi, [F_\chi, G_\chi] \rangle, \tag{335}$$

where $[\, , \,]$ is a Lie algebra product, which takes $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, and we have introduced the shorthand

$$F_\chi := \frac{\delta F}{\delta \chi}, \quad G_\chi := \frac{\delta G}{\delta \chi}; \tag{336}$$

F_χ and G_χ are, of course, in $\mathfrak{g}$. We refer to $\{ \, , \, \}$ as the "outer" bracket and $[\, , \,]$ as the "inner" bracket.

Now we define the binary operator $[\, , \,]^\dagger$ as follows:

$$\langle \chi, [f,g] \rangle =: \langle [\chi,g]^\dagger, f \rangle, \tag{337}$$

where evidently $\chi \in \mathfrak{g}^*$, $g,f \in \mathfrak{g}$, and $[\, , \,]^\dagger : \mathfrak{g}^* \times \mathfrak{g} \to \mathfrak{g}^*$. The operator $[\, , \,]^\dagger$ is necessary for obtaining the equations of motion from a Lie-Poisson bracket. The bilinear Clebsch transformation analogous to Eq. (331) is given by

$$\chi = [\Pi, Q]^\dagger. \tag{338}$$

In order to effect the reduction, let us consider a variation of Eq. (338),

$$\delta\chi=[\delta\Pi,Q]^{\dagger}+[\Pi,\delta Q]^{\dagger}, \tag{339}$$

which is used to relate functional derivatives as

$$\delta F=\langle\delta\chi,F_{\chi}\rangle=\langle[\delta\Pi,Q]^{\dagger},F_{\chi}\rangle+\langle[\Pi,\delta Q]^{\dagger},F_{\chi}\rangle. \tag{340}$$

Manipulation of the second equality of Eq. (340) yields

$$\begin{aligned}
\delta F&=\langle\delta\Pi,[F_{\chi},Q]\rangle+\langle\Pi,[F_{\chi},\delta Q]\rangle\\
&=\langle\delta\Pi,[F_{\chi},Q]\rangle-\langle\Pi,[\delta Q,F_{\chi}]\rangle\\
&=\langle\delta\Pi,[F_{\chi},Q]\rangle-\langle[\Pi,F_{\chi}]^{\dagger},\delta Q\rangle,
\end{aligned} \tag{341}$$

where the antisymmetry of [,] and the definition of [,]† have been used. Upon comparing

$$\delta F=\langle\delta\Pi,F_{\Pi}\rangle+\langle F_{Q},\delta Q\rangle,$$

which is just the variation of any functional $F[Q,\Pi]$, with the last equality of Eq. (341), we obtain

$$F_{\Pi}=[F_{\chi},Q],\quad F_{Q}=-[\Pi,F_{\chi}]^{\dagger}. \tag{342}$$

The canonical bracket in terms of Q and Π can be written as

$$\{F,G\}=\langle F_{Q},G_{\Pi}\rangle-\langle G_{Q},F_{\Pi}\rangle. \tag{343}$$

Inserting Eq. (342) into Eq. (343) produces

$$\begin{aligned}
\{F,G\}&=-\langle[\Pi,F_{\chi}]^{\dagger},[G_{\chi},Q]\rangle+\langle[\Pi,G_{\chi}]^{\dagger},[F_{\chi},Q]\rangle\\
&=-\langle\Pi,[[G_{\chi},Q],F_{\chi}]+[[Q,F_{\chi}],G_{\chi}]\rangle\\
&=\langle\Pi,[[F_{\chi},G_{\chi}],Q]\rangle=\langle[\Pi,Q]^{\dagger},[F_{\chi},G_{\chi}]\rangle\\
&=\langle\chi,[F_{\chi},G_{\chi}]\rangle,
\end{aligned} \tag{344}$$

where use has been made of the Jacobi identity of [,].

Thus we have a quite general abstract construction for inflating Lie-Poisson brackets into canonical Poisson brackets. Let us consider some examples.

3. Fluid examples

Two examples from fluid mechanics are given: the first is the two-dimensional Euler equation, while the second is related to the three-dimensional ideal fluid.

a. Two-dimensional Euler equation

As observed above, the structure constants for the free-rigid-body noncanonical bracket are ϵ_{ijk}, which is completely antisymmetric. The structure operator for the 2D Euler noncanonical bracket, which was given in Sec. IV, shares this property; this is clear from the fgh identity of Eq. (198), from which we also observe that

$$[f,g]^{\dagger}=-[f,g]. \tag{345}$$

Here no distinction is made between the vector space and its dual. (We are only working at a formal mathematical level.) For this case the inner bracket is

$$[f,g]=\frac{\partial f}{\partial x}\frac{\partial g}{\partial y}-\frac{\partial f}{\partial y}\frac{\partial g}{\partial x}, \tag{346}$$

and the pairing is

$$\langle,\rangle=\int_{D}d^{2}r. \tag{347}$$

The Clebsch variables $Q(r,t)$ and $\Pi(r,t)$ are related to the scalar vorticity via

$$\omega(r,t)=[Q,\Pi], \tag{348}$$

which is the analog of the finite-dimensional formula of Eq. (331).

The reduction from the canonical variables defined by Eq. (348) to the 2D Euler bracket parallels exactly the calculation of the previous subsection. There are two ways to obtain the equations of motion for $Q(r,t)$ and $\Pi(r,t)$. One way is to insert Eq. (348) into the Hamiltonian $H[\omega]$ of Eq. (194) and then calculate

$$\frac{\partial Q}{\partial t}=\frac{\delta H}{\delta\Pi},\quad\frac{\partial\Pi}{\partial t}=-\frac{\delta H}{\delta Q}. \tag{349}$$

The other way is to insert Eq. (348) directly into the two sides of the equation of motion for ω, viz.,

$$\frac{\partial\omega}{\partial t}=-[\psi,\omega], \tag{350}$$

[cf. Eq. (192)] and then manipulate as follows:

$$\begin{aligned}
\frac{\partial\omega}{\partial t}&=\left[\frac{\partial Q}{\partial t},\Pi\right]+\left[Q,\frac{\partial\Pi}{\partial t}\right]-[\psi,\omega]=-[\psi,[Q,\Pi]]\\
&=[Q,[\Pi,\psi]]+[\Pi,[\psi,Q]],
\end{aligned} \tag{351}$$

where the Jacobi identity was used to obtain the last equality. From Eqs. (351) and (350) we obtain

$$\left[\frac{\partial\Pi}{\partial t}+[\psi,\Pi],Q\right]+\left[\Pi,\frac{\partial Q}{\partial t}+[\psi,Q]\right]=0, \tag{352}$$

which is satisfied if

$$\begin{aligned}
\frac{\partial\Pi}{\partial t}&=-[\psi,\Pi]+\frac{\partial\Upsilon}{\partial Q},\\
\frac{\partial Q}{\partial t}&=-[\psi,Q]-\frac{\partial\Upsilon}{\partial\Pi},
\end{aligned} \tag{353}$$

for arbitrary functions $\Upsilon(Q,\Pi)$.[30] Solutions of Eq. (353) for any Υ can be used to construct solutions of Euler's equation (350). Given an initial vorticity $\omega(r,0)$, one need only find an initial Q and Π such that $\omega(r,0)=[Q(r,0),\Pi(r,0)]$. If $Q(r,t)$ and $\Pi(r,t)$ are solutions to Eq. (353), with the initial conditions $Q(r,0)$ and $\Pi(r,0)$, then $\omega(r,t)=[Q(r,t),\Pi(r,t)]$ constructed from these solutions is a solution of Eq. (350).

b. Three-dimensional fluid

Now consider the following bracket, which is a portion of the noncanonical bracket for the ideal fluid [cf. Eq. (209)]:

---

[30] The terms involving Υ can be eliminated by a time-dependent gauge transformation, where by gauge transformation we mean a transformation $Q=Q(\bar{Q},\bar{\Pi},t)$ and $\Pi=\Pi(\bar{Q},\bar{\Pi},t)$ that does not change ω as given by $\omega=[Q,\Pi]$.

$$\{F,G\} = \int_D M_i \left(\frac{\delta F}{\delta M_j} \frac{\partial}{\partial x^j} \frac{\delta G}{\delta M_i} - \frac{\delta G}{\delta M_j} \frac{\partial}{\partial x^j} \frac{\delta F}{\delta M_i} \right) d^3 r$$

$$=: \langle M, [F_M, G_M] \rangle. \tag{354}$$

It is obvious that this bracket will satisfy the Jacobi identity if Eq. (209) does. The inner bracket in this case is given by

$$[f,g]^i = f^j \frac{\partial g^i}{\partial x^j} - g^j \frac{\partial f^i}{\partial x^j}, \tag{355}$$

where, evidently, f and g now have three components. Integration by parts and neglect of surface terms yields

$$[\chi,g]_i^\dagger = \chi_j \frac{\partial g^j}{\partial x^i} + \frac{\partial(\chi_i g^j)}{\partial x^j}, \tag{356}$$

whence the Clebsch variables are seen to be related to M by

$$M_i = \Pi_j \frac{\partial Q^j}{\partial x^i} + \frac{\partial(\Pi_i Q^j)}{\partial x^j}. \tag{357}$$

In reality the decomposition above is not quite that due to Clebsch, whose transformation did not have the second term of Eq. (357).[31] It is an interesting fact the reduction occurs for other transformations besides Eq. (357). It works if the last term of Eq. (357) is dropped, Clebsch's original decomposition, and it works if the last term of Eq. (357) has the opposite sign. Also, some velocity fields do not need three Q's and Π's to be represented, i.e., it may only be necessary for the index i to take on one or two values. We shall expand a little on this below in the last subsection of this lecture.

4. Semidirect product reductions

The semidirect product is an example of an *extension*, a group-theoretic notion for making bigger groups out of a given group. We shall not discuss this in detail here, so the interested reader is referred to the references (Sudarshan and Mukunda, 1974; Holm and Kupershmidt, 1983; Marsden and Morrison, 1984). However, this notion makes its way from Lie groups to Lie algebras, and thus to Lie-Poisson brackets. We shall briefly discuss Lie-Poisson brackets that have inner algebras that are extensions.

Suppose the functional F in Eq. (340) has, in addition to its χ dependence, dependence on Π, i.e., $\bar{F}[\chi,\Pi] = F[Q,\Pi]$. (We have included the overbar now, as in Sec. III, to avoid confusion.) Invoking the chain rule with this additional dependence yields

$$F_\Pi = [\bar{F}_\chi, Q] + \bar{F}_\Pi, \quad F_Q = -[\Pi, \bar{F}_\chi]^\dagger, \tag{358}$$

[31]We note, however, that a transformation of the form of Eq. (357) is closely related to the Clebsch decomposition introduced for magnetohydrodynamics (MHD) (Zakharov and Kuznetsov, 1971; Morrison and Greene, 1980(E); Morrison, 1982).

which upon substitution into Eq. (343) produces instead of Eq. (344) the following:

$$\{\bar{F},\bar{G}\} = \langle \chi, [\bar{F}_\chi, \bar{G}_\chi] \rangle + \langle [\Pi, \bar{G}_\chi]^\dagger, \bar{F}_\Pi \rangle - \langle [\Pi, \bar{F}_\chi]^\dagger, \bar{G}_\Pi \rangle$$

$$= \langle \chi, [\bar{F}_\chi, \bar{G}_\chi] \rangle + \langle \Pi, [\bar{F}_\Pi, \bar{G}_\chi] - [\bar{G}_\Pi, \bar{F}_\chi] \rangle, \tag{359}$$

where the second equality follows from manipulations similar to those performed above.

Many systems possess brackets of the form of Eq. (359). The rigid body in a gravitational field is an example of finite dimension. An example of infinite dimension, which was given in the context of reduced magnetohydrodynamics (RMHD) (Morrison and Hazeltine, 1984; Zeitlin, 1992), but which also occurs in fluid mechanics, is the semidirect product extension of the noncanonical bracket for the 2D Euler fluid. For this example the bracket of Eq. (359) has an inner bracket given by Eq. (346) and a pairing, $\langle,\rangle$, given by Eq. (347).

5. Other Clebsch reductions: The ideal fluid

In this subsection we present some other forms of Clebsch reductions. The first is another way to reduce to the RMHD bracket mentioned above. This emphasizes the fact that reductions are not unique. Following this we show another way to reduce to a portion of the ideal-fluid bracket, also treated above. Finally we reduce to the complete ideal-fluid noncanonical bracket.

Suppose we have a system with canonical variables $[Q_i(r,t), \Pi_i(r,t)]$, in which $i=1,2$ and $r=(x,y)$. The canonical Poisson bracket is then

$$\{F,G\} = \int_D (F_Q \cdot G_\Pi - G_Q \cdot F_\Pi) d^2 r. \tag{360}$$

The following transformation gives a reduction:

$$\chi = [Q_1, \Pi_1] + [Q_2, \Pi_2],$$

$$\psi = [\Pi_1, \Pi_2], \tag{361}$$

where $[\,,\,]$ is given by Eq. (346). We leave it as an exercise to show via the chain rule that with Eq. (361), Eq. (360) reduces to a bracket of the form of Eq. (359).

Now consider the portion of the fluid bracket discussed above in Eq. (354), but instead of Eq. (357) we let

$$M = \Pi_i \nabla Q^i, \tag{362}$$

where $i=1,2,\ldots,N$ and N is arbitrary. We also leave it as an exercise to show via the chain rule that, with Eq. (362), a canonical bracket in terms of $[Q_i(r,t), \Pi_i(r,t)]$, where now $r=(x,y,z)$, reduces to a bracket of the form of Eq. (354).

Finally, suppose in addition to Eq. (362) that

$$\rho = -\Pi_1, \quad \sigma = -\Pi_2. \tag{363}$$

We leave it as a last exercise to show via the chain rule that with Eqs. (362) and (363), a canonical bracket in terms of $[Q_i(r,t), \Pi_i(r,t)]$ reduces to the ideal 3D fluid bracket of Eq. (209). One can choose N large enough to describe the velocity field of interest. Any velocity field can be represented with $N=3$, since Eq. (362) is just a

curvilinear coordinate decomposition. For nondivergent or irrotational velocity fields, fewer terms are needed.

VI. STABILITY AND HAMILTONIAN SYSTEMS

This section concerns notions of stability in Hamiltonian systems. In Sec. VI.A canonical systems are considered. Here, basic definitions are reviewed, energy arguments for stability are discussed, and the notion of a *negative-energy mode* (NEM) is introduced. An example is given of a gyroscopic system that illustrates properties of NEMs, in which context simple Hamiltonian bifurcation theory is reviewed. Finally in this subsection, these ideas are applied to the ideal fluid in the Lagrangian variable description. Section VI.B is concerned with stability in noncanonical Hamiltonian systems. The energy-Casimir method[32] is described and two examples are given: a modification of the rigid body and the 2D Euler equation. The examples exhibit a pathology related to the rank-changing behavior of the cosymplectic form, which is discussed. In Sec. VI.C the notion of *dynamical accessibility*, which can be used to make statements about stability, in spite of the rank-changing behavior, is introduced. Finally, it is shown how Eulerian variations, constrained by the condition of dynamical accessibility, lead to the same expression for the potential energy, $\delta^2 W$, as Lagrangian variations.

A. Stability and canonical Hamiltonian systems

Consider a dynamical system of the form

$$\dot{z}^i = V^i(z), \quad i = 1, 2, \ldots, M, \tag{364}$$

where, as in Sec. IV, we shall not get into what is required of $V(z)$ for existence and uniqueness of solutions, but just assume everything is all right. An *equilibrium point* z_e is a type of solution of Eq. (364) that satisfies $V(z_e) = 0$. Stability concerns the behavior of solutions near such equilibrium points. Roughly speaking, z_e is *stable* if solutions starting "close" to z_e at $t = 0$ remain close to z_e for all later times. This idea is formalized as follows:

The equilibrium point z_e is said to be stable if, for any neighborhood N of z_e there exists a subneighborhood $S \subset N$ of z_e such that if $z(t=0) \in S$ then $z(t) \in N$ for all time $t > 0$.

At first one might wonder why such a fancy definition is needed. Why introduce the subneighborhood? Why not just say that, if it starts in a set and stays in the set, then it is stable? The answer to this is illustrated in Fig. 16, which is the phase portrait for the simple harmonic oscillator.

In this figure the circles are surfaces of constant energy. Here we have chosen as a neighborhood N the

[32]See, for example, Holm *et al.* (1985), Morrison and Eliezer (1986), McIntyre and Shepherd (1987), Morrison and Kotschenreuther (1990), and Shepherd (1992).
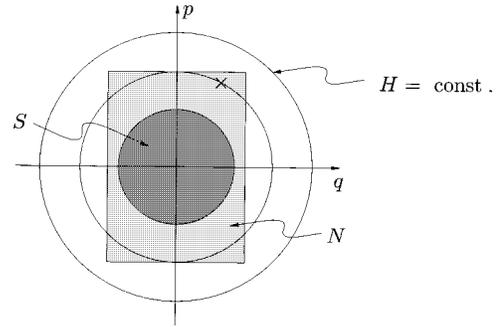
FIG. 16. Depiction of neighborhoods used in the definition of stability.

rectangular region in which we have marked an initial condition by the symbol $\times$. Since trajectories move round and round on the circles of constant H, it is clear that in a short time the trajectory starting at $\times$ will leave N, in spite of the fact that the equilibrium point at the origin is stable. However, if we choose initial conditions inside the subneighborhood S, which is defined as the region bounded by an $H = $ const surface contained in N, then the trajectory will remain in N for all time. Thus $H = $ const surfaces serve as a handy means of defining subneighborhoods.

Observe that the neighborhood N can be chosen to be *any* neighborhood N of z_e. We can make N smaller and smaller and in this way probe the stability property of the point z_e. In the example above we can always find tiny circular energy surfaces inside any N, no matter how small.

When $z(t)$ is determined from the linearized dynamics,

$$\delta \dot{z}^i = \frac{\partial V^i}{\partial z^j}(z_e) \delta z^j, \tag{365}$$

where now $z(t) \approx z_e + \delta z$, and this dynamics is stable according to the above definition, we say that Eq. (365) or z_e is *linearly stable*.

One might think, since N can be made as small as we like, that stability and linear stability are equivalent; but this is not the case, as we shall see below. To distinguish, we sometimes call stability under the full nonlinear dynamics, $V(z)$, *nonlinear stability*. Equilibria that are unstable under nonlinear dynamics, yet stable under linear dynamics, are said to be *nonlinearly unstable*. This is different from *finite-amplitude instability*, in which the equilibrium point is nonlinearly stable until it is pushed hard enough. In a sense (almost) all physical systems are finite-amplitude unstable; for example, any laboratory experiment is unstable to a perturbation caused by a large enough earthquake.

One last definition is that of *spectral stability*. A linear system such as Eq. (365) has this type of stability if, upon substituting $\delta z = \delta \hat{z} e^{i\omega t}$ and solving the resulting linear algebra problem for $\omega := \omega_R + i\gamma$, there exist no solutions with $\gamma < 0$. Clearly, linear stability implies spectral stability, but beware, the converse is not true.

A nice thing about Hamiltonian systems is that they

have a built-in method for proving nonlinear stability. When the Hamiltonian has a *separable* form, $H = p^2/2 + V(q)$, an old theorem due to Lagrange states that an equilibrium point with $p_e = 0$ and q_e being a local minimum of V is stable. It is tempting to think that the converse should be true, but a counterexample from the book of Wintner (1947) shows this not to be the case. Consider

$$V(q) = \begin{cases} e^{-1/q^2}\cos(1/q) & q \neq 0, \\ 0 & q = 0. \end{cases} \qquad (366)$$

The equilibrium position $q_e = 0$ is stable, but due to the wild oscillation that occurs as $q \to 0$, the origin is not a local minimum. However, with some relatively mild restrictions on V, *Lagrange's theorem* is both necessary and sufficient for stability for Hamiltonians of this restricted form. Sufficiency follows since surfaces of constant H serve to define subneighborhoods, as in the example of the simple harmonic oscillator above. Necessity is more difficult to see, but rests upon the idea that there exists a direction in which the trajectory can fall down to a state of lower potential energy.

For "well-behaved" $V(q)$, stability can be determined by analyzing the potential-energy matrix, $\partial^2 V(q_e)/\partial q^i \partial q^j$. If all the eigenvalues of this matrix are greater than zero, then H defines good subneighborhoods (topological $2N-1$ dimensional spheres) and the equilibrium is stable—in fact nonlinearly stable. If there exists a negative eigenvalue the system is unstable.

One might be fooled into thinking that nonlinear stability implies linear stability; however, with a little thought one can see that this is not true. The one-degree-of-freedom system with potential

$$V(q) = \frac{q^4}{4} \qquad (367)$$

has an equilibrium point $q_e = 0$, and it is clear that this is nonlinearly stable since H defines good subneighborhoods. However, the linear dynamics is governed by

$$\dot{\delta p} = 0, \quad \dot{\delta q} = \delta p \qquad (368)$$

and thus

$$\delta p = \delta p_0, \quad \delta q = \delta q_0 + \delta p_0 t. \qquad (369)$$

Obviously, trajectories leave any neighborhood of the equilibrium point provided $\delta p \neq 0$. This example also reveals why spectral stability does not imply linear stability. Adding another degree of freedom (q', p') and defining the potential $V(q, q') = q^4/4 + q'^2/2$ produces a linearly unstable, yet spectrally stable, system.

In the 1950s, project Matterhorn was begun at Princeton for the purpose of investigating controlled fusion reactions as a source of energy. The idea was (and still is) to confine hot plasmas by means of magnetic fields. Since the dominant force balance is governed by MHD, a great many stability analyses using this model were undertaken in a variety of confinement configurations invoking different magnetic-field geometries. What is in essence the infinite-degree-of-freedom version of Lagrange's theorem was worked out for MHD.[33] This goes by the name of the energy principle or "δW" (which is in fact the second variation of the potential energy). Extremization techniques applied to this quantity were used to determine stability and instability, and such procedures were automated in PEST, the Princeton Equilibrium and Stability Code, and elsewhere. Early MHD calculations were successful in explaining and eliminating the fastest plasma instabilities.

Often (as we shall see) Hamiltonian systems are not of the separable form $H(q, p) = p^2/2 + V(q)$, but are instead general functions of q and p. When this is the case another old theorem, which is sometimes called *Dirichlet's theorem*, gives a sufficient condition for stability (Dirichlet, 1846). It should be no surprise to you now that if, in the vicinity of an equilibrium point, surfaces of $H = \text{const}$ define a family of good neighborhoods, then the equilibrium is nonlinearly stable. For well-behaved Hamiltonians one need only analyze the matrix $\partial^2 H(z_e)/\partial z^i \partial z^j$, where $z := (q, p)$. If this matrix is definite, i.e., it has no zero eigenvalues and all of its eigenvalues have the same sign, then we have stability. Observe that H could in fact be an energy maximum. This can occur for rigid-body dynamics and is typically the case for a localized vortex in fluid mechanics.

There is an important example due to Cherry (1925) that illustrates two things: that Dirichlet's theorem is not necessary and sufficient and that linear stability does not imply nonlinear stability. Cherry's Hamiltonian is

$$H = \tfrac{1}{2}\omega_2(p_2^2 + q_2^2) - \tfrac{1}{2}\omega_1(p_1^2 + q_1^2) + \tfrac{1}{2}\alpha[2q_1 p_1 p_2 - q_2(q_1^2 - p_1^2)], \qquad (370)$$

where $\omega_{1,2} > 0$ and α are constants. If α is set to zero, Cherry's system reduces to a linear system of two stable simple harmonic oscillators. However, because of the minus sign, $\partial^2 H/\partial z_i \partial z_j$ is not definite. Observe that this minus sign cannot be removed by a time-independent canonical transformation and in the typical case cannot be removed by any canonical transformation. Oscillator *1* of this system is a *negative-energy mode* (NEM).

Negative-energy modes are important because, when dissipation is added, they tend to become linearly unstable: If energy is removed from an NEM its amplitude increases.[34] Also, with the inclusion of nonlinearity NEM's can be driven unstable. The example of Cherry demonstrates this; assuming $\alpha \neq 0$ and $\omega_2 = 2\omega_1$, we find that Eq. (370) possesses a solution[35] of the form

---

[33]There is an enormous literature on this subject. Early important papers are those of Hain *et al.* (1957), Bernstein *et al.* (1958), and Laval *et al.* (1965).

[34]This is a fairly old idea that is sometimes called the Kelvin-Tait theorem. See Thompson and Tait (1921), part 1, p. 388. An early application appears in Poincaré (1885).

[35]See Whittaker (1937), Sec. 136, p. 101. Be warned, for there are typographical errors.

$$q_1 = \frac{\sqrt{2}}{\epsilon - \alpha t} \sin(\omega_1 t + \gamma),$$

$$p_1 = -\frac{\sqrt{2}}{\epsilon - \alpha t} \cos(\omega_1 t + \gamma)$$

$$q_2 = -\frac{1}{\epsilon - \alpha t} \sin(2\omega_1 t + 2\gamma),$$

$$p_2 = -\frac{1}{\epsilon - \alpha t} \cos(2\omega_1 t + 2\gamma). \tag{371}$$

This is a two-parameter (α, γ) subfamily of the general four-parameter solution set of Cherry's system. These solutions are of interest since they can diverge in finite time. In fact, in any neighborhood of the equilibrium point $q_1 = q_2 = p_1 = p_2 = 0$ there exist initial conditions for solutions that diverge in finite time. Such behavior is referred to as explosive growth and is characteristic of systems that possess both NEMs and resonance. Another example is the well-known "three-wave" problem.[36] The three-wave problem and Cherry's "two-wave" problem are examples of systems with order three resonances that are driven unstable by cubic terms in the Hamiltonian. These are in fact normal forms that are obtained upon averaging a general class of Hamiltonians.[37] Thus explosive behavior is to be expected in systems with both positive- and negative-energy modes that are in resonance. When the resonance is detuned these systems generally are finite amplitude unstable and systems with three or more degrees of freedom may in fact be unstable, although with very small growth (Kueny, 1993).

1. Gyroscopic systems

One might think that systems with NEMs are artifacts or unphysical, purely mathematical, oddities; this, however, is not the case. They occur in fluid and plasma systems[38] for a reason that will become clear below. Generally, they occur in mechanical systems with gyroscopic forces, like the Coriolis force, and they occur in the dynamics of particles in magnetic fields. An example that exhibits both of these is described by a Lagrangian of the form

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + G(\dot{y}x - \dot{x}y) + \tfrac{1}{2}k(x^2 + y^2), \tag{372}$$

---

[36]See, for example, Weiland and Wilhelmsson (1977), Kueny (1993), and Kueny and Morrison (1995a; 1995b), and many references cited therein. Also, see Pfirsch (1993).

[37]The classification of normal forms for linear systems was accomplished by Williamson (1936). For nonlinear systems see Birkhoff (1927).

[38]In the context of fluids see Lamb (1945) Chaps. VIII and XII, Cairns (1979), MacKay and Saffman (1986), Ripa (1993); for MHD see Greene and Coppi (1965); for Vlasov theory see Morrison and Pfirsch (1989, 1990, 1992); and in general see Morrison and Kotschenreuther (1990).

where G is a constant that is either proportional to the constant angular speed of a rotating coordinate system or to a constant magnetic field. Note that for $k > 0$ the potential-energy term corresponds to a hill and thus without the gyroscopic term the system would be unstable. Upon Legendre transforming and scaling, the following Hamiltonian is obtained:

$$H = \tfrac{1}{2}(p_1^2 + p_2^2) + \omega_G(q_2 p_1 - q_1 p_2)$$
$$+ \tfrac{1}{2}(\omega_G^2 - \omega_k^2)(q_1^2 + q_2^2), \tag{373}$$

where the two time scales of the problem are determined by the frequencies

$$\omega_G := \frac{G}{m}, \qquad \omega_k := \sqrt{\frac{k}{m}}. \tag{374}$$

Assuming $q_{1,2}$, $p_{1,2} \sim e^{i\omega t}$, it is easy to solve for the eigenvalues,

$$\omega = \pm \omega_k(\sqrt{\epsilon - 1} \pm \sqrt{\epsilon}), \tag{375}$$

where $\epsilon := \omega_G^2 / \omega_k^2$. This system possesses the three types of Hamiltonian spectra:

(1) $\omega = \pm \omega_R$ stable,
(2) $\omega = \pm i\gamma$ unstable,
(3) $\omega = \pm \omega_R \pm i\omega_I$ unstable.

In Hamiltonian systems eigenvalues occur in doublets or quartets. Case (1) is the only stable case. It occurs in the example when $\epsilon = \omega_G^2 / \omega_k^2 > 1$, which means the rotation or magnetic field is large enough to make the system stable in spite of the destabilizing potential energy. In this case we have two stable doublets, a fast one and a slow one. The slow one is an NEM. For $\epsilon > 1$ there exists a canonical transformation $(q, p) \to (Q, P)$ that takes H into

$$H(Q, P) = -\tfrac{1}{2}\omega_s(P_s^2 + Q_s^2) + \tfrac{1}{2}\omega_f(P_f^2 + Q_f^2), \tag{376}$$

which is the linear part of Cherry's Hamiltonian. The canonical transformation is effected by the following mixed-variable generating function:

$$F_2(q_1, q_2, P_f, P_s) = c(q_1 P_s + q_2 P_f) + P_f P_s + \tfrac{1}{2}c^2 q_1 q_2, \tag{377}$$

where $c := [4(\omega_G^2 - \omega_k^2)]^{1/4}$.

Case (2) occurs if G is set to zero. There exist two unstable doublets, corresponding to the two directions for falling off the hill.

Case (3) occurs when $\epsilon < 1$. This case of the quartet obviously requires two degrees of freedom and is obviously unstable.

A nice feature of the above example is that it displays the two kinds of bifurcations that are generic to Hamiltonian systems. The first occurs when a doublet makes a transition between cases (1) and (2). There is a steady-state bifurcation where the frequencies go through the origin of the ω plane as shown in Fig. 17.

Here the stable pair is indicated by $\times$ and the unstable pair by $\otimes$. This bifurcation generally occurs in systems where the Hamiltonian is separable, i.e., $H = p^2/2 + V(q)$, for which Lagrange's theorem applies. It occurs in one-degree-of-freedom systems where the
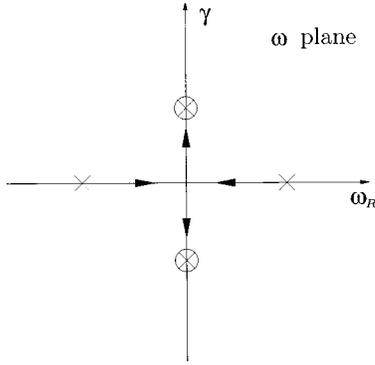
FIG. 17. The eigenvalue plane showing a bifurcation through the origin (zero frequency).

FIG. 18. The eigenvalue plane showing a bifurcation between eigenvalues of positive and negative signature.

potential goes from concave up to concave down. The arrows of the figure correspond to this case. For the system of Eq. (373) it occurs when $G=0$ and $\omega_k^2 \rightarrow - \omega_k^2$.

The other bifurcation, which is something called a *Kreĭn crash*, is illustrated in Fig. 18. The arrows indicate the path followed by the eigenvalues of system (373) as ε is decreased from some value greater than unity. At $\varepsilon=1$ the fast and slow modes coalesce at a value $|\omega_k|$ $\neq 0$. Two possibilities exist: either the modes go through each other and remain on the real axis or they can migrate off the real axis as shown in the figure. *Kreĭn's theorem*[39] states that a necessary condition for this latter case is that the signature of the colliding modes be different, i.e., one of them must be an NEM. The proof of Kreĭn's theorem is not difficult; it relies on the fact that definite Hamiltonians cannot have instabilities.

Kreĭn's theorem provides a means for detecting the occurrence of NEMs. If you have performed an eigenanalysis in some nondissipative system, one that you believe is Hamiltonian, and you observe the bifurcation described above, there must exist an NEM. This bifurcation is very common in fluid and plasma models. Why?

2. Ideal-fluid perturbation energy

To answer this question we return to the Hamiltonian formulation of the ideal fluid in terms of the Lagrangian variables q and π that we discussed in Sec. III. Since we have defined an equilibrium point of a dynamical system to be a solution obtained by setting time derivatives to zero, it is evident that the sets of Lagrangian and Eulerian equilibria are not equivalent. Although static Eulerian equilibria, i.e., ones for which $v=0$ for all r, certainly correspond to Lagrangian equilibria with $\pi=0$ and $q=$ const, stationary Eulerian equilibria, i.e., ones for which $v=v(r)$, do not correspond to Lagrangian equilibria, but to a particular kind of time-dependent trajectory, which we denote by

---

[39]Clear expositions are given by Moser (1958, 1968), who rediscovered the theorem. English translations of Kreĭn's original papers are given in Kreĭn and Jakubovič (1980).
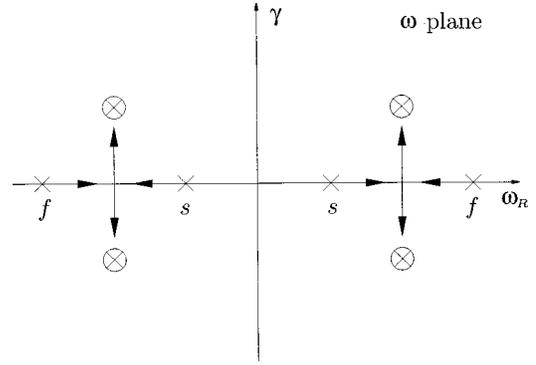
$$q_e=q_e(a,t), \quad \pi_e=\pi_e(a,t). \tag{378}$$

The functions above are particular in that they have the properties

$$\left.\frac{\rho_0(a)}{\mathcal{J}(a,t)}\right|_{a=q_e^{-1}(r,t)}=\rho_e(r), \tag{379}$$

$$s_0(a)|_{a=q_e^{-1}(r,t)}=s_e(r), \tag{380}$$

$$\left.\frac{\pi_e(a,t)}{\rho_0}\right|_{a=q_e^{-1}(r,t)}=\dot{q}_e(a,t)|_{a=q_e^{-1}(r,t)}=v_e(r), \tag{381}$$

where we emphasize that, upon doing the substitutions indicated on the left-hand sides of the above equations, the resulting functions ρ_e, s_e, and v_e are independent of time.

Although (q_e,π_e) does not constitute a Lagrangian equilibrium state, it is a reference state about which we can linearize. We set

$$q(a,t)=q_e(a,t)+\xi(a,t),$$
$$\pi(a,t)=\pi_e(a,t)+p_\xi(a,t) \tag{382}$$

and expand Eq. (120); however, the resulting equation will have explicit time dependence due to that in (q_e,π_e). Even when the time dependence is periodic, analysis of such linear equations is not trivial (recall Mathieu's equation).

We can get out of this bind of having to deal with explicit time dependence by a trick (see, for example, Frieman and Rotenberg, 1960, Newcomb, 1962, Van Kampen and Felderhof, 1967, and Pfirsch and Sudan, 1993). To see how this goes we turn to the action principle of Eq. (104), insert

$$q(a,t)=q_e(a,t)+\xi^{(1)}(a,t)+\xi^{(2)}(a,t)+\cdots,$$

and expand as follows:

$$S[q]=S[q_e]+\delta S[q_e;\xi^{(1)}]+\delta S[q_e;\xi^{(2)}]$$
$$+\delta^2 S[q_e;\xi^{(1)}]+\cdots . \tag{383}$$

The first term of Eq. (383) is merely a number, while the second term vanishes if the reference trajectory q_e is assumed to be a solution and if in addition $\xi^{(1)}$ is assumed to satisfy the boundary condition $\xi^{(1)}\cdot\hat{n}=0$,

where $\hat{n}$ is a unit vector normal to the boundary. The third and fourth terms can be combined to give an action principle for the linear dynamics relative to the reference state q_e.

Newcomb (1962) observed a subtlety associated with the derivation of this action principle. Usually action principles are designed so that surface terms, obtained from integration by parts, vanish upon variation. However, with the boundary condition $\xi^{(1)} \cdot \hat{n} = 0$ there are surface terms that come from the fourth term of Eq. (383) that do not vanish. To cure this malady it is necessary to extract a piece from $\delta S[q_e; \xi^{(2)}]$ that is a perfect divergence, and hence a surface term, and then add this to the third term. A "good" action principle is obtained by letting $\xi^{(2)} = (1/2)(\xi^{(1)} \cdot \nabla \xi^{(1)}) + \tilde{\xi}^{(2)}$ and gathering together all the terms that depend quadratically on $\xi^{(1)}$.

We denote this action principle, which gives the linear dynamics upon variation with respect to ξ, by $s[\xi]$, where for convenience we drop the superscript (1) on ξ. Using[40]

$$\delta^2 S[q_e; \xi] = \int_{t_0}^{t_1} dt \int_D d^3 a \left\{ \frac{\rho_0}{2} \dot{\xi}^2 - [(\partial_i \xi^j)(\partial_j \xi^i) \right.$$

$$+ (\partial_i \xi^i)(\partial_j \xi^j)] \frac{\rho_0^2}{2} \frac{U_\rho}{\mathcal{J}} \bigg|_{q_e}$$

$$\left. - (\partial_i \xi^i)(\partial_j \xi^j) \frac{\rho_0^3}{2} \frac{U_{\rho\rho}}{\mathcal{J}^2} \bigg|_{q_e} \right\}, \tag{384}$$

where $\partial_i \xi^j := \partial \xi^j / \partial q_e^i$, together with the surface term described above gives

$$s[\xi] = \int_{t_0}^{t_1} dt \int_D \mathcal{J} d^3 a \left[\frac{\rho}{2} \dot{\xi}^2 - (\partial_i \xi^i)^2 \rho_e \frac{\partial p_e}{\partial \rho_e} \right.$$

$$\left. - (\partial_i \xi^i)(\xi^j \partial_j p_e) + (\xi^i \partial_i \xi^j)(\partial_j p_e) \right], \tag{385}$$

Note that Eq. (385) can be obtained from Eq. (384) by integrating the term involving $(\partial_i \xi^j)(\partial_j \xi^i)$ by parts and blindly neglecting the surface term.

---

[40]The first term of Eq. (384) obviously comes from the kinetic energy; the other terms come from expanding the potential energy to second order:

$$\delta^2 W[\xi] = \int_D \left[\rho^2 U_\rho (\mathcal{J}_1^2/\mathcal{J} - \mathcal{J}_2) + \frac{\rho^3}{2} U_{\rho\rho} (\mathcal{J}_1^2/\mathcal{J}) \right] d^3 a,$$

where $\mathcal{J}_1$ and $\mathcal{J}_2$ are the first-and second-order pieces of the Jacobian, respectively, obtained upon inserting $q = q_e + \xi$ and expanding. They are given by

$$\mathcal{J}_1 = \frac{1}{2} \epsilon_{kjl} \epsilon^{imn} \frac{\partial \xi^k}{\partial a^i} \frac{\partial q_e^j}{\partial a^m} \frac{\partial q_e^l}{\partial a^n} = A_k^i \frac{\partial \xi^k}{\partial a^i} = \mathcal{J} \frac{\partial \xi^i}{\partial q_e^i}$$

and

$$\mathcal{J}_2 = \frac{1}{2} \epsilon_{jkl} \epsilon^{imn} \frac{\partial q_e^l}{\partial a^n} \frac{\partial \xi^k}{\partial a^i} \frac{\partial \xi^j}{\partial a^m} = \frac{\mathcal{J}}{2} \left(\frac{\partial \xi^i}{\partial q_e^i} \frac{\partial \xi^j}{\partial q_e^j} - \frac{\partial \xi^j}{\partial q_e^i} \frac{\partial \xi^i}{\partial q_e^j} \right).$$

It is important to observe that in Eq. (384), or Eq. (385), the terms involving U_ρ and $U_{\rho\rho}$, or p_e and ρ_e, possess the explicit time dependence arising from $q_e(a,t)$. The trick mentioned above is to view the perturbation of a trajectory in a frame of reference moving with the reference trajectory. This can be done since $q_e = q_e(a,t)$ is invertible. Thus we define

$$\eta(r,t) := \xi(a,t)|_{a = q_e^{-1}(r,t)}. \tag{386}$$

The quantity $\eta(r,t)$ is a sort of Eulerian field for the Lagrangian displacement variable. A time derivative of Eq. (386) yields

$$\dot{\xi}(a,t) = \frac{\partial \eta(r,t)}{\partial t} + \frac{\partial \eta(r,t)}{\partial r} \cdot \dot{q}_e \bigg|_{a = q_e^{-1}(r,t)} \tag{387}$$

or in light of Eq. (381)

$$\dot{\xi}(a,t) = \frac{\partial \eta(r,t)}{\partial t} + v_e(r) \cdot \nabla \eta(r,t). \tag{388}$$

Note that we have used $\cdot$ for time derivatives at constant a and $\partial/\partial t$ for time derivatives at constant r. Since in Eq. (388) $v_e(r)$, the equilibrium velocity, is time independent, no explicit time dependence is introduced by this transformation.

It is interesting and revealing to compare Eq. (388) with the transformation for time derivatives when going into a rotating frame of reference,

$$\frac{\partial}{\partial t} \bigg|_{\text{fixed}} = \frac{\partial}{\partial t} \bigg|_{\text{rot}} + \Omega \times. \tag{389}$$

Just as the second term of Eq. (389) gives rise to noninertial (or fictional) forces, notably the Coriolis force that gives rise to the gyroscopic term in the Hamiltonian of Eq. (373), the second term of Eq. (388) will give rise to a noninertial type of force in the fluid Hamiltonian. Transforming Eq. (385), using Eqs. (386) and (388), yields

$$s[\eta] = \frac{1}{2} \int_{t_0}^{t_1} dt \int_D d^3 r [\rho_e |\dot{\eta} + v_e \cdot \nabla \eta|^2 - \eta \cdot \mathfrak{V}_e \cdot \eta], \tag{390}$$

where $\mathfrak{V}_e$ is an operator, although one without explicit time dependence because it is now a function of the equilibrium quantities ρ_e and s_e. The second term of Eq. (390), the potential energy, can be written as

$$\delta^2 W'[\eta] := \frac{1}{2} \int_D d^3 r [\eta \cdot \mathfrak{V}_e \cdot \eta]$$

$$= \frac{1}{2} \int_D d^3 r \left[(\nabla \cdot \eta)^2 \rho_e \frac{\partial p_e}{\partial \rho_e} \right.$$

$$\left. + (\nabla \cdot \eta)(\eta \cdot \nabla p_e) - (\eta \cdot \nabla \eta) \cdot \nabla p_e \right], \tag{391}$$

where $p_e(\rho_e, s_e)$ is the equilibrium pressure expressed as a function of the equilibrium density and entropy.

We can now obtain the (time-independent) Hamil-

tonian by Legendre transformation. The canonical momentum is given by

$$p_\eta = \frac{\delta L}{\delta \dot{\eta}} = \rho_e(\dot{\eta} + v_e \cdot \nabla \eta), \tag{392}$$

whence the Hamiltonian is seen to be

$$\delta^2 H[p_\eta, \eta] = \frac{1}{2} \int_D d^3 r \left[\frac{p_\eta^2}{\rho_e} - 2p_\eta \cdot (v_e \cdot \nabla \eta) \right.$$

$$\left. + \eta \cdot \mathfrak{V}_e \cdot \eta \right], \tag{393}$$

which has the "noninertial" term $-p_{\eta i} \, v_e^j \, \partial \eta^i / \partial x^j$ that is reminiscent of the gyroscopic term of Eq. (373).

Now, it should come as no surprise that ideal fluid equilibria typically have negative-energy modes, and generally $\delta^2 H$ is not positive definite as required for Dirichlet's theorem. In spite of the indefiniteness of $\delta^2 H$ the system can be spectrally stable; Lagrange's theorem, which is a necessary and sufficient condition for stability, does not in general apply, since the Hamiltonian is not of the separable form.

B. Stability and noncanonical Hamiltonian systems

In noncanonical Hamiltonian systems it is still the case that equilibria occur at extremal points of the Hamiltonian,

$$\dot{z}^i = J^{ij} \frac{\partial H}{\partial z^j} = [z^i, H] = 0, \tag{394}$$

but the situation is more complicated. To see that something is amiss, consider the variation of the energy for a barotropic fluid, where

$$H[\rho, v] = \int_D [\tfrac{1}{2} \rho v^2 + \rho U(\rho)] d^3 r; \tag{395}$$

namely,

$$\frac{\delta H}{\delta v} = \rho v,$$

$$\frac{\delta H}{\delta \rho} = \frac{v^2}{2} + U(\rho) + \rho U_\rho(\rho). \tag{396}$$

Setting the right-hand side of Eq. (396) to zero results in the trivial equilibrium state with $v = 0$ and $\rho = $ const.[41] If this were the only equilibrium state, fluid mechanics would not be a very interesting discipline. Where are the other equilibria? Why are they not extremal points of the Hamiltonian? Since Dirichlet's stability theorem uses the fact that equilibria are extremal points of the Hamiltonian, is there a noncanonical analog of this theorem?

----

[41]The term equilibrium is used here to mean both static and stationary equilibria.

1. General formulation

To answer these questions, compare Eq. (394) with its counterpart for the canonical case:

$$\dot{z}^i = J_c^{ij} \frac{\partial H}{\partial z^j} = 0. \tag{397}$$

Since $\det J_c = 1$, it is evident that $\dot{z} = 0$ implies $\partial H / \partial z^j = 0$. Thus all equilibria are extremal points of the Hamiltonian. However, in the noncanonical case this is not so when $\det J = 0$. In the vicinity of points where the rank of J does not change, the null space of J is spanned by $\partial C^\alpha / \partial z^i$, $\alpha = 1, 2, \ldots, \nu$, where ν is the co-rank of J. In this case the general solution to Eq. (397) is given by

$$\left. \frac{\partial F}{\partial z^i} \right|_{z_e} := \left. \frac{\partial H}{\partial z^i} \right|_{z_e} + \lambda_\alpha \left. \frac{\partial C^\alpha}{\partial z^i} \right|_{z_e} = 0. \tag{398}$$

Here λ_α are Lagrange multipliers, which are determined by choosing the values of the constants of motion C^α. Thus Eq. (398) gives those equilibria that lie on the symplectic leaf with the chosen values. Later we shall see how $F := H + \lambda_\alpha C^\alpha$ is related to the energy.

Not surprisingly, the linear dynamics obtained by setting $z = z_e + \delta z$ and expanding to first order exhibit behavior arising from $\det J = 0$, namely, the existence of zero-frequency modes (Morrison and Eliezer, 1986). The equation for the linear dynamics is easily seen to be

$$\delta \dot{z}^i = A_k^i(z_e) \delta z^k, \tag{399}$$

where

$$A_k^i(z_e) := J^{ij}(z_e) \frac{\partial^2 F(z_e)}{\partial z^j \partial z^k} =: J_e^{ij} F_{,jk}. \tag{400}$$

Note, this linear dynamics has a Hamiltonian structure with the Poisson bracket defined by J_e (which is constant) and the Hamiltonian given by

$$\delta^2 F := \tfrac{1}{2} F_{,jk} \delta z^j \delta z^k. \tag{401}$$

Upon assuming $\delta z \sim e^{i\omega t}$, we obtain an eigenvalue problem with a characteristic equation given by

$$\det(i\omega I - A) = 0, \tag{402}$$

where, evidently, zero frequency modes satisfy

$$\det A = 0. \tag{403}$$

In the canonical case, A is given by

$$A_{ck}^i = J_c^{ij} H_{,jk} \tag{404}$$

and

$$\det(A_{ck}^i) = \det(J_c^{ij}) \det(H_{,jk}) = \det(H_{,jk}). \tag{405}$$

Thus all the zero eigenvalues of A_c arise from $\det(H_{,jk}) = 0$. These zero eigenvalues correspond to (local) troughs in the energy surface.

In the noncanonical case zero eigenvalues can arise from two places, namely, $\det(J^{ij}) = 0$ and $\det(F_{,ij}) = 0$. An accounting of these zero eigenvalues is given by
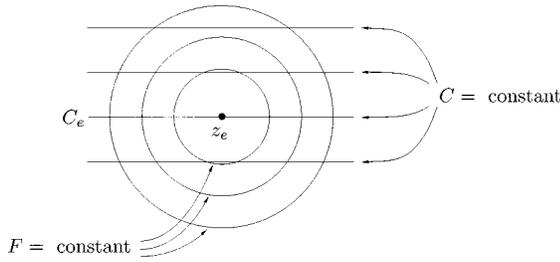
FIG. 19. Depiction of phase space, showing energy and Casimir surfaces, in the vicinity of an equilibrium point.

$$\text{Rank}(A^i_k) \leq \min\{ \text{Rank}(J^{ij}), \text{Rank}(F_{,jk})\}. \qquad (406)$$

Thus for every Casimir there exists a null eigenvector, which we denote by δz^k_0. To avoid complication suppose that $\det(F_{,jk}) \neq 0$, i.e., that there are no troughs in F; then all the null eigenvectors come from degeneracy in the bracket and they are given by

$$\delta z^k_0 = (F^{-1})^{,kj} \frac{\partial C(z_e)}{\partial z^j}, \qquad (407)$$

where $(F^{-1})^{,kj} F_{,jl} = \delta^k_l$. Evidently, with δz_0 given by Eq. (407),

$$A^j_k \delta z^k_0 = J^{ij} F_{,jk} (F^{-1})^{,kl} \frac{\partial C}{\partial z_l} = J^{ij} \frac{\partial C}{\partial z_j} = 0. \qquad (408)$$

In spite of the existence of null eigenvalues, a version of Dirichlet's theorem goes through in the noncanonical case. Since F is a constant of motion, it potentially could be used to define the subneighborhoods in the definition of stability given above. When $\delta^2 F = \frac{1}{2} F_{,jk} \delta z^j \delta z^k = $ const defines compact surfaces (as depicted in Fig. 19) in the vicinity of z_e, then z_e is a stable equilibrium point. This will be the case if $F_{,jk}(z_e)$ is definite.

It is of interest to note that this prescription for stability places no restrictions on δz, even though dynamically δz is confined to surfaces of constant C^α (as depicted in Fig. 19). We shall see in the next subsection that sometimes it is useful to take advantage of this information.

Although the picture described above for the equilibrium and stability of noncanonical Hamiltonian systems may seem nice and tidy, there is a complication that occurs at places where the rank of J changes. Generally, this happens at isolated points, but it can happen on curves or surfaces. When the rank changes it is no longer true that setting $\dot{z}^i = 0$ and solving for z_e is equivalent to solving Eq. (398) for all choices of λ_α. If the rank were to change and become smaller on an open set, but remain constant there, then there would be no problem in obtaining new Casimirs whose gradients span the null space of J on the open set. However, when the rank changes at (for example) a point, new null eigenvectors of J appear, eigenvectors that cannot be written as gradients in the normal way.

2. Examples

Now consider some examples that display how Dirichlet's theorem works for noncanonical systems, and the effect of the rank-changing pathology. We consider here two examples: a modification of the rigid body and the 2D Euler equation. We note, however, that examples that demonstrate the character of stability in noncanonical systems are not difficult to find: as is clear from Sec. IV, essentially all systems that describe fluids and plasmas in terms of Eulerian variables possess a description in this form, and the features present in the examples given here are typical.

a. Rigid body

We take up the free rigid body of Sec. IV, but with a Hamiltonian that has the following modified form (Kandrup and Morrison, 1993):

$$H = \sum_{i=1}^{3} \left(\frac{\ell^2_i}{2 I_i} + B_i \ell_i \right), \qquad (409)$$

where we have added the second term that is linear in ℓ_i. The quantities B_i are assumed to be constant and nonzero for $i = 1,2,3$. This Hamiltonian is a sort of mixture between that of a spin system and that of a free rigid body. This form serves our purpose and we shall not dwell on the physics, although it is not hard to imagine a physical system in which Hamiltonians of this form might arise. The equations of motion are now

$$\dot{\ell}_i = -\epsilon_{ijk} \ell_k \frac{\partial H}{\partial \ell_j} = -\epsilon_{ijk} \ell_k \left(\frac{\ell_j}{I_j} + B_j \right), \qquad (410)$$

and it is clear that equilibria must satisfy

$$\ell_1(I^{-1}_2 \ell_2 + B_2) - \ell_2(I^{-1}_1 \ell_1 + B_1) = 0,$$

$$\ell_1(I^{-1}_3 \ell_3 + B_3) - \ell_3(I^{-1}_1 \ell_1 + B_1) = 0,$$

$$\ell_2(I^{-1}_3 \ell_3 + B_3) - \ell_3(I^{-1}_2 \ell_2 + B_2) = 0. \qquad (411)$$

From Eqs. (411) it is also clear that a nonrotating configuration with $\ell_1 = \ell_2 = \ell_3 = 0$ is an equilibrium point, but there are other, uniformly rotating equilibria as well.

Now, consider the equilibria that arise upon extremizing $F = H + \lambda C$, where C is given by Eq. (165). (Note that the Casimir remains the same as in Sec. IV since we have not altered the bracket—only the Hamiltonian.) From $\partial F / \partial \ell_i = 0$ for $i = 1,2,3$, respectively, we obtain

$$\ell_1(I^{-1}_1 + \lambda) = -B_1,$$

$$\ell_2(I^{-1}_2 + \lambda) = -B_2,$$

$$\ell_3(I^{-1}_3 + \lambda) = -B_3. \qquad (412)$$

It is evident from Eqs. (412) that there exists no choice of λ for which the equilibrium point

$$\ell_1 = \ell_2 = \ell_3 = 0 \tag{413}$$

extremizes F. Observe, also, that the inequivalence of Eqs. (411) and (412) occurs for an equilibrium, namely Eq. (413), that corresponds to a point where $J^{ij} = -\epsilon_{ijk}\ell_k$ changes from rank 2 to rank 0.

Those equilibria that are extremal points of F can be tested by calculating $F_{,ij}$ to see if Dirichlet's theorem provides a sufficient condition for stability; the other equilibria for this example will be discussed in Sec. VI.C.

b. Two-dimensional Euler equation

The second example, the 2D Euler equation for fluid motion[42] (see Sec. IV), also has the property that the variation $\delta F = 0$ does not yield all equilibria. Because this example is infinite dimensional, there are some additional issues that arise, which we briefly comment on.

The equation of motion yields the equilibrium relation

$$\frac{\partial \omega}{\partial t} = [\omega, \psi] = 0, \tag{414}$$

which is satisfied if ω and ψ are functionally dependent: if $S = S(x,y)$ is an arbitrary function, then the equilibrium relation is satisfied if $\omega_e = \omega_e(S)$ and $\psi_e = \psi_e(S)$. Note that ω_e need not be a monotonic function of ψ_e and vice versa. Assuming $\omega_e = \omega_e(\psi_e)$ we can write

$$\nabla^2 \psi_e = \omega_e(\psi_e), \tag{415}$$

where $\omega_e(\psi_e)$ is an arbitrary, not necessarily monotonic, function of ψ_e.

Let us contrast the equilibrium equation (415) with the equation obtained upon varying the functional $F = H + C$, which for the 2D Euler equation is given by

$$F[\omega] = -\frac{1}{2}\int_D \psi\omega\, d^2r + \int_D \mathcal{C}(\omega)d^2r. \tag{416}$$

The functional derivative $\delta F/\delta\omega = 0$ implies

$$\psi_e = \mathcal{C}'(\omega_e). \tag{417}$$

Assuming $\mathcal{C}'(\omega)$ is monotonic we can solve for ω as follows:

$$\omega_e = \nabla^2 \psi_e = \mathcal{C}'^{-1}(\psi_e). \tag{418}$$

Thus here, in contrast to Eq. (415), the vorticity must be

a monotonic function of the stream function—if it is not, then it does not satisfy Eq. (417) and hence is not extremal.[43]

In order to apply Dirichlet's theorem it is necessary for the equilibrium to be extremal. When this is the case, as it is for the monotonic equilibria above, one can calculate the second variation,

$$\delta^2 F[\omega_e;\delta\omega] = \frac{1}{2}\int_D [|\nabla\,\delta\psi|^2 + \mathcal{C}''(\delta\omega)^2]d^2r$$

$$= \frac{1}{2}\int_D \Bigg[|\nabla\,\delta\psi|^2$$

$$+ (\delta\omega)^2\bigg(\frac{\partial\omega_e(\psi_e)}{\partial\psi_e}\bigg)^{-1}\Bigg]d^2r, \tag{419}$$

where the second equality follows upon differentiation of Eq. (417) with respect to ψ_e. Formally, if we have an equilibrium for which $\partial\omega_e(\psi_e)/\partial\psi_e > 0$, then $\delta^2 F$ is positive definite, and in analogy with finite-degree-of-freedom systems we could claim stability. In infinite dimensions, care must be taken in defining what is meant by a neighborhood: a natural definition is in terms of the "norm" defined by $\delta^2 F$.

It is also possible for $\delta^2 F$ to be positive definite when $\partial\omega_e(\psi_e)/\partial\psi_e < 0$; this would happen if the second term of Eq. (419) could be shown always to dominate the first when $\delta\omega$ lies in some space of functions. This case, which is typical of localized vortices, corresponds to an energy maximum.

In either case, the situation would be pretty good, but in infinite dimensions things can still be slippery. Recall that in Sec. III.A we gave an example of a functional with positive second variation at a point that was not a minimum. The condition of strong positivity is needed to show convexity. A rigorous stability analysis requires more than formal manipulations like those given above; in particular, it requires the definition of a Banach space in which the solution must be shown to exist. Convexity is one technical piece that is needed in a complete rigorous proof of stability.

If the first variation exists and does not vanish on the equilibrium of interest, then it is impossible for $F[\omega_e]$ to be convex and thus impossible to obtain a norm as discussed above. It can turn out, however, that the functional is not differentiable at the equilibrium of interest, but still can be proven to be stable by using F to obtain appropriate bounds (Rein, 1994). Another technique is to restrict the class of variations so that they lie within symplectic leaves, which we have mentioned above. In the next subsection we shall see how this removes problems related to the rank-changing behavior of J.

C. Dynamic accessibility

Here we discuss dynamically accessible variations, variations for which *all* the Casimir invariants are

---

[42]This example is credited to Arnold (1966a, 1966b, 1969), whose name is invoked in the popular terminology "Arnold's method" or "Arnold's theorem" for the application of these ideas to this and other situations. However, the method was used in earlier papers: Fjortoft (1950), Newcomb, in Appendix of Bernstein (1958), Kruskal and Oberman (1958), Fowler (1963), Gardner quoted in Fowler (1963), and Gardner (1963). A better terminology is the "energy-Casimir method."

---

[43]It is easy to prove this by contradiction. Suppose $\psi_0 \neq \psi_1$ and $\omega(\psi_0) = \omega(\psi_1) = \omega_*$. Then Eq. (417) implies $\psi_0 = \mathcal{C}'(\omega_*) = \psi_1$, which is the contradiction.
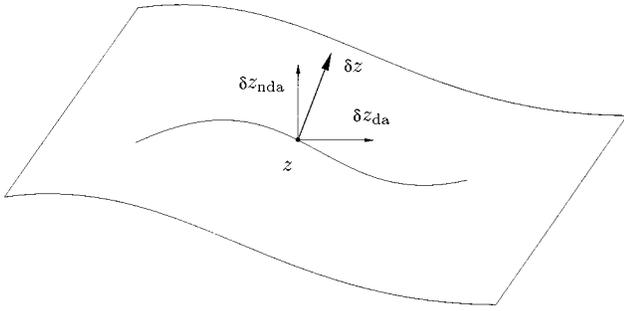
FIG. 20. Dynamically accessible and nonaccessible variations.

unchanged.[44] As depicted in Fig. 20, these variations lie in the surfaces defined by $C^\alpha = $ const for all α. In Sec. VI.C.1 we show how to generate these variations and how to use them to obtain the missing equilibria as extremals. We revisit the two examples of Sec. VI.B.2. In Sec. VI.C.2 we describe the analog of Dirichlet's theorem for dynamically accessible variations, and in Sec. VI.C.3 we treat the 3D ideal fluid; in particular, the second-order energy, constrained to dynamically accessible variations, is calculated and compared to the energy expression of Eq. (391).

1. General discussion

In the prescription described above for obtaining equilibria of noncanonical systems from a variational principle, the energy was extremized subject to a selection of Casimir invariants, and the values of these invariants were determined by the Lagrange multipliers (and vice versa). In contrast, dynamically accessible variations are *direct* variations, i.e., they have a form that automatically satisfies the constraints. Particular values of the Casimir constraints are not selected *a priori*, but are selected *a posteriori* by the choice of the equilibrium point, rather than by the choice of Lagrange multipliers.

---

[44]The essential idea here has been rediscovered many times in varied contexts. Gardner (1963) uses it in essence in his rearrangement argument for stability of Vlasov equilibria; it is (at least) implicit in Arnold's (1966b,1969) treatment of the 2D Euler equation; it is fully developed in the context of stellar dynamics as governed by the Vlasov (Jeans) equation by Bartholomew (1971). Filippov and Yan'kov (1986) use it in their discussion of electron vortices. In Morrison (1987), Morrison and Pfirsch (1989,1990,1992), and Morrison and Kotschenreuther (1990) it is described in general and used in a variety of plasma and fluid contexts; Nycander (1992) discusses it in the context of modon stability; and Zeitlin and Kambe (1993) use it in the context of 2D MHD. The discussion here follows the general development given in Kandrup and Morrison (1993), which also applies the idea in the context of the Vlasov-Einstein equation. The terminology "dynamically accessible" was introduced by Morrison and Pfirsch (1990) because such variations can be generated by inserting some Hamiltonian into the Poisson bracket, and thus these variations are accessible by some dynamics, while variations that leave the symplectic leaves cannot be generated this way.

Since the cosymplectic form, J^{ij}, projects (co)vectors onto the symplectic leaves, it is natural to use this object to construct dynamically accessible variations: such variations to first order are given by

$$\delta z_{da}^i = [\mathcal{G}, z^i] = J^{ji}(z)g_j, \qquad (420)$$

where $\mathcal{G} := z^i g_i$. Here the arbitrariness in the variation is embodied in the arbitrariness of the generating function g_j. However, because of the presence of J^{ij}, the variation δz_{da} is arbitrary only within the symplectic leaf; its components that are not tangent to the leaf are forced to vanish. Observe that J^{ij} can be evaluated at any point z, but in practice it will be evaluated at a candidate equilibrium point, which will be determined after setting the first variation to zero.

It is easy to see that dynamically accessible variations of the form of Eq. (420) automatically preserve the Casimir constraints to first order. Inserting Eq. (420) into the first variation of $C(z)$ yields

$$\delta C(z) = \frac{\partial C}{\partial z^i} \delta z_{da}^i = \frac{\partial C}{\partial z^i} J^{ji} g_j \equiv 0, \qquad (421)$$

where the last equality follows for any Casimir C, by its definition, and this equality is true for arbitrary g.

Let us return to the two examples of Sec. VI.B.2 and show how dynamically accessible variations yield *all* equilibria. The example of the rigid body with the modified Hamiltonian of Eq. (409) has dynamically accessible variations of the form

$$\delta \ell_i^{da} = \epsilon_{ijk} \ell_k g_j. \qquad (422)$$

Using this we obtain

$$\delta F = \frac{\partial F}{\partial \ell_i} \delta \ell_i^{da} = \frac{\partial H}{\partial \ell_i} \delta \ell_i^{da}$$

$$= (I_i^{-1} \ell_i + B_i) \epsilon_{ijk} \ell_k g_j^{(1)} = 0, \qquad (423)$$

for the extremal equilibrium condition. Equation (423) yields a result that is identical to Eq. (411), the equilibrium condition obtained upon setting $\dot{\ell}_i = 0$ in the equation of motion.

In the case of the 2D Euler fluid (Arnold, 1969)

$$\delta \omega_{da} = \{\mathcal{G}, \omega\} = -[g, \omega], \qquad (424)$$

where $\mathcal{G} := \int_D \omega g \, d^2 r$ with g arbitrary, and

$$\delta F_{da} := \delta F[\omega; \delta \omega_{da}] = -\int_D \psi \delta \omega_{da} d^2 r$$

$$= \int_D \psi [g, \omega] d^2 r = -\int_D g[\psi, \omega] d^2 r = 0, \qquad (425)$$

which implies $[\psi, \omega] = 0$—the condition obtained upon setting $\partial \omega / \partial t = 0$ in Eq. (414).

Although it is clearly nice how dynamically accessible variations produce all equilibria, before proceeding to a stability analysis, which we discuss in Sec. VI.C.2, it makes sense to stop and consider whether or not it is appropriate to restrict to such variations. After all, dynamically accessible variations correspond physically

to a restricted class of perturbations of the equilibrium state. Whether or not the physics calls for such restriction may need to be determined on a case-by-case basis.

In some situations the restriction may be called for. This occurs if the equilibrium state is physically well prepared, i.e., it is known in precise mathematical terms, and it is also known that fluctuations or external agents that cause perturbations to the equilibrium state preserve the constraint of dynamical accessibility. For example, in the case of the 2D Euler equation, dynamical accessibility means that the only permissible dynamics are those which rearrange the vorticity. If no physical mechanism exists for violating this constraint, then dynamically accessible variations are appropriate. This would be the case if viscosity or other means of dissipation such as Ekman friction caused changes on a time scale much longer than that of interest.

In the situation where external agents or fluctuations produce variations that are not dynamically accessible, the first comment to be made is that one should reexamine the model equation to see if the constraint-breaking physics should be incorporated into the dynamics. However, there are physical situations where such incorporation is not necessary. For example, it may be that there is an initial shock to a system that produces a nondynamically accessible perturbation, but the subsequent dynamics lies within a Casimir constraint surface. When this is the case it may turn out that there is an *equivalent dynamics* that corresponds to a new equilibrium with a perturbation that *is* dynamically accessible.

Equivalent dynamics means that a nondynamically accessible variation δz, about an equilibrium point z_e, is equivalent to a dynamically accessible variation about a new equilibrium point z_e^*; i.e., $z_e + \delta z_{\mathrm{nda}} = z_e^* + \delta z_{\mathrm{da}}$. If linear stability is under consideration, then we can decompose the perturbation as follows: $\delta z = \delta z_{\mathrm{nda}} + \delta z_{\mathrm{da}}$, where δz_{nda} is equal to δz_0, a null eigenvector of the cosymplectic form (cf. Sec. VI.B.1 and Fig. 20). The new equilibrium point, to within linear order, is given by $z_e^* = z_e + \delta z_0$. Because of the existence of the null eigenvector there exists a family of neighboring equilibria, and thus a nondynamically accessible perturbation followed by dynamics at constant Casimirs can be converted into the situation described above. This procedure fails if the neighboring equilibrium has different stability properties from the original one, i.e., if the equilibrium is at criticality.

In order to investigate stability by a Dirichlet-like argument it is necessary to obtain an expression for the perturbation energy, which is a second-order quantity. Thus it is necessary to obtain expressions for dynamically accessible variations to (at least) second order. The following is the needed second-order expression:

$$\delta z_{\mathrm{da}}^i = J^{ij} g_j^{(2)} + \tfrac{1}{2} J^{il} \frac{\partial J^{ti}}{\partial z^l} g_t^{(1)} g_j^{(1)} . \qquad (426)$$

Here we have added the superscripts (1) and (2) to distinguish the first- and second-order generating functions.

A variation of the form of Eq. (426) preserves the Casimir invariants to second order. This can be seen upon inserting Eq. (426) into $\delta^2 C$, manipulating, and using the fact that $J^{ij} \partial C / \partial z^i = 0$.

When the Poisson bracket is of Lie-Poisson type, $J_{ij} = c_{ij}^k w_k$, the first- and second-order dynamically accessible variations have the forms

$$\delta w_i^{\mathrm{da}} = c_{ji}^k w_k g_{(1)}^j ,$$
$$\delta^2 w_i^{\mathrm{da}} = c_{ji}^k w_k g_{(2)}^j + \tfrac{1}{2} c_{ti}^l c_{jl}^k w_k g_{(1)}^t g_{(1)}^j . \qquad (427)$$

Here we are using the notation introduced in Sec. V.B.2.

Sometimes it is useful to use finite-amplitude dynamically accessible variations. For Lie-Poisson brackets a convenient form for this is given by

$$\hat{w}_i = e^{g^j c_{ji}^k} w_k , \qquad (428)$$

where $\Delta w := \hat{w} - w$ is a finite variation. Observe that expansion of both $g = g_{(1)} + g_{(2)} + \cdots$ and the exponential of Eq. (428) yields Eqs. (427) to second order. Thus it is clear now how to obtain expressions for dynamically accessible variations, analogous to Eqs. (427), to all orders. An expression similar to Eq. (428) exists for general noncanonical Poisson brackets.

Finite variations can be used for probing phase space to find the boundaries for finite-amplitude stability; i.e., one can use these variations to find the boundary of the neighborhood of an equilibrium point, for which the intersection of the Casimir constraint surface with the energy surface is compact. Also, finite leaf variations of the form of Eq. (428) can be used in infinite-dimensional systems for proving convexity (see Sec. III.A). [However, it should be noted that in infinite dimensions, in contrast to finite dimensions, there are mathematical issues regarding variations of the form of Eq. (428); not all variations can be represented in this form.]

2. Energy and stability: $\delta^2 F_{\mathrm{da}} \equiv \delta^2 H_{\mathrm{da}}$

Thus far we have argued that the positive definiteness of two different kinds of energy expressions is sufficient for stability: $\delta^2 H[q_e, p_e; \delta q, \delta p]$ of Sec. VI.A for canonical systems and $\delta^2 F[z_e; \delta z]$ of Sec. VI.B for noncanonical systems. In the latter case δz was arbitrary, but the equilibrium z_e needed to be one of those obtainable from $\delta F = \delta H + \lambda \delta C = 0$ for some C. Here we consider another expression, $\delta^2 H_{\mathrm{da}}$, which is obtained upon expanding the energy about *any* equilibrium, but with the expansion restricted to the dynamically accessible variations of Sec. VI.C.1. If $\delta^2 H_{\mathrm{da}}$ is positive definite, then we have stability of a restricted type (at least).

We describe the restricted expansion procedure here, revisit the rigid-body-like example, and then show that $\delta^2 H_{\mathrm{da}} \equiv \delta^2 F[z_e; \delta z_{\mathrm{da}}] =: \delta^2 F_{\mathrm{da}}$. This means that inserting first-order dynamically accessible variations into the unrestricted expansion $\delta^2 F$ is equivalent to the expansion of the energy H with *second-order* dynamically accessible variations. This is a good thing to know, since in practice it is easier to calculate $\delta^2 F_{\mathrm{da}}$.

Expanding H through second order and then restricting the variations to be dynamically accessible yields

$$\Delta^2 H_{da} := \frac{\partial H}{\partial z^i} \delta z^i_{da} + \frac{\partial H}{\partial z^i} \delta^2 z^i_{da} + \frac{1}{2} \frac{\partial^2 H}{\partial z^i \partial z^j} \delta z^i_{da} \delta z^j_{da}.$$

(429)

Inserting Eqs. (420) and (426) into Eq. (429), observing that the first term of Eq. (429) vanishes when evaluated on the equilibrium, and likewise for those terms involving $g^{(2)}$, results in

$$\delta^2 H_{da} = \frac{1}{2} \frac{\partial^2 H}{\partial z^i \partial z^j} J_e^{li} g_l^{(1)} J_e^{kj} g_k^{(1)} + \frac{1}{2} \frac{\partial H}{\partial z^i} \frac{\partial J_e^{ti}}{\partial z_e^l} J_e^{jl} g_t^{(1)} g_j^{(1)}.$$

(430)

Equation (430) represents the energy content of a dynamically accessible perturbation. If this quantity is positive definite for all $g^{(1)}$, then the system is stable against this type of perturbation.

Let us now proceed to the task of showing that $\delta^2 H_{da} \equiv \delta^2 F_{da}$. Decomposing a general first-order variation as

$$\delta z = \delta z_{da} + \delta z_{nda}$$

(431)

and inserting into $\delta^2 F$ yields

$$\delta^2 F = \delta^2 F_{da} + \delta^2 F_{nda},$$

(432)

where

$$\delta^2 F_{da} = \frac{1}{2} \left(\frac{\partial^2 H(z_e)}{\partial z^i \partial z^j} + \lambda_\alpha \frac{\partial^2 C^\alpha(z_e)}{\partial z^i \partial z^j} \right)$$
$$\times J^{li}(z_e) g_l^{(1)} J^{kj}(z_e) g_k^{(1)}.$$

(433)

It is always the case that $\delta^2 F_{da}$ depends only on the first-order g's.

We note that $\delta^2 F$ can be indefinite because of the presence of $\delta^2 F_{nda}$, even if $\delta^2 F_{da}$, which involves only variations of the form $J^{ij}(z_e) g_i^{(1)}$, is of definite sign. An example is given by the free-rigid-body-like system with the equilibrium

$$\ell_1^e = -\frac{B I_1}{\lambda I_1 + 1}, \qquad \ell_2^e = \ell_3^e = 0,$$

(434)

where we set $B_1 = B$ and $B_2 = B_3 = 0$. In this case

$$\delta^2 F = -\frac{B}{2 \ell_1^e} (\delta \ell_1)^2 + \frac{1}{2} \left(\frac{1}{I_2} - \frac{1}{I_1} \right) (\delta \ell_2)^2$$
$$+ \frac{1}{2} \left(\frac{1}{I_3} - \frac{1}{I_1} \right) (\delta \ell_3)^2.$$

(435)

If $I_1 < I_2 < I_3$, the last two terms are positive; however, the first term can have either sign. Dynamically accessible variations satisfy

$$\delta \ell_i^{da} = \epsilon_{ijk} \ell_j^e g_k = \epsilon_{ijk} \ell_j^e g_k;$$

(436)

hence $\delta \ell_1^{da} = 0$. Therefore $\delta^2 F_{da}$ is definite, even though $\delta^2 F$ need not be.[45] Observe, also, that the nondynamically accessible variation corresponds to the null eigenvector described above.

To see that in general $\delta^2 F_{da}$ is identical to $\delta^2 H_{da}$, we expand some Casimir C^α to second order about the equilibrium:

$$\Delta^2 C^\alpha = \frac{\partial C^\alpha}{\partial z^i} \delta z^i + \frac{\partial C^\alpha}{\partial z^i} \delta^2 z^i + \frac{1}{2} \frac{\partial^2 C^\alpha}{\partial z^i \partial z^j} \delta z^i \delta z^j, \quad (437)$$

which when restricted to the constraint surface reduces to

$$\Delta^2 C_{da}^\alpha = \frac{\partial C^\alpha}{\partial z^i} J_e^{li} g_l^{(1)} + \frac{\partial C^\alpha}{\partial z^i} J_e^{li} g_l^{(2)}$$

$$+ \frac{1}{2} \frac{\partial C^\alpha}{\partial z^i} \frac{\partial J_e^{ti}}{\partial z_0^l} J_e^{jl} g_t^{(1)} g_j^{(1)}$$

$$+ \frac{1}{2} \frac{\partial^2 C^\alpha}{\partial z^i \partial z^j} J_e^{li} g_l^{(1)} J_e^{kj} g_k^{(1)}.$$

(438)

The first and second terms in Eq. (438) clearly vanish because of

$$J^{ij} \frac{\partial C^\alpha}{\partial z^j} = 0.$$

(439)

By the definition of dynamical accessibility, the last two terms must cancel; thus, to second order $\Delta^2 C_{da}^\alpha$ vanishes identically. Indeed, one can realize Eq. (439) as a Taylor series about the equilibrium point z_e and observe that, since this equation holds for all z (at least in a neighborhood of z_e), each power of δz in the expansion

$$0 = J^{ij} \frac{\partial C^\alpha}{\partial z^j}$$

(440)

$$= J_e^{ij} \frac{\partial C^\alpha}{\partial z_e^j} + \delta z^l \left(\frac{\partial J_e^{ij}}{\partial z_e^l} \frac{\partial C^\alpha}{\partial z_e^j} + J_e^{ij} \frac{\partial^2 C^\alpha}{\partial z_e^l \partial z_e^j} \right) + \cdots$$

must vanish identically. The first term in Eq. (440) is clearly zero, while the vanishing of the second term, the one linear in δz^l, yields a desired relation

$$\frac{\partial J_e^{ij}}{\partial z_e^l} \frac{\partial C^\alpha}{\partial z_e^j} = -J_e^{ij} \frac{\partial^2 C^\alpha}{\partial z_e^l \partial z_e^j},$$

(441)

between the first and second partial derivatives of C^α. In this way it follows immediately that $\Delta^2 C_{da} \equiv 0$.

It is evident that the first term of Eq. (430) is the same as the first term of the free energy $\delta^2 F_{da}$ of Eq. (433), but in order to compare the second terms in these equations, one must use Eq. (441) and the equilibrium condition (398) involving the Lagrange multipliers. Indeed, by summing Eq. (441) over λ_α and then exploiting Eq. (398), we conclude that

---

[45]It is important to keep in mind that the existence of an indefinite conserved quadratic form does not imply instability. Dirichlet-like theorems give only a sufficient condition for stability; negative-energy modes provide the counterexample that precludes necessity (cf. Sec. VI.A).

$$\lambda_\alpha J_e^{ij} \frac{\partial^2 C^\alpha}{\partial z_e^l \partial z_e^j} = -\lambda_\alpha \frac{\partial J_e^{ij}}{\partial z_e^l}\frac{\partial C^\alpha}{\partial z_e^j} = \frac{\partial J_e^{ij}}{\partial z_e^l}\frac{\partial H}{\partial z_e^j}. \tag{442}$$

It thus follows, as was asserted, that $\delta^2 H_{da} \equiv \delta^2 F_{da}$.

Note that although the two energy expressions are equivalent, the derivation of the $\delta^2 F_{da}$ is valid only for those equilibria that are extremal points of $F = H + \lambda C$ for some C, while $\delta^2 H_{da}$ has no such restriction.

3. Dynamically accessible fluid energy: $\delta^2 H_{da}$

By now it should be clear that there exist three different kinds of energy expressions: $\delta^2 H$, $\delta^2 F$, and $\delta^2 H_{da}$. The expression $\delta^2 H[q_e, p_e; \delta q, \delta p]$ occurs in the canonical description and lives on the largest phase space, the expression $\delta^2 F[w_e; \delta w]$ occurs in the noncanonical description and lives on the reduced phase space, and the expression $\delta^2 H_{da}[w_e; g]$ occurs in the noncanonical description and, because it is constrained to symplectic leaves which are subsurfaces of the reduced phase space, lives on the smallest phase space. Below we compute $\delta^2 H_{da}$ for the general 3D compressible fluid. To this end we construct first- and second-order dynamically accessible variations, show how such first-order variations give rise to all fluid equilibria as extremals of the energy, calculate the second variations of the potential and kinetic energies, and thus obtain $\delta^2 H_{da}$. Lastly, we compare $\delta^2 H_{da}$ with $\delta^2 H$, the Lagrangian variable expression for the second-order fluid energy as calculated in Sec. VI.A.2, and briefly comment on its utility.

a. Dynamically accessible variations

To obtain dynamically accessible variations for the fluid, the functional

$$\mathcal{G} := \int_D (M \cdot \eta + h\rho + k\sigma)\, d^3 r \tag{443}$$

can be inserted into the bracket of Eq. (209). Here the arbitrariness of variation within the symplectic leaf is described by the functions of η, h, and k, which are free functions of position. Following are expressions for the first-order variations of the momentum density, mass density, and entropy per unit mass, respectively:

$$\delta M_{da} = \{\mathcal{G}, M\} = [M, \eta]^\dagger + \rho\nabla h + \sigma\nabla k,$$

$$\delta\rho_{da} = \{\mathcal{G}, \rho\} = \nabla\cdot\left(\rho\frac{\delta\mathcal{G}}{\delta M}\right) = \nabla\cdot(\rho\eta),$$

$$\delta\sigma_{da} = \{\mathcal{G}, \sigma\} = \nabla\cdot\left(\sigma\frac{\delta\mathcal{G}}{\delta M}\right) = \nabla\cdot(\sigma\eta). \tag{444}$$

Recall that $[M, \eta]_i^\dagger := M\cdot\partial_i\eta + \nabla\cdot(M_i\eta)$, where $\partial_i := \partial/\partial x^i$, satisfies the identities $\langle M, [f, g]\rangle = \langle[M, g]^\dagger, f\rangle = -\langle[M, f]^\dagger, g\rangle$, where $[\ ,\ ]$ is defined by Eq. (355). These identities require that one of $M\cdot\hat{n}$, $f\cdot\hat{n}$, and $g\cdot\hat{n}$ vanish on the boundary, where we recall that $\hat{n}$ is a vector normal to the boundary. We assume $M_e\cdot\hat{n}\equiv 0$ and $\eta\cdot\hat{n}\equiv 0$, and thus $\delta\rho_{da}$ and $\delta\sigma_{da}$ do not change the total mass and entropy.

Observe that the variations $\delta\rho_{da}$ and $\delta\sigma_{da}$ of Eq. (444) are identical to $\delta\rho$ and $\delta\sigma$ of Eq. (323) if we make the identification $\delta q = \xi(a, t) = \eta(r, t)$. Therefore, for these quantities, first-order dynamically accessible variations are precisely the same as Eulerian variations induced by underlying Lagrangian variable variations. Also, upon identifying $M = \pi/\mathcal{J}$ and comparing δM of Eq. (323) with δM_{da} of Eq. (444), we obtain

$$\delta\pi_i = -\pi\cdot\partial_i\eta - \rho_0\partial_i h - \sigma_0\partial_i k, \tag{445}$$

i.e., if $\delta\pi$ satisfies Eq. (445), then the δM obtained from Eq. (323) will be dynamically accessible. Because $\delta\pi$ is arbitrary on the Lagrangian level, not all Lagrangian variations will be of the form of Eq. (445), and thus not all induced Eulerian variations will be dynamically accessible.

In order to calculate the energy expression, we need the following expressions for dynamically accessible variations to second order:

$$\delta^2 M_{dai} = \tfrac{1}{2}\{\mathcal{G}, \{\mathcal{G}, M\}\}_i$$
$$= \tfrac{1}{2}\{[[M, \eta]^\dagger, \eta]_i^\dagger + \nabla\cdot(\rho\eta)\partial_i h$$
$$\quad + \nabla\cdot(\rho\eta\partial_i h) + \rho(\partial_j h)(\partial_i\eta_j)$$
$$\quad + \nabla\cdot(\sigma\eta)\partial_i k + \nabla\cdot(\sigma\eta\partial_i k)$$
$$\quad + \sigma(\partial_j k)(\partial_i\eta_j)\},$$

$$\delta^2\rho_{da} = \tfrac{1}{2}\{\mathcal{G}, \{\mathcal{G}, \rho\}\} = \tfrac{1}{2}\nabla\cdot[\eta\nabla\cdot(\rho\eta)],$$

$$\delta^2\sigma_{da} = \tfrac{1}{2}\{\mathcal{G}, \{\mathcal{G}, \sigma\}\} = \tfrac{1}{2}\nabla\cdot[\eta\nabla\cdot(\sigma\eta)]. \tag{446}$$

These second-order variations should include dependence upon second-order generating functions as in Sec. VI.C.1, but the second-order pieces do not contribute below, so we neglect them.

b. Equilibria

Now we show how all 3D fluid equilibria are gotten from the vanishing of the first-order dynamically accessible variation of the fluid energy,

$$H[\rho, \sigma, M] = \int_D\left(\frac{M^2}{2\rho} + \rho\widetilde{U}(\rho, \sigma)\right)d^3 r, \tag{447}$$

where we recall that $\widetilde{U}(\rho, \sigma)$ is the internal energy per unit mass. Here we have used the tilde to indicate that the dependence is upon ρ and σ instead of ρ and s. In terms of $\widetilde{U}(\rho, \sigma)$ the equation of state for the pressure is given by

$$\widetilde{p}(\rho, \sigma) = \rho^2\left(\frac{\partial\widetilde{U}}{\partial\rho} + \frac{\sigma}{\rho}\frac{\partial\widetilde{U}}{\partial\sigma}\right). \tag{448}$$

The first variation of H is

$$\delta H = \int_D\left[\frac{M}{\rho}\cdot\delta M - \frac{M^2}{2\rho^2}\delta\rho + \delta\rho\left(\widetilde{U} + \rho\frac{\partial\widetilde{U}}{\partial\rho}\right)\right.$$
$$\left. + \delta\sigma\left(\rho\frac{\partial\widetilde{U}}{\partial\sigma}\right)\right]d^3 r. \tag{449}$$

Inserting Eq. (444), integrating by parts, and using $M = \rho v$ yields

$$\delta H_{\mathrm{da}}[\eta, h, k] = \int_D (\eta_j[-M \cdot \partial_j v - \partial_i(\rho v_i v_j)$$

$$+ \rho \partial_j(v^2/2) - \rho \partial_j(\tilde{U} + \rho \tilde{U}_\rho)$$

$$- \sigma \partial_j(\rho \tilde{U}_\sigma)] - h[\nabla \cdot M]$$

$$- k[\nabla \cdot (\sigma M/\rho)]) d^3 r. \tag{450}$$

Since η, h, and k are arbitrary, $\delta H_{\mathrm{da}} \equiv 0$ implies that the [...] terms of Eq. (450) vanish individually. These are easily shown to imply the usual Eulerian equilibrium equations for a compressible fluid:

$$\partial_j(\rho v_i v_j) + \partial_i p = 0,$$

$$\partial_j(\rho v_j) = 0,$$

$$v_j \partial_j s = 0. \tag{451}$$

c. Potential energy

Now consider the expansion of the potential-energy functional,

$$W[\rho, \sigma] = \int_D \rho \tilde{U}(\rho, \sigma) d^3 r, \tag{452}$$

to second order. Upon Taylor expansion, we obtain

$$\delta^2 W = \frac{1}{2} \int_D [(\delta \rho)^2 (\rho \tilde{U}_{\rho\rho} + 2 \tilde{U}_\rho) + (\delta \sigma)^2 (\rho \tilde{U}_{\sigma\sigma})$$

$$+ 2(\delta \sigma \delta \rho)(\rho \tilde{U}_{\rho\sigma} + \tilde{U}_\sigma) + 2(\delta^2 \rho)(\rho \tilde{U}_\rho + \tilde{U})$$

$$+ 2(\delta^2 \sigma)(\rho \tilde{U}_\sigma)] d^3 r, \tag{453}$$

where subscripts denote partial differentiation. Inserting Eq. (444) and (446) into Eq. (453) creates a relatively complicated formula, one with several terms that look similar, but simplification is not obvious. What we have is a sort of integration-by-parts puzzle. We shall not give all the details here of a calculation that gets us to the desired end, but only a few "moves" and "landmarks." The first move is to integrate the second-order variations by parts. Next, the terms are grouped as follows:

$$\delta^2 W_{\mathrm{da}}[\eta] = \frac{1}{2} \int_D [(\nabla \cdot \eta)^2 (\rho^3 \tilde{U}_{\rho\rho} + 2\rho^2 \tilde{U}_\rho + \sigma^2 \rho \tilde{U}_{\sigma\sigma}$$

$$+ 2\sigma \rho^2 \tilde{U}_{\sigma\rho} + 2\sigma \rho \tilde{U}_\sigma)$$

$$+ (\nabla \cdot \eta)(\eta \cdot \nabla \sigma)(\sigma \rho \tilde{U}_{\sigma\sigma} + \rho^2 \tilde{U}_{\rho\sigma} + \rho \tilde{U}_\rho)$$

$$+ (\nabla \cdot \eta)(\eta \cdot \nabla \rho)(\rho^2 \tilde{U}_{\rho\rho} + 2\rho \tilde{U}_\rho$$

$$+ \sigma \rho \tilde{U}_{\rho\sigma} + \sigma \tilde{U}_\sigma)] d^3 r, \tag{454}$$

which upon making use of Eq. (448) can be put into the form

$$\delta^2 W_{\mathrm{da}}[\eta] = \frac{1}{2} \int_D [(\nabla \cdot \eta)^2 (\rho \tilde{p}_\rho + \sigma \tilde{p}_\sigma)$$

$$+ (\nabla \cdot \eta)(\eta \cdot \nabla \tilde{p})] d^3 r. \tag{455}$$

The definition $p(\rho, s) := \tilde{p}(\rho, \sigma)$ and the chain rule imply $\rho \tilde{p}_\rho + \sigma \tilde{p}_\sigma = \rho p_\rho$, which when used in Eq. (455) yields, finally,

$$\delta^2 W_{\mathrm{da}}[\eta] = \frac{1}{2} \int_D [(\nabla \cdot \eta)^2 (\rho p_\rho) + (\nabla \cdot \eta)(\eta \cdot \nabla p)] d^3 r. \tag{456}$$

This expression, when evaluated on $\rho = \rho_e$ and $\sigma = \sigma_e$, is precisely that calculated in the footnote following Eq. (384), which was obtained in the strictly Lagrangian variable context. This is to be expected, since the dynamically accessible variations of density and entropy are the same as those induced by Lagrangian variations.

d. Kinetic energy

Now consider the expansion of the kinetic-energy functional,

$$T[\rho, M] = \int_D \frac{M^2}{2\rho} d^3 r, \tag{457}$$

to second order. Upon Taylor expansion, we obtain

$$\delta^2 T = \int_D \left[\frac{M}{\rho} \cdot \delta^2 M - \frac{M^2}{2\rho^2} \delta^2 \rho + \frac{M^2}{2\rho^3} (\delta \rho)^2 \right.$$

$$\left. - \delta \rho \frac{M}{\rho^2} \cdot \delta M + \frac{(\delta M)^2}{2\rho} \right] d^3 r. \tag{458}$$

We shall reduce this expression to a familiar form by solving another integration-by-parts puzzle. Making use of $\delta M = \rho \delta v + v \delta \rho$, which is a valid identity whether or not the variations $\delta \rho$ and δM are dynamically accessible,[46] yields the simplified expression

$$\delta^2 T = \int_D \left[v \cdot \delta^2 M - \frac{v^2}{2} \delta^2 \rho + \frac{\rho}{2} (\delta v)^2 \right] d^3 r. \tag{459}$$

Now restricting ourselves to dynamically accessible variations by inserting Eqs. (444) and (446) into Eq. (459) yields

---

[46]Using Eq. (444) yields $\delta v_{\mathrm{da}i} = v \cdot \partial_i \eta + \eta \cdot \nabla v_i + \partial_i h + s \partial_i k$, which is different from the expression for δv in terms of the displacement: $\delta v = \dot{\eta} + v \cdot \nabla \eta - \eta \cdot \nabla v$.

$$\delta^2 T_{\mathrm{da}}[\eta,h,k] = \frac{1}{2}\int_D \Bigg\{ v\cdot[[M,\eta]^\dagger]^\dagger + \nabla\cdot(\rho\,\eta)(v\cdot\nabla h)$$

$$+ v_i\nabla\cdot(\rho\,\eta\partial_i h) + \rho(v\cdot\nabla\,\eta_i)(\partial_i h)$$

$$+ \nabla\cdot(\sigma\,\eta)(v\cdot\nabla k) + v_i\nabla\cdot(\sigma\,\eta\partial_i k)$$

$$+ \sigma(v\cdot\nabla\,\eta_i)(\partial_i k) - \frac{v^2}{2}\nabla\cdot[\,\eta\nabla\cdot(\rho\,\eta)]$$

$$+ \rho(\delta v_{\mathrm{da}})^2 \Bigg\} d^3 r. \tag{460}$$

Making use of the identity $\langle[\chi,g]^\dagger,f\rangle = \langle\chi,[f,g]\rangle$, effecting some cancellation, and rearranging, gives

$$\delta^2 T_{\mathrm{da}}[\eta,h,k] = \frac{1}{2}\int_D [\rho\,\delta v_{\mathrm{da}}\cdot([v,\eta]+\delta v_{\mathrm{da}})$$

$$+ \delta\rho_{\mathrm{da}}[v\cdot\nabla h + v\cdot(v\cdot\nabla\,\eta)]$$

$$+ \delta\sigma_{\mathrm{da}}(v\cdot\nabla k)]d^3 r. \tag{461}$$

Introducing

$$P_i := \rho v\cdot\partial_i\eta + \rho\partial_i h + \sigma\partial_i k, \tag{462}$$

which satisfies $\rho\,\delta v_{\mathrm{da}} = \rho\,\eta\cdot\nabla v + P$, we obtain from Eq. (461)

$$\delta^2 T_{\mathrm{da}}[\eta,h,k] = \frac{1}{2}\int_D \Bigg[\frac{P^2}{\rho} + P_i(v\cdot\nabla\,\eta_i + \eta\cdot\nabla v_i)$$

$$+ \rho(\eta\cdot\nabla v_i)(v\cdot\nabla\,\eta_i)$$

$$+ \delta\rho_{\mathrm{da}}[v\cdot\nabla h + v\cdot(v\cdot\nabla\,\eta)]$$

$$+ \delta\sigma_{\mathrm{da}}(v\cdot\nabla k)\Bigg]d^3 r. \tag{463}$$

After several integrations by parts and making use of the equilibrium relations of Eq. (451), we obtain our desired result,

$$\delta^2 T_{\mathrm{da}}[\eta,h,k] = \frac{1}{2}\int_D \Bigg[\frac{P^2}{\rho} + 2P\cdot(v\cdot\nabla\,\eta)$$

$$- (\eta\cdot\nabla\,\eta)\cdot\nabla p \Bigg]d^3 r. \tag{464}$$

e. Total energy

Adding Eqs. (456) and (464) gives the following expression for the total, dynamically accessible energy of a perturbation to an Eulerian equilibrium state:

$$\delta^2 H_{\mathrm{da}}[\eta,h,k] = \frac{1}{2}\int_D \Bigg[\frac{P^2}{\rho_e} + 2P\cdot(v_e\cdot\nabla\,\eta)$$

$$+ (\nabla\cdot\eta)^2(\rho_e p_{e\rho}) + (\nabla\cdot\eta)(\eta\cdot\nabla p_e)$$

$$- (\eta\cdot\nabla\,\eta)\cdot\nabla p_e \Bigg]d^3 r. \tag{465}$$

It is interesting to compare the energy $\delta^2 H_{\mathrm{da}}$ of Eq. (465) with its counterpart $\delta^2 H$ of Eq. (393), which was derived within the Lagrangian variable context. If we

observe that $P_i \equiv -\delta\pi_i/\mathcal{J}$, where $\delta\pi_i$ is given by Eq. (445), and identify P with $-p_\eta$, then these energy expressions are identified. However, as noted above there is a difference in that p_η at fixed η is completely arbitrary in the Lagrangian variable context, while the freedom of P at fixed η is constrained to be determined by Eq. (462), which contains only the two free functions h and k.

The energy expression $\delta^2 H_{\mathrm{da}}$ can be rewritten in many different ways by integration by parts and by making use of the equilibrium equations (451). We shall not pursue this further here, but it is clear how to obtain Dirichlet-like stability criteria: if for some equilibrium $\delta^2 H_{\mathrm{da}}[\eta,h,k]$ can be manipulated into a positive definite form, then we have stability against dynamically accessible perturbations. One can proceed with $\delta^2 H_{\mathrm{da}}$ in a manner similar to that used in MHD analyses of the stability of plasma devices (see, for example, Newcomb, 1960; Hazeltine and Meiss, 1992), where minimizations are carried out in order to ascertain whether or not the energy functional is positive definite.[47]

The sufficient conditions for stability that result from the procedure described above are equivalent to Rayleigh or Fjortoft-type criteria, which are usually stated as necessary conditions for instability. Criteria such as the Rayleigh inflection point criterion for fluid shear flow are in essence conditions for the positive definiteness of the energy. This explains why similar criteria occur in a variety of fluid and plasma contexts. The expression $\delta^2 H_{\mathrm{da}}$ is a quite general starting point for doing this kind of stability analysis for a wide variety of equilibria and dynamics.

VII. CONCLUSION

One thing that is clear from the above section, indeed from this entire article, is that there are many different ways of writing the equations that describe fluid phenomena. These different ways amount to the choice of variables employed, such as Lagrangian, Eulerian, or Clebsch, and the different variables give rise to the different kinds of Hamiltonian structure. Although one can argue that underlying it all is Newton's second law together with constraints, such as mass conservation or the constancy of entropy, the different ways of describing a fluid imply different phrasings of the constraints. This difference in phrasing gives rise to differences in the usual analyses of fluid phenomena, notably stability analyses. Generally speaking, when two things are different, each has advantages. In closing we explore this and related stability issues a bit.

One consequence of phrasing the constraints in different ways is the occurrence of the different kinds of energy expressions that we have described. The use of $\delta^2 F$ for obtaining stability criteria, the energy-Casimir method, is relatively simple, direct, and extendable to

---

[47]Sun and Finn (1987), Pfirsch (1990), and Isichenko (1997) have done calculations of this sort for MHD.

nonlinear stability analysis, but is limited to a restricted class of equilibrium states, viz., those that arise as extremals of $F = H + \lambda C$. However, the use of $\delta^2 H_{da}$ has no such restriction and it can be calculated for all equilibria, but this approach has difficulties that need to be overcome in a rigorous mathematical analysis.[48]

A consequence of the above is that the restricted class of equilibria obtained by extremizing $F = H + \lambda C$ is singled out. Thus it is natural to ask, is there something special about these equilibria? One thing for certain is that this class includes the dynamical equilibrium state that corresponds to thermodynamic equilibrium. For example, for the 3D compressible fluid, the equilibrium equations (396) yield the state of thermodynamic equilibrium corresponding to zero mean flow and uniform density. In fact the condition $\delta F = 0$, in the case where $v = 0$, can be made to be identical to the thermodynamic minimum-energy principle (Callen, 1960), and $\delta^2 F$ positive definite corresponds in this case to thermodynamic stability (Le Châtelier's principle). Similarly for Vlasov theory, $\delta F = 0$ can be made to yield the Maxwellian phase-space density. Generally speaking it appears that these special equilibria are more likely to be stable than other equilibria, and proofs of nonlinear stability for them are possible in many cases. Indeed for both the 2D Euler equation and for Vlasov theory the restricted class corresponds to equilibria where the vorticity and phase-space density are monotonic functions of the stream function and energy, respectively. These equilibria do not have negative-energy modes (NEMs).

Any of the various energy expressions can be used to determine the existence of NEMs, but additional information is required. In particular, if it is known by some means that the system is spectrally stable and that the energy functional, whichever type, is not positive definite, then the system can have a NEM. By definition, if the eigenfunction of a NEM is substituted into the energy functional, then one obtains a negative expression. Similarly, if the eigenfunction of a positive-energy mode is substituted, then one obtains a positive expression. In both these cases mode means a stable pure oscillation—the energy associated with an unstable eigenfunction necessarily is zero. Thus associated with an eigenfunction is a signature.

Signature is important in that it gives an indication of possible behavior, as illustrated by the examples of Secs. VI.A and VI.A.1. From many examples, both finite and infinite, one can infer a general two-part principle regarding systems with NEMs (see, for example, Morrison and Kotschenreuther, 1990): (i) ideal stable systems with both positive- and negative-energy modes tend to be nonlinearly unstable, and (ii) if an ideal system with a

NEM is perturbed by adding terms to the dynamical system that extract energy, then a NEM (which by definition is neutrally stable) becomes unstable. This principle is clearly not a theorem: additional invariants can preclude the existence of nonlinear instability, and one must be sure that the dissipation actually extracts energy from a NEM in order for it to be destabilized.

Systems with symmetry possess other invariants, constants of motion that commute with the particular Hamiltonian (unlike Casimir invariants, which commute with all Hamiltonians). Extremization of the energy with these constants held fixed, which can be achieved by using Lagrange multipliers, yields relative equilibria, i.e., equilibria in frames of reference generated by the invariants. For example, extremization of $H + \lambda \cdot P$, where P is the momentum, gives a state that is uniformly translating at the velocity λ. Alternatively, one can interpret the extremal points obtained from extremizing the Hamiltonian at fixed invariants as being the equilibrium of interest observed in a different frame of reference. If there exists any frame of reference in which the energy functional is definite, then both linear and, for finite systems, nonlinear instability are precluded. Further, one can analyze the energy with the variations restricted to lie within the surfaces defined by the invariants, in a manner similar to that used in Sec. VI.C where dynamical accessibility was discussed.[49] In general, if any combination of known invariants implies the existence of a family of compact invariant sets about an equilibrium, then from the definition of Sec. VI.A that equilibrium is nonlinearly stable.

The existence of different frames of reference also affects the second part of our two-part principle because dissipation is frame dependent. This is easily demonstrated by an example due to Poincaré (1885). Consider a particle subjected to gravity and confined within a clear, spherical, rotating bowl. Clearly the point at the bottom of the bowl is an equilibrium point in any frame of reference. If there is no friction between the particle and the bowl, then the particle will oscillate about the bottom point, whether or not the bowl is rotating. However, if friction is added (by, for example, applying hairspray), then a stationary observer who is unable to discern the bowl, will see a strange effect of the ''dissipation.'' Rather than seeing the friction damp the oscillation, the stationary observer will see the particle leave the vicinity of the bottom equilibrium point as it grabs onto the side of the bowl. Clearly care must taken when defining dissipation, since in this example the stationary observer sees energy being put into the particle, while a rotating observer sees the particle attempting to come to rest with the rotating bowl.

The energy-Casimir (Arnold's) method or any energy method for obtaining sufficient conditions for stability provides an incomplete analysis of a problem. It is incomplete because the indefiniteness of the energy func-

---

[48]Namely, to attempt a nonlinear dynamically accessible stability analysis one must define the class of functions that are representable by the exponential of Sec. VI.C.1, and this is difficult. Also, this class of functions is not closed under the usual operations, and so the taking of limits is difficult to define (see, for example, de la Llave, 1993).

---

[49]See, for example, Ilgisonis and Pastukhov (1996) and Ilgisonis (1996) for a recent use of this idea in the context of MHD.

tional implies either the existence of NEMs or instability, and, as noted above, there is no way in general of distinguishing from any of the energy expressions alone which is the case. More importantly, it is incomplete because stability is not the whole issue: a complete analysis would determine the dynamics given an initial condition. Thus to finish the linear problem one must find a complete set of eigenfunctions; i.e., a normal-mode analysis must be done.

This normal-mode analysis is rather difficult because usually fluid and plasma systems possess continuous eigenspectra, unlike the complete discrete spectrum that occurs in elementary quantum mechanics problems. In the Hamiltonian setting, finishing the linear problem amounts to finding a transformation to action angle variables, variables that are related to the eigenfunctions of the continuous spectrum. For systems that are noncanonical this transformation has two parts: canonizing and diagonalizing. One must find a set of canonical variables on the symplectic leaf and then one must find the canonical transformation that diagonalizes the Hamiltonian. This program has been carried out for the Vlasov system (Morrison and Pfirsch, 1992; Morrison and Shadwick, 1994) using Van Kampen singular eigenfunctions (see, for example, Van Kampen and Felderhoff, 1967). The necessary mathematical tools have been generalized (Balmforth and Morrison, 1995) and applied to inviscid shear flow with a broad class of fluid equilibrium flow profiles. This is a long story, one that will be reported elsewhere (Balmforth and Morrison, 1998).

ACKNOWLEDGMENTS

REFERENCES

Arnold, V. I., 1963, Usp. Mat. Nauk [Sov. Math. Usp.] **18**, 85.

Arnold, V. I., 1966a, Izv. Vyssh. Uchebn. Zaved. Mat. [Sov. Math. J.] **5**, 3.

Arnold, V. I., 1966b, Ann. Inst. Fourier **16**, 319.

Arnold, V. I. 1969, Usp. Mat. Nauk. [Sov. Math. Usp.] **24**, 225.

Arnold, V. I., 1978, *Mathematical Methods of Classical Mechanics* (Springer, Berlin).

Arnold, V. I., V. V. Kozlov, and A. I. Neishtadt, 1990, *Dynamics III: Mathematical Aspects of Classical and Celestial Mechanics* (Springer, Berlin).

Balmforth, N. J., and P. J. Morrison, 1995 (Institute for Fusion Studies, University of Texas, Austin), Report No. 692. See National Technical Information Service Document No. DE95008952INZ. Copies may be obtained from NTIS, 5285 Port Royal Rd., Springfield, VA 22161, http://www.ntis.gov.

Balmforth, N. J., and P. J. Morrison, 1998, *Mathematics of Atmosphere and Ocean Dynamics*, Proceedings of the Isaac Newton Institute Program, edited by I. Roulstone and J. Norburg (Cambrige University, Cambridge), in press.

Bartholomew, P., 1971, Mon. Not. R. Astron. Soc. **151**, 333.

Bernstein, I., 1958, Phys. Rev. **109**, 10.

Bernstein, I., E. A. Frieman, M. D. Kruskal, and R. M. Kulsrud, 1958, Proc. R. Soc. London, Ser. A **244**, 17.

Berry, M. V., 1978, in *Topics in Nonlinear Dynamics*, AIP Conference Proceedings No. 46, edited by S. Jorna (AIP, New York), p. 16.

Bialynicki-Birula, I., and P. J. Morrison, 1991, Phys. Lett. A **158**, 453.

Birkhoff, G. D., 1927, *Dynamical Systems* (American Mathematical Society, New York).

Bretherton, F. P., 1970, J. Fluid Mech. **44**, 19.

Cairns, R. A., 1979, J. Fluid Mech. **92**, 1.

Calkin, M., 1963, Can. J. Phys. **41**, 2241.

Callen, H. B., 1960, *Thermodynamics* (Wiley, New York).

Cherry, T. M., 1925, Trans. Cambridge Philos. Soc. **23**, 199.

Clebsch, A., 1857, J. Reine Angew. Math. **56**, 1.

Courant, R., and D. Hilbert, 1953, *Methods of Mathematical Physics* (Wiley Interscience, New York).

Clebsch, A., 1859, J. Reine Angew. Math. **54**, 293.

de la Llave, R., 1993, in *Proceedings II Granada Seminar in Computational Physics*, edited by P. L. Garrido and J. Marro (World Scientific, Singapore), pp. 73–106.

del-Castillo-Negrete, D., J. M. Greene, and P. J. Morrison, 1997, Physica D **100**, 311.

del-Castillo-Negrete, D., and P. J. Morrison, 1993, Phys. Fluids A **5**, 948.

Dirichlet, P. G. L., 1846, Crelle **32**, 3.

Dubrovin, B. A., and S. P. Novikov, 1984, Sov. Math. Dokl. **30**, 651.

Dzyaloshinskii, I. E. and G. E. Volovick, 1980, Ann. Phys. (N.Y.) **125**, 67.

Eckart, C., 1960, Phys. Fluids **3**, 421.

Eisenhart, L. P., 1961, *Continuous Groups of Transformations* (Dover, New York).

Falcolini, C., and R. de la Llave, 1992, J. Stat. Phys. **67**, 609.

Filippov, D. V., and V. V. Yan'kov, 1986, Sov. J. Plasma Phys. **12**, 548.

Fjortoft, R., 1950, Geofys. Publ. **17**, 1.

Fowler, K., 1963, J. Math. Phys. **4**, 559.

Frieman, E., and M. Rotenberg, 1960, Rev. Mod. Phys. **32**, 898.

Gardner, C. S., 1963, Phys. Fluids **6**, 839.

Gardner, C. S., 1971, J. Math. Phys. **12**, 1548.

Gelfand, I. M. and S. V. Fomin, 1963, *Calculus of Variations* (Prentice-Hall, Englewood Cliffs, NJ).

Greene, J. M., 1979, J. Math. Phys. **20**, 1183.

Greene, J. M., and B. Coppi, 1965, Phys. Fluids **8**, 1745.

Greenspan, H. P., 1968, *Theory of Rotating Fluids* (Cambridge University, Cambridge, England).

Gürsey, F., 1963, in *Relativity, Groups and Topology*, edited by C. DeWitt and B. DeWitt (Gordon Breach, New York), p. 91.

Hain, V. K., R. Lüst, and A. Schlüter, 1957, Z. Naturforsch. Teil A **12**, 833.

Hammermesh, M., 1962, *Group Theory and its Applications to Physical Problems* (Addison-Wesley, Reading, MA).

Hazeltine, R. D., and J. D. Meiss, 1992, *Plasma Confinement* (Addison-Wesley, Redwood City, CA).

Holm, D. D., and B. Kupershmidt, 1983, Physica D **6**, 347.

Holm, D. D., J. E. Marsden, T. Ratiu, and A. Weinstein, 1985, Phys. Rep. **123**, 1.

Ilgisonis, V. I., 1996, Phys. Plasmas **3**, 4577.

Ilgisonis, V. I., and V. P. Pastukhov, 1996, Plasma Phys. Rep. **22**, 209.

Isichenko, M., 1997, private communication.

Jacobson, N., 1962, *Lie Algebras* (Wiley Interscience, New York).

Kandrup, H. E., and P. J. Morrison, 1993, Ann. Phys. (N.Y.) **225**, 114.

Kreĭn, M. G., and V. A. Jakubovič, 1980, *Four Papers on Ordinary Differential Equations* (American Mathematical Society, Providence, Rhode Island).

Kruskal, M. D., and C. Oberman, 1958, Phys. Fluids **1**, 275.

Kueny, C., Ph.D. thesis, The University of Texas at Austin, Austin, Texas, 1993.

Kueny, C. S., and P. J. Morrison, P. J., 1995a, Phys. Plasmas **2**, 1926.

Kueny, C. S., and P. J. Morrison, 1995b, Phys. Plasmas **2**, 4149.

Kuznetsov, E. A., and A. Mikhailov, 1980, Phys. Lett. A **77**, 37.

Lagrange, J. L., 1788, *Mécanique Analytique* (Desaint, Paris).

Lamb, H., 1945, *Hydrodynamics* (Dover, New York).

Laval, G., C. Mercier, and R. Pellat, 1965, Nucl. Fusion **5**, 156.

Lax, P., 1975, Commun. Pure Appl. Math. **28**, 141.

Lewis, D., J. Marsden, R. Montgomery, and T. Ratiu, 1986, Physica D **18**, 391.

Likins, P. W., 1971, AGARD Lecture Series **3b**, 1.

Littlejohn, R., 1981, private communication.

Littlejohn, R., 1982, in *Mathematical Methods in Hydrodynamics and Integrability in Related Dynamical Systems*, AIP Conference Proceedings No. 88, edited by M. Tabor and Y. Treve (AIP, New York), p. 47.

Loewner, C., 1971, *Theory of Continuous Groups*, notes by H. Flanders and M. Protter (MIT Press, Cambridge, MA).

MacKay, R. S., 1982, Ph.D. thesis (Princeton University).

MacKay, R. S., and P. G. Saffman, 1986, Proc. R. Soc. London, Ser. A **406**, 115.

Magri, F., 1978, J. Math. Phys. **19**, 1156.

Marsden, J. E., 1993, *Lectures on Mechanics* (Cambridge University, Cambridge, England).

Marsden, J. E., and P. J. Morrison, 1984, Contemp. Math. **28**, 133.

Marsden, J. E., and A. Weinstein, 1974, Rep. Math. Phys. **5**, 121.

Marsden, J. E., and A. Weinstein, 1983, Physica D **7**, 305.

Marsden, J. E., A. Weinstein, T. Ratiu, R. Schmid, and R. Spencer, 1983, in *Proceedings of IUTAM-IS1MM Symposium on Modern Developments in Analytical Mechanics*, Atti della Accad. Sci. Torino **117**, 289.

McIntyre, M. E., and T. G. Shepherd, 1987, J. Fluid Mech. **181**, 527.

Meacham, S. P., G. R. Flierl, and P. J. Morrison, 1997, Phys. Fluids **9**, 2310.

Meiss, J. D., 1992, Rev. Mod. Phys. **64**, 795.

Montgomery, D., and R. Kraichnan, 1979, Rep. Prog. Phys. **43**, 35.

Morrison, P. J., 1980, Phys. Lett. A **80**, 383.

Morrison, P. J., 1981a, "Hamiltonian field description of two-dimensional vortex fluids and guiding center plasmas," Princeton University Plasma Physics Laboratory Report, PPPL-1783. Available as American Institute of Physics Document No. PAPS-PFBPE-04-771-24, AIP Auxiliary Publication Service, 500 Sunnyside Blvd., Woodbury, NY 11797-2999, http://www.aip.org/epaps/.

Morrison, P. J., 1981b, "Hamiltonian field description of One-Dimensional Poisson-Vlasov Equation," Princeton University Plasma Physics Laboratory Report, PPPL-1788. Available as American Institute of Physics Document No. PAPS-PFBPE-04-771-14, AIP Auxiliary Publication Service, 500 Sunnyside Blvd., Woodbury, NY 11797-2999, http://www.aip.org/epaps/.

Morrison, P. J., 1982, in *Mathematical Methods in Hydrodynamics and Integrability in Related Dynamical Systems*, AIP Conference Proceedings No. 88, edited by M. Tabor and Y. Treve (AIP, New York), p. 13.

Morrison, P. J., 1987, Z. Naturforsch. **42a**, 1115.

Morrison, P. J., 1994, "Hamiltonian Description of the Ideal Fluid," in *Geophysical Fluid Dynamics*, edited by R. Salmon and B. Ewing-Deremer (Woods Hole Oceanographic Institute, Woods Hole, 1994), Technical Report No. WHOI-94-12. See National Technical Information Service Document No. DE94007372INZ. Copies may be obtained from the NTIS, 5285 Port Royal Rd., Springfield, VA 22161; http://www.ntis.gov.

Morrison, P. J., and S. Eliezer, 1986, Phys. Rev. A **33**, 4205.

Morrison, P. J., and J. M. Greene, 1980, Phys. Rev. Lett. **45**, 790; **48**, 569(E).

Morrison, P. J., and R. D. Hazeltine, 1984, Phys. Fluids **27**, 886.

Morrison, P. J., and M. Kotschenreuther, 1990, in *Nonlinear World: IV International Workshop on Nonlinear and Turbulent Processes in Physics*, edited by V. G. Bar'yakhtar, V. M. Chernousenko, N. S. Erokhin, A. B. Sitenko, and V. E. Zakharov (World Scientific, Singapore), p. 910.

Morrison, P. J., and D. Pfirsch, 1989, Phys. Rev. A **40**, 3998.

Morrison, P. J., and D. Pfirsch, 1990, Phys. Fluids B **2**, 1105.

Morrison, P. J., and D. Pfirsch, 1992, Phys. Fluids B **4**, 3038.

Morrison, P. J., and B. A. Shadwick, 1994, Acta Phys. Pol. A **85**, 759.

Moser, J. K., 1958, Commun. Pure Appl. Math. **11**, 81.

Moser, J. K., 1968, Mem. Am. Math. Soc. **81**, 1.

Moser, J. K., 1973, *Stable and Random Motions in Dynamical Systems* (Princeton University, Princeton).

Nambu, Y., 1973, Phys. Rev. D **7**, 2405.

Newcomb, W. A., 1960, Ann. Phys. (N.Y.) **10**, 232.

Newcomb, W. A., 1962, Nucl. Fusion Suppl. **2**, 451.

Newcomb, W. A., 1967, Proc. Symp. Appl. Math. **18**, 152.

Nycander, J., 1992, Phys. Fluids A **4**, 467.

Olver, P. J., 1982, J. Math. Anal. Appl. **89**, 233.

Ottino, J. M., 1990, Annu. Rev. Fluid Mech. **22**, 207.

Padhye, N., and P. J. Morrison, 1996a, Phys. Lett. A **219**, 287.

Padhye, N., and P. J. Morrison, 1996b, Plasma Phys. Rep. **22**, 869.

Pedlosky, J., 1987, *Geophysical Fluid Dynamics*, 2nd edition (Springer, New York).

Pfirsch, D., 1990, private communication.

Pfirsch, D., 1993, Phys. Rev. A **48**, 1428.

Pfirsch, D., and R. N. Sudan, 1993, Phys. Fluids B **5**, 2052.

Poincaré, H., 1885, Acta Math. **VII**, 259.

Ramani, A., B. Grammaticos, and T. Bountis, 1989, Phys. Rep. **180**, 161.

Rein, G., 1994, Math. Methods Appl. Sci. **17**, 1129.

Ripa, P., 1981, *Nonlinear Properties of Internal Waves*, AIP Conference Proceedings No. 76, edited by B. J. West (AIP, New York), p. 281.

Ripa, P., 1993, Geophys. Astrophys. Fluid Dyn. **70**, 85.

Saletan, E., and A. Cromer, 1971, *Theoretical Mechanics* (Wiley, New York).

Salmon, R., 1982, in *Mathematical Methods in Hydrodynamics and Integrability in Related Dynamical Systems*, AIP Conference Proceedings No. 88, edited by M. Tabor and Y. Treve (AIP, New York), p. 127.

Salmon, R., 1988a, Annu. Rev. Fluid Mech. **20**, 225.

Salmon, R., 1988b, J. Fluid Mech. **196**, 345.

Sanz-Serna, J. M., and M. P. Calvo, 1994, *Numerical Hamiltonian Problems* (Chapman and Hall, London, New York).

Serrin, J., 1959, in *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin), Vol. VIII pt. 1, p. 125.

Shepherd, T. G., 1990, Adv. Geophys. **32**, 287.

Shepherd, T. G., 1992, in, *Topological Aspects of the Dynamics of Fluids and Plasmas*, edited by H. K. Moffatt (Kluwer Academic, Dordrecht), p. 275.

Sommeria, J. S. D. Meyers, and H. L. Swinney, 1991, in *Nonlinear Topics in Ocean Physics*, edited by A. Osborne (North-Holland, Amsterdam).

Souriau, J.-M., 1970, *Structure des systèmes dynamiques* (Dunod, Paris).

Sudarshan, E., 1963, ''Principles of Classical Dynamics,'' University of Rochester, Department of Physics and Astronomy Report NYO 10250.

Sudarshan, E., and N. Mukunda, 1974, *Classical Dynamics: A Modern Perspective* (Wiley, New York).

Sun, G.-Z., and J. M. Finn, 1987, Phys. Fluids **30**, 770.

Thompson, W., and P. G. Tait, 1921, *Treatise on Natural Philosophy* (Cambridge University, Cambridge, England), pt. 1, p. 388.

Truesdell, C., 1966, *Six Lectures on Modern Natural Philosophy* (Springer, New York).

Van Kampen, N. G., and B. U. Felderhof, 1967, *Theoretical Methods in Plasma Physics* (North-Holland, Amsterdam).

Weiland, J., and H. Wilhelmsson, 1977, *Coherent Nonlinear Interactions of Waves in Plasmas* (Pergamon, New York).

Weinstein, A., 1983a, Phys. Fluids **26**, 388.

Weinstein, A., 1983b, J. Diff. Geom. **18**, 523; **22**, 255(E).

Whitham, G., 1974, *Linear and Nonlinear Waves* (Wiley-Interscience, New York).

Whittaker, E. T., 1937, *Analytical Dynamics* (Cambridge University, London, England), Sec. 136, p. 101.

Williamson, J., 1936, Am. J. Math. **58**, 141.

Wiggins, S., 1992, *Chaotic Transport in Dynamical Systems* (Springer, New York).

Wintner, A., 1947, *The Analytical Foundations of Celestial Mechanics* (Princeton University, Princeton, New Jersey).

Zakharov, V. E., and E. A. Kuznetsov, 1971, Sov. Phys. Dokl. **15**, 913.

Zakharov, V. E., and L. I. Piterbarg, 1988, Phys. Lett. A **126**, 497.

Zeitlin, V., 1992, J. Phys. A **25**, 171.

Zeitlin, V., and T. Kambe, 1993, J. Phys. A **26**, 5025.