Hamiltonian Description of the Ideal Fluid

Lectures presented at the
Geophysical Fluid Dynamics Summer Program
Woods Hole Oceanographic Institution
Woods Hole, MA — June 1993

P.J. MORRISON
Department of Physics and
Institute for Fusion Studies
The University of Texas at Austin
Austin, Texas 78712

January 1994

Hamiltonian Description of the Ideal Fluid

P. J. MORRISON

Department of Physics and Institute for Fusion Studies
The University of Texas at Austin
Austin, Texas 78712
morrison@hagar.ph.utexas.edu

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 will 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, over the past 13 years 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 there are various kinds of calculations that are done again and again, for example, calculations regarding:

1. waves and instabilities by means of linear eigenanalyses;

2. parameter dependencies 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. the nature of the spectrum is not arbitrary, but possesses limitations;

2. upon collision of eigenvalues only certain types of bifurcations can occur;

3. the existence of 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. 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 (dissipationless) 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 (that I won't 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.
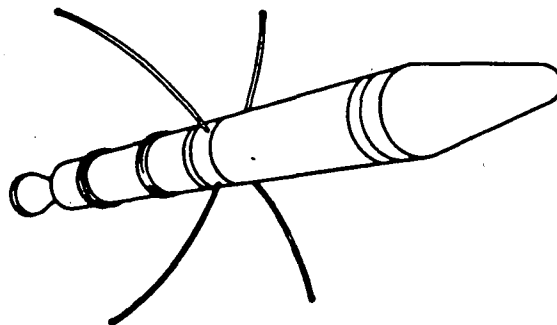

Figure 1:

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 Figure 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 soon began to tumble. The reason for this is attributed

to energy dissipation in the small antennae shown in the figure. Thus unlike the simple oscillator, where 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 Figure 2),* which has an obvious design difference. This satellite behaved like the damped linear oscillator in the sense that dissipation merely caused it to spin down. I would like to emphasize that the difference between the behavior of Explorer I and Alouette I lies in a mathematical property of the Hamiltonian dynamics of these spacecraft; it could have been predicted.
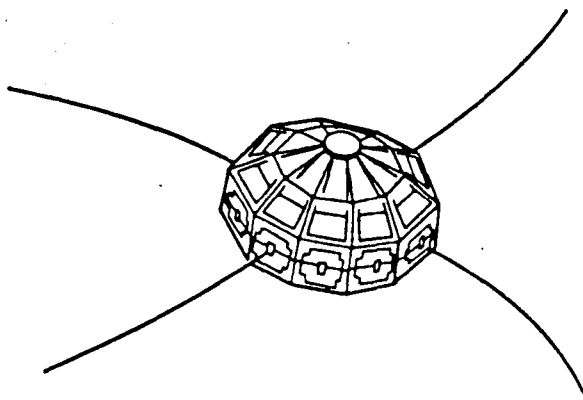

Figure 2:

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. Lecture I 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. What should you expect of the dynamics? Lecture I, being a sketch of low degree-of-freedom Hamiltonian dynamics, answers this to some degree. The remaining four lectures are concerned with the structure of infinite degree-of-freedom Hamiltonian systems, although I will often use finite systems for means of exposition. To see how the Lectures are organized, consult the Table of Contents.

References are given both as footnotes and at the end of various sections. Those at the ends of sections are typically of a more general and comprehensive nature. The referencing should not be taken as complete, but as a guide to the literature.

Acknowledgments

First and foremost I would like to thank Rick Salmon for being an excellent Director of the summer program and for facilitating this cathartic outburst. I would like to thank the Fellows for being a jolly good audience and for their aid in the preparation of this

---

*Both Figures 1 and 2 are after P. W. Likins, AGARD Lecture Series, 3b-1 (1971).

Table of Contents

I. Rudiments of Few Degree-of-Freedom Hamiltonian Systems Illustrated by Passive Advection in Two-Dimensional Fluids

In this introductory lecture we will review some basic aspects of Hamiltonian systems with a finite number of degrees of freedom. We illustrate, in particular, properties of one, two, and three degree-of-freedom systems by considering the passive advection of a tracer in two-dimensional incompressible fluid flow. 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 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 passive advection problem that is treated in this lecture; the former is governed by partial differential equations while the later is governed by ordinary differential equations.

A. A Model for Two-Dimensional Fluid Motion

In various situations fluids are adequately described by models where motion only occurs in two spatial dimensions. An important example is that of rotating fluids where 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 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 the following:

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

where $q(x, y, t)$ is a vorticity-like variable, $\psi(x, y, t)$ is a stream function, both of which 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, which is also the Jacobian, is used:

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

and we have assumed incompressible flow, which implies that the two components of the velocity field are given by

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

---

*J. Pedlosky, *Geophysical Fluid Dynamics*, 2nd ed. (Springer-Verlag, New York, 1987).

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:

- The two-dimensional Euler equation where $q = \nabla^2 \psi$.

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

In the former case q is the vorticity, while in the latter case q is the potential vorticity.

For convenience we will suppose that the domain D is an annular region as depicted in Figure 1 below. Many experiments have been performed in this geometry* where the fluid swirls about in the θ and r directions and is predominantly 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 (I.2) becomes

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

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 manifest by the factor of $1/r$. (In Lecture III we will discuss this in detail.) To preserve the canonical form we replace r by a new coordinate $J := r^2/2$ and the bracket becomes

$$[f, g] = \frac{\partial f}{\partial \theta} \frac{\partial g}{\partial J} - \frac{\partial f}{\partial J} \frac{\partial g}{\partial \theta}. \tag{I.5}$$

These coordinates are convenient, since they can be action-angle variables, as we will see.
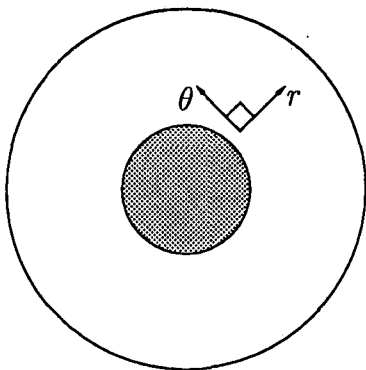
Figure 1:

A solution to (I.1) provides a stream function, $\psi(\theta, J, t)$. In this lecture we will assume that various forms of ψ are known, without going into detail as to whether or not these forms are solutions with particular choices of $\mathcal{L}$, $\mathcal{S}$, or $\mathcal{D}$. Here we will just suppose that the tracers

---

*See e.g. H. P. Greenspan, *Theory of Rotating Fluids* (Cambridge Univ. Press, Cambridge, 1968) and J. Sommeria, S. D. Meyers and H. L. Swinney, *Nonlinear Topics in Ocean Physics*, edited by A. Osborne, (North-Holland, Amsterdam, 1991).

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 $\psi = $ constant for some particular time defines an instantaneous stream line whose tangent is the velocity field. (See Figure 2 below.)
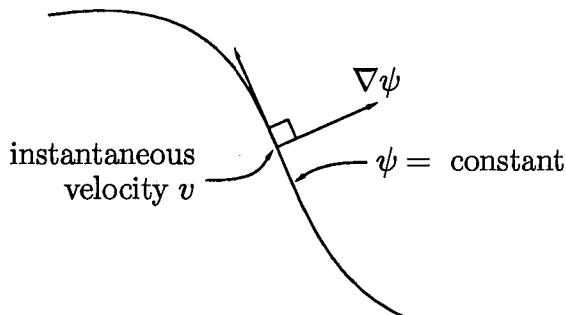
Figure 2:

B.   Passive Advection

Imagine that a tiny piece of the fluid is labeled, somehow, in 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

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

or in terms of the (θ, J) variables

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

These are special cases of Hamilton's equations, which are usually written as

$$\dot{p_i} = [p_i, H] = -\frac{\partial H}{\partial q_i}, \qquad \dot{q_i} = [q_i, H] = \frac{\partial H}{\partial p_i}, \qquad i = 1, 2, \dots N, \qquad (I.8)$$

where $[\,,\,]$, the *Poisson bracket*, is defined by

$$[f, g] = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right). \qquad (I.9)$$

Here (q_i, p_i) constitutes a canonically conjugate pair 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.

8

We emphasize, once again, that the coordinates (x, y) are coordinates for something labelling a fluid element, and the motion of the fluid element 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 Lecture 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 as to 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 (I.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 so precise, since one might want to distinguish between different types of time dependencies. We will 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 old-fashioned mechanics texts make them the center piece (if not the only piece). A theorem often credited to Siegel* 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."

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.

$$\frac{dI_i}{dt} = [I_i, H] = 0, \qquad i, j = 1, 2, \dots N, \tag{I.10}$$

that are in *involution*, i.e.

$$[I_i, I_j] = 0, \qquad i, j = 1, 2, \dots N. \tag{I.11}$$

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 = $ constant for $i = 1, 2, \dots N$ be *compact* and *connected*. 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

$$\frac{dJ_i}{dt} = -\frac{\partial H}{\partial \theta_i} = 0, \qquad \frac{d\theta_i}{dt} = \frac{\partial H}{\partial J_i} = \Omega_i(J), \qquad i, j = 1, 2, \dots N. \tag{I.12}$$

---

*See e.g. J. Moser, *Stable and Random Motions in Dynamical Systems*, (Princeton University Press, Princeton, 1973).

The first of Eqs. (I.12) implies $H = H(J)$ 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. (I.12) is easy to integrate, yielding

$$\theta_i = \theta_i^0 + \Omega_i(J)t \qquad i, j = 1, 2, ...N , \qquad (I.13)$$

where θ_i^0 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 to try and reduce equations to the action-angle form above. If one could find a coordinate transformation, in particular a canonical transformation (c.f. Lecture III), that takes the system of interest into the form of (I.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, now it is 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 that lying on an N-dimensional torus. A distinct feature of trajectories is that they display *sensitive dependence* on initial conditions. We will say a little about this below.

To conclude this section 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:

$$\dot{\psi} = \frac{\partial \psi}{\partial x} \dot{x} + \frac{\partial \psi}{\partial y} \dot{y} = 0 , \qquad (I.14)$$

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

$$\psi(x, y) = \psi_0 \qquad (I.15)$$

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

$$\dot{y} = \left. \frac{\partial \psi}{\partial x}(x, y) \right|_{x = f(\psi_0, y)} =: D(\psi_0, y) . \qquad (I.16)$$

Equation (I.16) is separable, which implies

$$\int_{y_0}^{y} \frac{dy'}{D(\psi_0, y)} = \int_{t_0}^{t} dt' . \qquad (I.17)$$

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 even invertible explicitly, and usually one cannot do the integral explicitly. Moreover, afterwards one must invert (I.17) to obtain the trajectory. These are only technical problems, ones that are easily surmounted with modern computers.

Generally equations of the form of (I.1) possess equilibrium solutions when ψ and q depend upon only a single coordinate. The case of special interest here is when 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 stream lines are "energy surfaces," which are merely concentric circles as depicted in Figure 3 below. The counterpart of (I.13), the equations of motion for the speck of dust in the fluid, are

$$\theta = \theta_0 + \Omega(r)\, t, \qquad r = r_0 \qquad\qquad (I.18)$$

where $v_\theta = \Omega r$. Note the speck goes round and round at a rate dependent upon its radius, but does not go in or out.
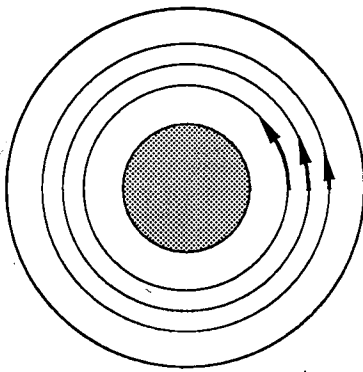
Figure 3:

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 is briefly described.

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) = $ constant $=: E$ is compact (i.e. contained within a three sphere). Since the motion is restricted to this surface p_2 can be eliminated in lieu 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 it 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 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

11
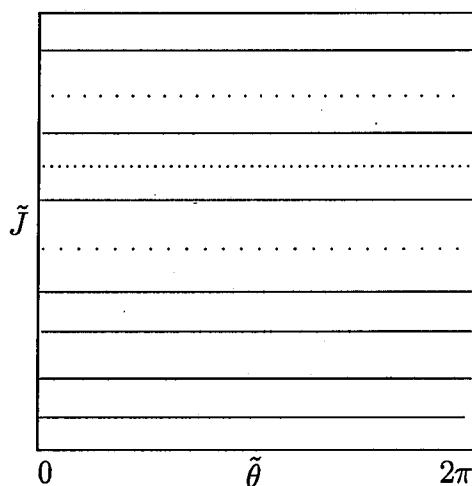
Figure 4:

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 Figure 4. In nonintegrable or chaotic systems this is not true as is illustrated Figures 5(a) and 5(b), where it is seen that orbits make "erratic" patterns.

(a)   (b)
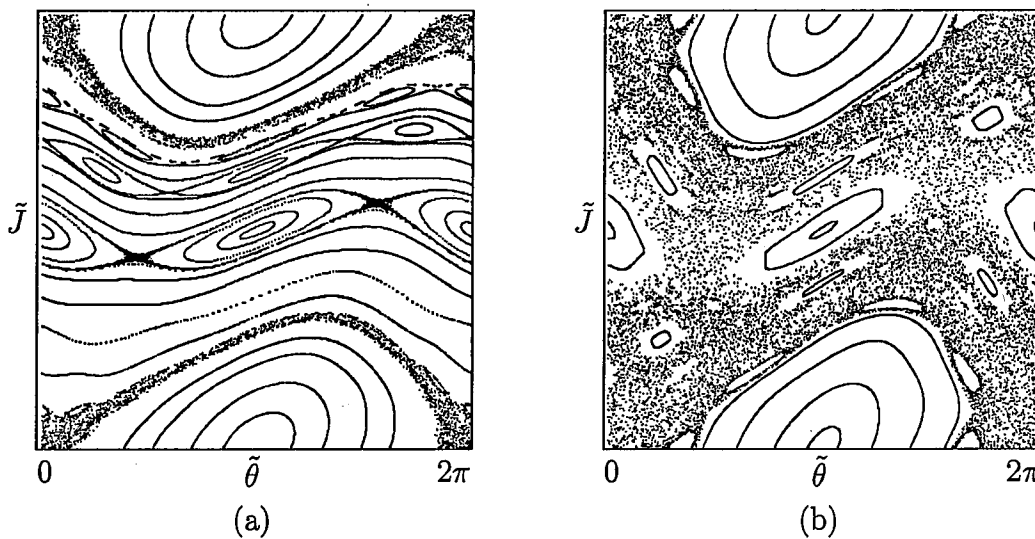
Figure 5:

Now what about the fluid mechanics illustration? Can chaos exist? How can we have a two degree-of-freedom system when we only have the two spatial coordinates, say (θ, J)? The answer is that explicit time dependence in ψ, the extra half of 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 two degree-of-freedom system, and vice verse.

12

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

$$H(\theta, J, \phi, I) = \psi(\theta, J, \phi) + I\,. \qquad (I.19)$$

The equations of motion for this Hamiltonian are

$$\frac{d\theta}{ds} = \frac{\partial H}{\partial J} = \frac{\partial \psi}{\partial J}\,, \qquad \frac{dJ}{ds} = -\frac{\partial H}{\partial \theta} = -\frac{\partial \psi}{\partial \theta}\,; \qquad (I.20)$$

$$\frac{d\phi}{ds} = \frac{\partial H}{\partial I} = 1\,, \qquad \frac{dI}{ds} = -\frac{\partial H}{\partial \phi} = -\frac{\partial \psi}{\partial \phi}\,. \qquad (I.21)$$

The first of Eqs. (I.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. (I.20) give the correct equations of motion. What is the role of the second of Eqs. (I.21)? This equation merely tells us that I has to change so as to make H=constant.

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 force or vector 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.

We will leave it as an exercise to show how to construct $1\frac{1}{2}$ degree-of-freedom Hamiltonian systems from two degree-of-freedom Hamiltonian systems.

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

$$\psi(J, \theta, t) = \psi_0(J) + \psi_1(J) \cos\left(m_1(\theta - \omega_1 t)\right)\,, \qquad (I.22)$$

where $m_1 \in \mathbb{N}$ 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.

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 into the frame of reference rotating with the wave. In this frame the stream function becomes

$$\tilde{\psi}(\tilde{J}, \tilde{\theta}, t) = \psi_0(\tilde{J}) + \psi_1(\tilde{J}) \cos\left(m_1\tilde{\theta}\right) - \omega_1\tilde{J}\,, \qquad (I.23)$$

where the canonical transformation is $\tilde{J} = J$, $\tilde{\theta} = \theta - \omega_1 t$. This transformation is derivable from the mixed variable generating function $F(\theta, \tilde{J}) = \tilde{J}(\theta - \omega_1 t)$. Note the term $-\omega_1 \tilde{J}$ 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 Figure 3 (or straight lines as plotted in Figure 4). However, from the form of (I.22) it is clear that something interesting is going to happen at stagnation points, that is, where

$$\frac{\partial \psi}{\partial \tilde{\theta}} = \frac{\partial \psi}{\partial \tilde{J}} = 0\,. \qquad (I.24)$$

13

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 Figure 6 below.
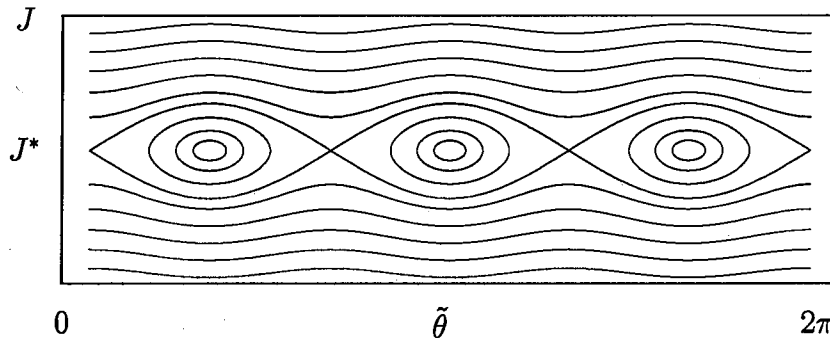
Figure 6:

From the picture it is clear that orbits lie on surfaces and from the form of the stream function given by Eq. (I.23) it is clear that the motion can be solved by quadrature. The use of the coordinate $\tilde\theta = \theta - \omega_1 t$ 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 coordinates. 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

$$\psi(J, \theta, t) = \psi_0(J) + \psi_1(J) \cos\left(m_1(\theta - \omega_1 t)\right) + \psi_2(J) \cos\left(m_2(\theta - \omega_2 t)\right) . \qquad (I.25)$$

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 below in Figure 7. This is a plot of (I.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 $\tilde\theta$ 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. Generally separatrices become fuzzy, but some continuous curves still exists as shown in Figure 5.

As stated above, Figure 7 is not a surface of section, because the resonances were plotted independently, but if they are far apart and the amplitude of the resonances are both small it looks about right. To see the real surface of section one could integrate the differential
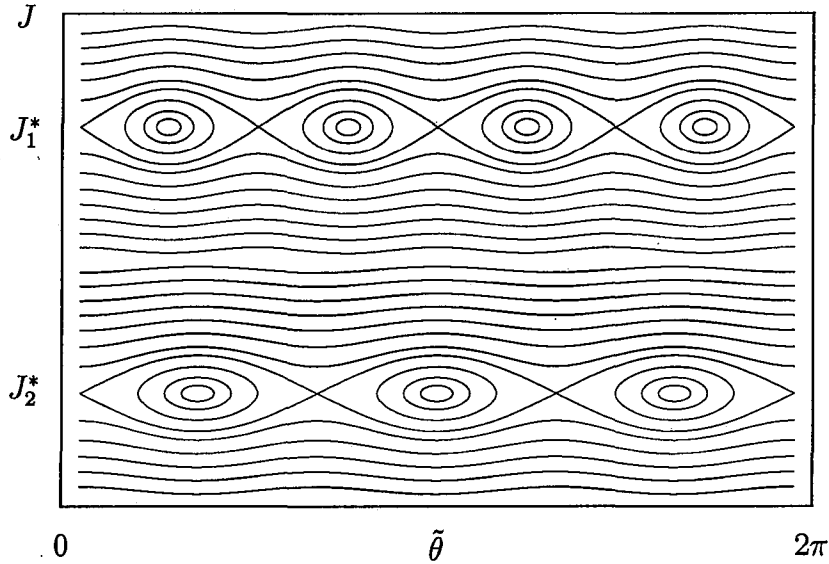
Figure 7:

equations numerically. * 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):

$$\tilde{\theta}_{n+1} = \tilde{\theta}_n + \tilde{J}_{n+1}$$

$$\tilde{J}_{n+1} = \tilde{J}_n - k \, \sin(\tilde{\theta}_n) \,, \tag{I.26}$$

where $\tilde{J}$ and $\tilde{\theta}$ are computed mod-2π. This is an example of an area preserving map; it was, in fact, used to obtain Figures 4, 5(a), and 5(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 where $k = 0$ was shown in Figure 4, 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 KAM theorem. For larger k more and more curves are broken, but some still exist (see Figure 5(a) where $k = .80$ and Figure 5(b) where $k = 1.2$). At a critical value of $k_c \approx .97$, curves that span $0 < \tilde{\theta} \leq 2\pi$ no longer exist. The critical value k_c was calculated by Greene* to many decimal places.

The question of when the last continuous curve breaks is an important one of Hamiltonian dynamics theory. In particular, it is of importance in the passive advection fluid mechanics

---

*To do this one can use standard Runge-Kutta packages. However, now more sophisticated *symplectic* integration algorithms exist. See e.g. C. Kueny, "Nonlinear Instability and Chaos in Plasma Wave-Wave Interactions," Ph.D. Thesis, University of Texas at Austin (1993) and many references cited therein.

*J. M. Greene, J. Math. Phys. **20**, 1183 (1979).

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,

$$\frac{W_1}{2} + \frac{W_2}{2} = |J_1^* - J_2^*| \qquad (I.27)$$

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 only a single lecture it is only possible to scratch the surface and hopefully whet your appetite. Conspicuously absent from this lecture is any discussion of the notions of universality and renormalization. There is much to be learned from the references given below.

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 Figure 8 below.
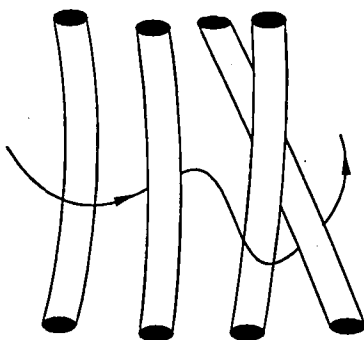
Figure 8:

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 (I.1) is not periodic, but *quasiperiodic*, a special case of which is represented by the following:

$$\psi(\theta, J, t) = f(\theta, J, \omega_2 t, \omega_3 t),$$ (I.28)

where f is a function that satisfies

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

If ω_1/ω_2 is irrational, then ψ is not periodic.

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

$$H(\theta_1, J_1, \theta_2, J_2, \theta_3, J_3) = f(\theta_1, J_1, \theta_2, \theta_3) + \omega_2 J_2 + \omega_3 J_3,$$ (I.30)

and introduce the false time s as before. Note the last two terms of (I.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

$$\frac{d\theta_1}{ds} = \frac{\partial f}{\partial J_1}, \qquad \frac{d\theta_2}{ds} = \omega_2, \qquad \frac{d\theta_3}{ds} = \omega_3;$$ (I.31)

$$\frac{dJ_1}{ds} = -\frac{\partial f}{\partial \theta_1}, \qquad \frac{dJ_2}{ds} = -\frac{\partial f}{\partial \theta_2}, \qquad \frac{dJ_3}{ds} = -\frac{\partial f}{\partial \theta_3}.$$ (I.32)

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

The kind of quasiperiodic system treated in this section is undoubtedly relevant for the study of transport in two dimensional fluids. Solutions of (I.1) are more likely quasiperiodic than periodic. A stream function that describes an azimuthally symmetric shear flow plus three waves with different speeds is quasiperiodic. Transport in systems like this and its generalization to more frequencies is not well understood.

References

[1] V. I. Arnold, "Small Denominators and Problems of Stability of Motion in Classical Celestial Motion," Usp. Mat. Nauk. **18**, 85-191 (1963).

[2] M. V. Berry, "Regular and Irregular Motion," in *Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems*, eds. M. Tabor and Y. Treve, AIP Conf. Proc. 88 (AIP, New York, 1982), pp. 16-120.

[3] J. D. Meiss, "Symplectic Maps, Variational Principles, and Transport," Rev. Mod. Phys. **64**, 795-848 (1992).

[4]R. S. MacKay, "Renormalization in Area-Preserving Maps," Ph.D. Thesis, Princeton University (1982).

[5]J. M. Ottino, "Mixing, Chaotic Advection and Turbulence," Ann. Rev. Fluid Mech. **22**, 207 (1990).

[6]D. del-Castillo-Negrete and P. J. Morrison, "Chaotic Advection by Rossby Waves in Shear Flow," Phys. Fluids A **5**, 948-965 (1993).

II. Functional Calculus, Two Action Principles of Mechanics, and the Action Principle and Canonical Hamiltonian Description of the Ideal Fluid

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 section we will not be concerned with rigor, but with learning how to perform various formal manipulations.

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

$$T[u] = \tfrac{1}{2} \int_{x_0}^{x_1} \rho_0 u^2 \, dx \, . \tag{II.1}$$

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 in (II.1), 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, which is given by

$$\delta K[u; \delta u] := \lim_{\epsilon \to 0} \frac{K[u + \epsilon \, \delta u] - K[u]}{\epsilon} = \frac{d}{d\epsilon} K[u + \epsilon \, \delta u] \Big|_{\epsilon = 0}$$

$$=: \int_{x_0}^{x_1} \delta u \, \frac{\delta K}{\delta u(x)} \, dx =: \left\langle \frac{\delta K}{\delta u}, \delta u \right\rangle . \tag{II.2}$$

We will assume that the limit exists and that there are no problems with the equalities above; later, however, we will give an exercise where 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 (II.2) is the *functional derivative* of the functional K. This notation for the functional derivative is chosen since it emphasizes the fact that $\delta K / \delta u$ is a gradient in function space. The reason why the arguments of u are sometimes displayed will become clear below.

For the example of (II.1) the first variation is given by

$$\delta T[u; \delta u] = \int_{x_0}^{x_1} \rho_0 u \, \delta u \, dx \, , \tag{II.3}$$

and hence the functional derivative is given by

$$\frac{\delta T}{\delta u} = \rho_0 u \,. \tag{II.4}$$

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

$$\delta f(x; \delta x) = \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x_i} \delta x_i =: \nabla f \cdot \delta x \,. \tag{II.5}$$

It is interesting to compare the definition of (II.5) with the last definition of (II.2). The "$\cdot$" in (II.5) 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

$$F[u] = \int_{x_0}^{x_1} \mathcal{F}(x, u, u_x, u_{xx}, \ldots) \, dx \,, \tag{II.6}$$

where $\mathcal{F}$ is an ordinary, sufficiently differentiable, function of its arguments. Note $u_x := du/dx$, etc. The first variation of (II.6) yields

$$\delta F[u; \delta u] = \int_{x_0}^{x_1} \left[\frac{\partial \mathcal{F}}{\partial u} \delta u + \frac{\partial \mathcal{F}}{\partial u_x} \delta u_x + \frac{\partial \mathcal{F}}{\partial u_{xx}} \delta u_{xx} + \cdots \right] dx \,, \tag{II.7}$$

which upon integration by parts becomes

$$\delta F[u; \delta u] = \int_{x_0}^{x_1} \left[\frac{\partial \mathcal{F}}{\partial u} - \frac{d}{dx} \frac{\partial \mathcal{F}}{\partial u_x} + \frac{d^2}{dx^2} \frac{\partial \mathcal{F}}{\partial u_{xx}} - \cdots \right] \delta u \, dx \; + \left[\frac{\partial \mathcal{F}}{\partial u_x} \delta u + \cdots \right] \Big|_{x_0}^{x_1} . \tag{II.8}$$

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 integrated 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, the boundary term vanishes, (II.8) becomes

$$\delta F[u; , \delta u] = \left\langle \frac{\delta F}{\delta u}, \delta u \right\rangle \,, \tag{II.9}$$

where

$$\frac{\delta F}{\delta u} = \frac{\partial \mathcal{F}}{\partial u} - \frac{d}{dx} \frac{\partial \mathcal{F}}{\partial u_x} + \frac{d^2}{dx^2} \frac{\partial \mathcal{F}}{\partial u_{xx}} - \cdots \,. \tag{II.10}$$

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 $\delta F[\hat{u}]/\delta u = 0$. 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,

$$L[u] = \int_{x_0}^{x_1} \sqrt{1 + u_x^2}\, dx \,. \tag{II.11}$$

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

$$u(x') = \int_{x_0}^{x_1} \delta(x - x')\, u(x)\, dx \,, \tag{II.12}$$

where $\delta(x - x')$ is the Dirac delta function and where we have departed from the "[ ]" notation. Applying the definition of (II.2) yields

$$\frac{\delta u(x')}{\delta u(x)} = \delta(x - x') \,. \tag{II.13}$$

This is the infinite dimensional analogue of $\partial x_i / \partial x_j = \delta_{ij}$.

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,

$$T[\rho, v] = \tfrac{1}{2} \int_D \rho v^2\, d^3x \,, \tag{II.14}$$

where the velocity has three Cartesian 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

$$\frac{\delta T}{\delta v_i} = \rho v_i \,, \qquad \frac{\delta T}{\delta \rho} = \frac{v^2}{2} \,. \tag{II.15}$$

We will use these later.

For a more general functional $F[\psi]$, where $\psi(x) = (\psi_1, \psi_2, \dots, \psi_\nu)$ and $x = (x_1, x_2, \dots, x_n)$, the analogue of (II.2) is

$$\delta F[\psi; \delta\psi] = \int_D \delta\psi_i \frac{\delta F}{\delta\psi_i(x)}\, d^n x =: \left\langle \frac{\delta F}{\delta\psi}, \delta\psi \right\rangle \,. \tag{II.16}$$

Here and henceforth we sum repeated indices.

As an exercise consider the pathological functional.

$$P[\psi] = \int_{-1}^{1} \mathcal{P}(\psi_1, \psi_2)\, dx \,, \tag{II.17}$$

where

$$\mathcal{P} = \begin{cases} \dfrac{\psi_1 \psi_2^2}{\psi_1^2 + \psi_2^4} \,, & \psi_1 \neq 0 \\[2ex] 0 \,, & \psi_1 = 0 \,. \end{cases} \tag{II.18}$$

21

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

$$u = \mathcal{O}w. \tag{II.19}$$

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

$$\mathcal{O} := \sum_{k=0}^{n} a_k(x)\frac{d^k}{dx^k}, \tag{II.20}$$

where, as usual, u, w, and a_k have as many derivatives as needed. We can define a functional on w by inserting (II.19) into $F[u]$:

$$\overline{F}[w] := F[u] = F[\mathcal{O}\,w]. \tag{II.21}$$

Equating variations yields

$$\left\langle \frac{\delta\overline{F}}{\delta w}, \delta w \right\rangle = \left\langle \frac{\delta F}{\delta u}, \delta u \right\rangle, \tag{II.22}$$

where the equality makes sense if δu and δw are connected by (II.19), i.e.

$$\delta u = \mathcal{O}\,\delta w, \tag{II.23}$$

where we assume an arbitrary δw induces a δu.

Inserting (II.23) into (II.22) yields

$$\left\langle \frac{\delta\overline{F}}{\delta w}, \delta w \right\rangle = \left\langle \frac{\delta F}{\delta u}, \mathcal{O}\,\delta w \right\rangle$$

$$= \left\langle \mathcal{O}^\dagger \frac{\delta F}{\delta u}, \delta w \right\rangle, \tag{II.24}$$

where $\mathcal{O}^\dagger$ is the formal adjoint of $\mathcal{O}$. Since δw is arbitrary

$$\frac{\delta\overline{F}}{\delta w} = \mathcal{O}^\dagger \frac{\delta F}{\delta u}. \tag{II.25}$$

This follows from the DuBois-Reymond lemma, which is proven by assuming (II.25) 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 needn't be for the chain rule to work in one direction. Functionals transform in the other direction. Given a relation $\psi[\chi]$ we can calculate an expression for $\delta F/\delta\chi$, where F is a functional of χ through ψ. However, given an arbitrary functional of χ, we cannot obtain a functional of ψ.

The above was clearly a special case in that the two functions u and w were linearly related. More generally, consider the functional $F[\psi]$ and suppose ψ is related to $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_1, \ldots, \mathcal{X}_\mu)$ in an arbitrary, not necessarily linear, way:

$$\psi_i = \psi_i[\mathcal{X}] \;, \qquad i = 1, 2, \ldots, \nu \,. \tag{II.26}$$

This "[ ]" notation could be confusing, but we have already stated that ψ and $\mathcal{X}$ are functions. A variation of ψ induced by $\mathcal{X}$ requires linearization of (II.26), which we write as

$$\delta\psi_i = \frac{\delta\psi_i}{\delta\mathcal{X}}[\mathcal{X}; \delta\mathcal{X}] \;, \qquad i = 1, 2, \ldots, \nu \,, \tag{II.27}$$

or simply, since $\delta\psi/\delta x$ is a linear operator on δx,

$$\delta\psi_i = \frac{\delta\psi_i}{\delta\mathcal{X}_j}\, \delta\mathcal{X}_j \qquad i = 1, 2, \ldots, \nu; \quad j = 1, 2, \ldots, \mu \,. \tag{II.28}$$

Inserting (II.28) into (II.22) implies

$$\left\langle \frac{\delta F}{\delta x}, \delta\mathcal{X} \right\rangle = \left\langle \left(\frac{\delta\psi}{\delta\mathcal{X}}\right)^\dagger \frac{\delta F}{\delta\psi}, \delta\mathcal{X} \right\rangle \,, \tag{II.29}$$

whence it is seen that

$$\frac{\delta F}{\delta\mathcal{X}_j} = \left(\frac{\delta\psi_i}{\delta\mathcal{X}_j}\right)^\dagger \frac{\delta F}{\delta\psi_i} \qquad i = 1, 2, \ldots, \nu; \quad j = 1, 2, \ldots, \mu \,. \tag{II.30}$$

Here we have dropped the overbar on F, as is commonly done. In (II.30) 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 (II.27) becomes

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

$$\delta v = \frac{\delta v}{\delta\psi}\, \delta\psi = \frac{\partial}{\partial x}\, \delta\psi \,, \tag{II.31}$$

and

$$\frac{\delta F}{\delta\psi} = \frac{\partial}{\partial y}\frac{\delta F}{\delta u} - \frac{\partial}{\partial x}\frac{\delta F}{\delta v} \,. \tag{II.32}$$

Now consider the *second variation*, $\delta^2 F$, and *second functional derivative*, $\delta^2 F/\delta\psi\delta\psi$. Since the first variation, $\delta F[\psi; \delta\psi]$, is a functional of ψ, a second variation can be made in

this argument:

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

$$=: \int_D \delta\psi_i \frac{\delta^2 F}{\delta\psi_i \delta\psi_j} \delta\hat\psi_j \, d^n x =: \left\langle \delta\psi, \frac{\delta^2 F}{\delta\psi\delta\psi}\delta\hat\psi \right\rangle. \qquad (\text{II}.33)$$

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 (II.33) defines $\delta^2 F/\delta\psi\delta\psi$, which is a linear operator that acts on $\delta\hat\psi$ but depends nonlinearly on ψ. It possesses a symmetry analogous to the interchange of the order of second partial differentiation. To see this observe

$$\delta^2 F[\psi; \delta\psi, \delta\hat\psi] = \frac{\partial^2}{\partial\eta\partial\epsilon}F[\psi + \eta\delta\hat\psi + \epsilon\delta\psi]\Big|_{\epsilon=0,\,\eta=0}. \qquad (\text{II}.34)$$

Since the order of differentiation in (II.33) is immaterial it follows that

$$\left(\frac{\delta^2 F}{\delta\psi_i \delta\psi_j}\right)^\dagger = \frac{\delta^2 F}{\delta\psi_j \delta\psi_i}, \qquad i = 1, 2, \ldots, \nu. \qquad (\text{II}.35)$$

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 (II.11). Performing the operations of (II.34) yields

$$\delta^2 L[u; \delta u; \delta\hat u] = \int_{x_0}^{x_1} \delta u_x \frac{1}{(1 + u_x^2)^{3/2}} \delta\hat u_x \, dx. \qquad (\text{II}.36)$$

Thus

$$\frac{\delta^2 L}{\delta u^2} = \frac{d}{dx} \frac{-1}{(1 + u_x^2)^{3/2}} \frac{d}{dx}. \qquad (\text{II}.37)$$

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 (II.6), is to suppose $(x_0, x_1) = (-\pi, \pi)$ and expand in a Fourier series,

$$u(x) = \sum_{k=-\infty}^{\infty} u_k e^{ikx}. \qquad (\text{II}.38)$$

Upon inserting (II.38) into (II.6) 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

$$F[u] = F(u_0, u_1, u_{-1}, \ldots). \qquad (\text{II}.39)$$

24

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

$$F(x_1, x_2, \ldots) = \sum_{k=1}^{\infty} \left(\tfrac{1}{2} a_k \, x_k^2 - \tfrac{1}{4} b_k \, x_k^4 \right) , \tag{II.40}$$

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

$$a_k = \frac{1}{k^6} , \qquad b_k = \frac{1}{k^2} . \tag{II.41}$$

Assuming that (II.40) converges uniformly, the first variation yields

$$\delta F = \sum_{k=1}^{\infty} \left(a_k \, x_k - b_k \, x_k^3 \right) \delta x_k , \tag{II.42}$$

which has three extremal points

$$x_k^{(0)} = 0, \qquad x_k^{(\pm)} = \pm (a_k / b_k)^{1/2} , \tag{II.43}$$

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

$$\delta^2 F = \sum_{k=1}^{\infty} a_k (\delta x_k)^2 . \tag{II.44}$$

where we assume (II.42) converges uniformly for x_k and δx_k. Since $a_k > 0$ for all k, (II.44) is positive definite; i.e.

$$\delta^2 F > 0 \quad \text{for} \quad \delta x_k \neq 0 , \quad \text{for all} \quad k . \tag{II.45}$$

However, consider ΔF defined by

$$\Delta F = F(x^{(0)} + \Delta x) - F(x^{(0)}) = \sum_{k=1}^{\infty} \left[\tfrac{1}{2} a_k \, (\Delta x_k)^2 - \tfrac{1}{4} b_k \, (\Delta x_k)^4 \right] , \tag{II.46}$$

which we evaluate at

$$\Delta x_k = \begin{cases} \dfrac{1}{m} , & k = m \\[2mm] 0 , & k \neq m, \end{cases} \tag{II.47}$$

and obtain

$$\Delta F < 0 , \tag{II.48}$$

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 = $ 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 an minimum. We will leave it to you to explain why the functional $F(x_1, x_2, \ldots)$ is not strongly positive. This example points to a mathematical technicality that is encountered when proving stability by Liapunov's method (cf. Lecture V).

References

[1]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Wiley Interscience, New York, 1953) Vol. I Chapt. IV.

[2]I. M. Gelfand and S. V. Fomin, *Calculus of Variations* (Prentice-Hall, Englewood Cliffs, N. J., 1963).

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 because 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 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 can't 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 will show how this story goes for the ideal fluid, but first we review some mechanics.

---

*C. Truesdell, *Six Lectures on Modern Natural Philosophy* (Springer-Verlag, New York, 1966).

One approach to producing the equations of motion for a mechanics problem is to first identify the configuration space, Q, with coordinates $q = (q_1, q_2, \ldots, q_N)$. 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:

$$S[q] = \int_{t_0}^{t_1} L(q, \dot{q}, t) \, dt \,, \tag{II.49}$$

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

$$\frac{\delta S[q]}{\delta q_i} = 0 \tag{II.50}$$

imply Lagrange's equations,

$$\frac{\partial L}{\partial q_i} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \,. \tag{II.51}$$

This is Hamilton's principle.

Since for particles in rectangular coordinates usually

$$T = \tfrac{1}{2} \sum_i m_i \dot{q}_i^2 \qquad V = V(q) \,, \tag{II.52}$$

Eqs. (II.51) yield

$$m_i \ddot{q}_i = -\frac{\partial V}{\partial q_i} \,, \qquad i = 1, 2, \ldots, N \,. \tag{II.53}$$

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 thermodynamics or, as is the case here, dynamics. It has a nice geometric interpretation, but we will skip this. Here we will use it to transform the N second order differential equations of (II.53) 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

$$\overline{H}(q, p, \dot{q}, t) := \sum_i p_i \dot{q}_i - L(q, \dot{q}, t) \,. \tag{II.54}$$

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

$$\delta \overline{H} = \sum_i \left[\frac{\partial \overline{H}}{\partial q_i} \delta q_i + \frac{\partial \overline{H}}{\partial \dot{q}_i} \delta \dot{q}_i + \frac{\partial \overline{H}}{\partial p_i} \delta p_i \right] + \frac{\partial \overline{H}}{\partial t} \delta t$$

$$= \sum_i \left[-\frac{\partial L}{\partial q_i} \delta q_i + \left(p_i - \frac{\partial L}{\partial \dot{q}_i} \right) \delta \dot{q}_i + \dot{q}_i \delta p_i \right] - \frac{\partial L}{\partial t} \delta t \,. \tag{II.55}$$

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

$$\frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i} \, ; \qquad \frac{\partial H}{\partial p_i} = \dot{q}_i \, ; \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \, . \qquad (\text{II}.56)$$

Lagrange's equations, (II.51), together with the definition of p_i and the middle of (II.56) give Hamilton's equations:

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \, , \qquad \dot{q}_i = \frac{\partial H}{\partial p_i} \, . \qquad (\text{II}.57)$$

In order to explicitly calculate $H(p, q, t)$ in the standard case of mechanics, one uses $p_i = \partial L / \partial \dot{q}_i$ to solve for $\dot{q} = \dot{q}(p)$ and then inserts this into (II.54). This requires L to be convex in $\dot{q}$. Since there exist important physical cases where L is not convex, Dirac and others developed a theory to handle this. An interesting new application of Dirac's constraint theory for filtering out fast motion in GFD models has recently been developed.*

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

$$S[q, p] = \int_{t_0}^{t_1} \left[\sum_i p_i \dot{q}_i - H(q, p, t) \right] \, dt \, , \qquad (\text{II}.58)$$

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 Figure 1 below, 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,

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \, , \qquad \dot{q}_i = \frac{\partial H}{\partial p_i} \, . \qquad (\text{II}.59)$$

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

References

[1] E. Saletan and A. Cromer, *Theoretical Mechanics* (John Wiley, New York, 1971).

[2] E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics: A Modern Perspective* (John Wiley, New York, 1974).

---

*R. Salmon, J. Fluid Mech. **196**, 345 (1988).
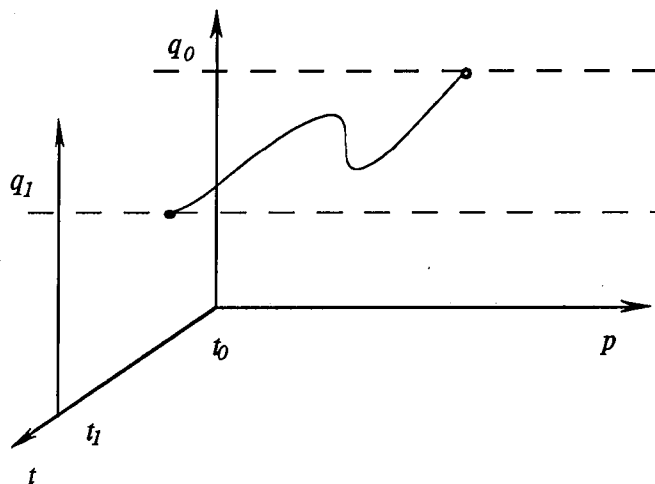
Figure 1:

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 fluid mechanics, but we're going to do so in terms of 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 don't exist: if they were of macroscopic size, they wouldn't 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 systems, is given by

$$q = q(a, t), \tag{II.60}$$

where $q = (q_1, q_2, q_3)$. This is the material or Lagrangian variable. Here $a = (a_1, a_2, a_3)$ is any label that identifies a fluid particle, which is often taken to be the position of the fluid particle at time $t = 0$.* The quantities $q_i(a, t)$ are coordinates for the configuration space

---

*Note, however, that the freedom to relabel particles is associated in an important way with the Casimir invariants that are discussed below. See e.g. M. Calkin, Can. J. Phys. **41**, 2241 (1963); P. Ripa, AIP Conf. Proc. **76** (1981); R. Salmon, AIP Conf. Proc. **88**, 127 (1982).

29

$\mathcal{Q}$, 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 will assume that as many derivatives of q with respect to a as needed exist, but we won't say more about $\mathcal{Q}$; in fact, not that much is known about the solution space for the 3-D fluid equations in Lagrangian variables. At this stage we will 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 the fluid can flow at nonzero velocity. Potential energy is stored in terms of pressure and temperature. More precisely we adapt the energy representation of thermodynamics where 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 later quantity is a measure of the volume. The intensive quantities, pressure and temperature, are obtained as follows:

$$T = \frac{\partial U}{\partial s}(s, \rho), \qquad p = \rho^2 \frac{\partial U}{\partial \rho}(s, \rho). \tag{II.61}$$

The second of (II.61) is a bit peculiar—it arises because the volume, the usual thermodynamics variable, is proportional to ρ^{-1}. Note also that special choices for U produce specific fluid flows: barotropic flow, adiabatic flow, etc.

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 sidetrack 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 follows:

$$\rho(r, t) = \int_D \delta(r - q(a, t)) \, \rho_0(a) \, d^3a. \tag{II.62}$$

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 (II.62) 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

$$\rho(r, t) = \frac{\rho_0(a)}{\mathcal{J}(a, t)}\bigg|_{a = q^{-1}(r, t)}, \tag{II.63}$$

where the Jacobian $\mathcal{J} = \det(\partial q_i/\partial a_j)$. That this is local mass conservation follows from

$$\rho\, d^3q = \rho_0\, d^3a\,, \tag{II.64}$$

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 will 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 (II.62) corresponding to this statement is given by

$$\sigma(r,t) = \int_D \sigma_0(a)\delta(r - q(a,t))\, d^3a\,, \tag{II.65}$$

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 (II.63) is

$$s(r,t) = s_0(a)\big|_{a=q^{-1}(r,t)}\,. \tag{II.66}$$

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 which 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 (II.62) and (II.65) is given by the following:

$$M(r,t) = \int_D \dot{q}(a,t)\,\delta(r - q(a,t))\,\rho_0(a)\, d^3a\,. \tag{II.67}$$

Performing the integration produces the counterpart of (II.63) and (II.64), viz.

$$v(r,t) = \dot{q}(a,t)\big|_{a=q^{-1}(r,t)}\,, \tag{II.68}$$

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

$$V[q] = \int_D \rho_0 U(s_0, \rho_0/\mathcal{J})\, d^3a\,. \tag{II.69}$$

Observe that (II.69) 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

$$T[q] = \tfrac{1}{2}\int_D \rho_0\, \dot{q}^2\, d^3a\,. \tag{II.70}$$

Observe that (II.70) is a functional of q that depends only upon $\dot{q}$.

31

From (II.69) and (II.70) the Lagrangian functional is obtained,

$$L[q, \dot{q}] = \int_D \left[\tfrac{1}{2}\rho_0 \, \dot{q}^2 - \rho_0 U(s_0, \rho_0/\mathcal{J}) \right] d^3a \qquad (\text{II}.71)$$

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

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

$$S[q] = \int_{t_0}^{t_1} L[q, \dot{q}] \, dt = \int_{t_0}^{t_1} \int_D \left[\tfrac{1}{2} \, \rho_0 \, \dot{q}^2 - \rho_0 U \right] d^3a \,. \qquad (\text{II}.72)$$

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.

$$\int_D d^3a \leftrightarrow \sum_i \, . \qquad (\text{II}.73)$$

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,

$$\delta q(a, t_0) = \delta q(a, t_1) = 0 \,. \qquad (\text{II}.74)$$

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. Later we will see what this implies.

In order to apply Hamilton's principle, we must functionally differentiate (II.72), thus, it is necessary to know something about differentiating determinants. Recall

$$\frac{\partial q_k}{\partial a_j} \frac{A_{ki}}{\mathcal{J}} = \delta_{ij} \qquad (\text{II}.75)$$

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

$$A_{ki} = \tfrac{1}{2} \, \epsilon_{kjl}\epsilon_{imn}\frac{\partial q_j}{\partial a_m} \frac{\partial q_l}{\partial a_n} \,, \qquad (\text{II}.76)$$

where ϵ_{ijk} is the skew symmetric tensor (density), which vanishes if any two of i, j, k are equal, 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 (II.72) we will require the following relation:

$$\frac{\partial \mathcal{J}}{\partial q_{i,j}} = A_{ij} \,, \qquad (\text{II}.77)$$

which follows from (II.75).

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

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) + \frac{\partial}{\partial a_j}\left(\frac{\partial \mathcal{L}}{\partial q_{i,j}}\right) - \frac{\partial \mathcal{L}}{\partial q_i} = 0, \tag{II.78}$$

provided the surface integral vanishes:

$$\int_{t_0}^{t_1}\int_{\partial D} p\, \delta q_i A_{ij} n_j \, d^2a = \int_{t_0}^{t_1}\int_{\partial D} p\, \delta q \cdot \hat{n}\, d^2q. \tag{II.79}$$

The equality above follows upon changing from integration over a to integration over q. Note $A_{ij}d^2a = d^2q$. Evidently 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. In the case where D is a box and periodic boundary conditions are imposed, the vanishing of the surface term is automatic. In the case where D is "all space" the physical boundary condition is (ii), which asserts that the pressure vanishes at infinity.

From (II.78) the equation of motion is obtained,

$$\rho_0\ddot{q}_i - A_{ij}\frac{\partial}{\partial a_j}\left(\frac{\rho_0^2}{\mathcal{J}^2}\frac{\partial U}{\partial \rho}\right) = 0. \tag{II.80}$$

Here we have used $\partial A_{ij}/\partial a_j = 0$, which you can work out using (II.76). Alternatively, upon using (II.75) the equation of motion can be written in the form

$$\rho_0\ddot{q}_j\frac{\partial q_j}{\partial a_i} - \mathcal{J}\frac{\partial}{\partial a_i}\left(\frac{\rho_0^2}{\mathcal{J}^2}\frac{\partial U}{\partial \rho}\right) = 0. \tag{II.81}$$

We leave it to you to show that (II.80) can be transformed into Eulerian form:

$$\rho\left(\frac{\partial v}{\partial t} + v \cdot \nabla v\right) = -\nabla p, \tag{II.82}$$

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

$$\frac{\partial}{\partial q_k} = \frac{1}{\mathcal{J}}A_{ki}\frac{\partial}{\partial a_i}. \tag{II.83}$$

With (II.83) it is clear that (II.80) is of the form of Newton's second law. The Legendre transform follows easily. The canonical momentum density is

$$\pi_i(a, t) := \frac{\delta L}{\delta \dot{q}_i(a)} = \rho_0\dot{q}_i, \tag{II.84}$$

33

and

$$H[\pi, q] = \int d^3a \left[\pi \cdot q - \mathcal{L}\right] = \int d^3a \left[\frac{\pi^2}{2\rho_0} + \rho_0 U\right] . \tag{II.85}$$

Hamilton's equations are then

$$\dot{\pi}_i = -\frac{\delta H}{\delta q_i}; \qquad \dot{q}_i = \frac{\delta H}{\delta \pi_i}. \tag{II.86}$$

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

$$\{F, G\} = \int \left[\frac{\delta F}{\delta q} \cdot \frac{\delta G}{\delta \pi} - \frac{\delta G}{\delta q} \cdot \frac{\delta F}{\delta \pi}\right] d^3a \tag{II.87}$$

viz.,

$$\dot{\pi}_i = \{\pi_i, H\}; \qquad \dot{q}_i = \{q_i, H\}. \tag{II.88}$$

Here $\delta q_i(a)/\delta q_j(a') = \delta_{ij}\delta(a - a')$ has been used, a relation that is analogous to $\partial q_j/\partial q_i = \delta_{ij}$ for finite systems [recall (II.13)].

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.

References

[1] J. Serrin, "Mathematical Principles of Classical Fluid Mechanics," in *Handbuch der Physik* VIII pt. 1, ed. S. Flügge (Springer-Verlag, Berlin, 1959) pp. 125-263.

[2] C. Eckart, "Variation Principles of Hydrodynamics," Phys. Fluids **3**, 421-427 (1960).

[3] W.A. Newcomb, "Lagrangian and Hamiltonian Methods in Magnetohydrodynamics," Nuclear Fusion Supplement, pt. 2, 451-463 (1962).

[4] R. Salmon, "Hamilton's Principle and Ertel's Theorem," in *Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems*, eds. M. Tabor and Y. Treve, AIP Conf. Proc. 88 (AIP, New York, 1982), pp. 127-135.

III. Noncanonical Hamiltonian Dynamics—Examples

A. Noncanonical Hamiltonian Dynamics

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

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

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

$$\sum_i \left[\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right] = \sum_i \left[\frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q_i} \right] = 0, \tag{III.2}$$

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

$$\mathcal{V}(t) = \int \prod_{i=1}^{N} dp_i \, dq_i, \tag{III.3}$$

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

$$\frac{d\mathcal{V}}{dt} = 0. \tag{III.4}$$

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 (III.1) is incompressibility of the vector field; i.e.

$$\sum_i \frac{\partial V^i}{\partial z^i} = 0, \tag{III.5}$$

whence it follows that

$$\mathcal{V}(t) = \int \prod_{i=1}^{N} dz_i \tag{III.6}$$

is constant in time.

Suppose (III.5) is not true, as is the case for the following example:

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

$$\dot{z}_2 = \frac{z_1}{z_2^2 + 1}. \tag{III.7}$$

For this system

$$\frac{\partial \dot{z}_1}{\partial z_1} + \frac{\partial \dot{z}_2}{\partial z_2} = -\frac{2z_1 z_2}{(z_2^2 + 1)^2} \neq 0. \tag{III.8}$$

You would be mistaken if, based on (III.8), you concluded that (III.7) is 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 will discuss below.

So, is there a method for determining whether a system is Hamiltonian in general? 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 for doing this.* Nevertheless we can say some things; however, 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.* This includes Liouville's equation for the dynamics of the phase space density of a collection of particles, the BBGKY hierarchy, the Vlasov equation, ideal fluid dynamics and various approximations thereof, magnetized fluids,...; it includes essentially every fundamental equation for classical media.

So with the above motivation, let us turn to discussing noncanonical Hamiltonian dynamics for finite degree of freedom systems, a formalism that extends back to Sophus Lie. The first step is to write Hamilton's equations in covariant form. Thus define

$$z^i = \begin{cases} q_i & \text{for } i = 1, 2, \ldots N, \\ p_{i-N} & \text{for } i = N+1, \ldots 2N. \end{cases} \tag{III.9}$$

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

$$\dot{z}^i = J_c^{ij} \frac{\partial H}{\partial z^j} = [z^i, H], \tag{III.10}$$

where the Poisson bracket is given by

$$[f, g] = \frac{\partial f}{\partial z^i} J_c^{ij} \frac{\partial g}{\partial z^j}, \tag{III.11}$$

with

$$(J_c^{ij}) = \begin{pmatrix} 0_N & I_N \\ -I_N & 0_N \end{pmatrix}. \tag{III.12}$$

Above, the repeated indices are to be summed over $1, 2, \ldots 2N$. In (III.12), 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

$$\bar{z}^i = \bar{z}^i(z). \tag{III.13}$$

---

*Techniques for finding constants of motion do exist, but necessarily possess limitations. See e.g. A. Ramani, B. Gramniaticos, and T. Bountis, Phys. Rep. **180**, 161 (1989), and references therein.

The Hamiltonian H transforms as a scalar:

$$H(z) = \overline{H}(\overline{z}) \, . \tag{III.14}$$

Taking time derivatives of (III.13) yields

$$\dot{\overline{z}}^l = \frac{\partial \overline{z}^l}{\partial z^i} \dot{z}^i = \frac{\partial \overline{z}^l}{\partial z^i} J_c^{ij} \frac{\partial H}{\partial z^j} = \left[\frac{\partial \overline{z}^l}{\partial z^i} J_c^{ij} \frac{\partial \overline{z}^m}{\partial z^j} \right] \frac{\partial \overline{H}}{\partial \overline{z}^m} \, . \tag{III.15}$$

Defining

$$J^{lm} := \frac{\partial \overline{z}^l}{\partial z^i} J_c^{ij} \frac{\partial \overline{z}^m}{\partial z^j} \, , \tag{III.16}$$

we see 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

$$\dot{\overline{z}}^l = J^{lm}(\overline{z}) \frac{\partial \overline{H}}{\partial \overline{z}^m} = [\overline{z}^l, H] \, , \tag{III.17}$$

where the Poisson bracket is now given by

$$[f, g] = \frac{\partial f}{\partial \overline{z}^l} J^{lm} \frac{\partial g}{\partial \overline{z}^m} \, . \tag{III.18}$$

Notice in (III.17) we have displayed the explicit $\overline{z}$ dependence in J^{lm}. This was done to emphasize an important distinction—that between *covariance* and *form invariance*. Equation (III.16) is a statement of covariance, while a statement of form invariance is given by

$$J_c^{lm} = \frac{\partial \overline{z}^l}{\partial z^i} J_c^{ij} \frac{\partial \overline{z}^m}{\partial z^j} \, . \tag{III.19}$$

This is in fact the 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 $\overline{z}^l$ 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 $\overline{z}^l$.

Let us now return to Liouville's theorem. Taking the divergence of (III.17) yields

$$\frac{\partial \dot{\overline{z}}^l}{\partial \overline{z}^l} = \frac{\partial J^{lm}}{\partial \overline{z}^l} \frac{\partial \overline{H}}{\partial \overline{z}^m} +^{lm} \frac{\partial^2 \overline{H}}{\partial \overline{z}^l \partial \overline{z}^m} \, . \tag{III.20}$$

The second term vanishes because J^{lm} is antisymmetric and $\partial^2 \overline{H} / \partial \overline{z}^l \partial \overline{z}^m$ 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:

$$\frac{\partial \dot{\overline{z}}^l}{\partial \overline{z}^l} = \frac{\partial J^{lm}}{\partial \overline{z}^l} \frac{\partial \overline{H}}{\partial \overline{z}^m} \neq 0 \, . \tag{III.21}$$

37

This was the situation for our little example of (III.7).

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 (III.3):

$$\mathcal{V}(t) = \int \mathcal{J} \prod_{l=1}^{2N} d\bar{z}^l . \tag{III.22}$$

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 (III.16) and using the determinant product rule:

$$\mathcal{J} = \frac{1}{\sqrt{\det J}} . \tag{III.23}$$

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 $\overline{H}$ is such that the two vectors of (III.21) 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.*

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 will play another sort of game. Suppose you have a system of the form of (III.1) 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 (III.17) with Poisson brackets like (III.18). Finally, suppose somehow you have found a constant of motion, call it H, and you think this is the energy and therefore a good bet for the Hamiltonian. Then you can write

$$V^i(z) = J^{ij} \frac{\partial H}{\partial z^j} \qquad i = 1, 2, \ldots M . \tag{III.24}$$

Everything in (III.24) is known except J^{ij}, which is required to be antisymmetric because of (III.16). 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 (III.8).] 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 (III.18) is assembled it satisfies

$$[f, g] = -[g, f] \qquad \forall\, f, g \tag{III.25}$$

$$[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0 \qquad \forall\, f, g \tag{III.26}$$

*See e.g. D. Montgomery and R. Kraichnan, Rep. Prog. Phys. **43**, 35 (1979).

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 will not prove Darboux's theorem, but will mention that Eq. (III.26) 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 properties (III.25) and (III.26), and elements that are functions defined on the phase space. We will say more about this later.

The bracket properties, (III.25) and (III.26), can be translated into properties required of the cosymplectic form. The first is evidently

$$J^{ij} = -J^{ji}. \tag{III.27}$$

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

$$S^{ijk} = J^{il}\frac{\partial J^{jk}}{\partial z^l} + J^{jl}\frac{\partial J^{ki}}{\partial z^l} + J^{kl}\frac{\partial J^{ij}}{\partial z^l} = 0. \tag{III.28}$$

In going from (III.26) to (III.28) it is observed that all the terms involving second derivatives that arise upon calculating $[f, [g, h]] + [g, [h, f]] + [h, [f, g]]$ cancel; only the terms where the derivative of the outer bracket acts upon the J of the inner bracket survive. This fact makes life much easier when verifying the Jacobi identity.

Now suppose everything worked out right except the J you found had det$J = 0$, with some rank $2N < M$. What then? A generalization of the Darboux theorem, which was proven[*] at least by the 1930s, says that J can be transformed into the following form:

$$(J_c) = \begin{pmatrix} 0_N & I_N & 0 \\ -I_N & 0_N & 0 \\ 0 & 0 & 0_{M-2N} \end{pmatrix}. \tag{III.29}$$

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

From (III.29) 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 below in Figure 1.

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 are 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 exists possibly $M - 2N$ independent null eigenvectors. A consequence of the Jacobi identity is that this is the case and, moreover, the null space can in fact be spanned by the gradients of the Casimir invariants, which satisfy

$$J^{ij}\frac{\partial C^{(\alpha)}}{\partial z^j} = 0, \tag{III.30}$$

---

[*]See e.g. L. P. Eisenhart, *Continuous Groups of Transformations* (Dover, New York, 1961) and R. Littlejohn, AIP Conf. Proc. 88, 47 (1982).
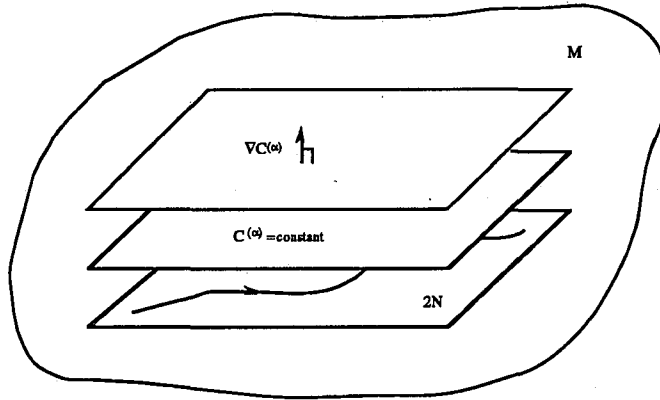
Figure 1:

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

$$\dot{C}^{(\alpha)} = \frac{\partial C^{(\alpha)}}{\partial z^i} J^{ij} \frac{\partial H}{\partial z^j} = 0 \,. \tag{III.31}$$

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 kinematical 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 $C^{(\alpha)} =$ constant. Dynamics generated by any H that begins on a particular $\mathcal{P}$ surface remains there.

The picture we have described above is the finite dimensional analogue of the Hamiltonian form possessed by Eulerian continuous media theories. We will describe the Poisson brackets for some of them soon, but now we mention that for these theories the J^{ij} has a special form that is linear in the z^k, i.e.

$$J^{ij} = c_k^{ij} z^k \,, \tag{III.32}$$

where the c_k^{ij} are constants—in fact, structure constants for a Lie algebra. In light of (III.27) and (III.28) they must satisfy

$$c_k^{ij} = -c_k^{ji} \tag{III.33}$$

and

$$c_m^{ij} c_\ell^{mk} + c_m^{jk} c_\ell^{mi} + c_m^{ki} c_\ell^{mj} = 0 \,. \tag{III.34}$$

Brackets with J's of the form of (III.32) are called *Lie-Poisson*.

It is interesting to reexamine the condition for Liouville's theorem (III.21) for J's of the above form,

$$\frac{\partial J^{lm}}{\partial z^l} \frac{\partial H}{\partial z^m} = c_\ell^{lm} \frac{\partial H}{\partial z^m} = 0 \,. \tag{III.35}$$

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 (III.B)]. In general semisimple Lie algebras can, by a coordinate change, be brought into a form

where the structure constants are completely antisymmetric.[*] In these coordinates there is Liouville's theorem without the need for inserting a Jacobian as in (III.23). This, as noted above, is typically the case for fluid theories in Eulerian variables.

In infinite dimensions the analogue of (III.18) is given by

$$\{F, G\} = \int_D \frac{\delta F}{\delta \psi^i} \mathfrak{J}^{ij} \frac{\delta G}{\delta \psi^j} \, d\mu =: \left\langle \frac{\delta F}{\delta \psi}, \mathfrak{J} \frac{\delta G}{\delta \psi} \right\rangle, \tag{III.36}$$

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

$$\{F, G\} = -\{G, F\} \qquad \forall \, F, G \tag{III.37}$$

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0 \qquad \forall \, F, G, H \tag{III.38}$$

where F and G are now functionals. Analogous to (III.27) the antisymmetry condition of (III.37) requires $\mathfrak{J}$ to be *skew-symmetric*, i.e.

$$\langle f, \mathfrak{J} \, g \rangle = \left\langle \mathfrak{J}^\dagger f, g \right\rangle = -\langle g, \mathfrak{J} \, f \rangle . \tag{III.39}$$

The Jacobi identity (III.38) for infinite dimensional systems has a condition analogous to (III.28); one need only consider functional derivatives of $\mathfrak{J}$ when calculating $\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\}$.[*] For Eulerian media, as noted above, the *cosymplectic operator* typically has the Lie-Poisson form

$$\mathfrak{J}^{ij} = C_k^{ij} \psi^k, \tag{III.40}$$

where C_k^{ij} are *structure operators*. We will clarify the meaning of these structure operators by examples; a bit more will be said in Lecture IV.

References

[1] E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics: A Modern Perspective*, (John Wiley, New York, 1974).

[2] P. J. Morrison, "Poisson Brackets for Fluids and Plasmas," in *Mathematical Methods in Hydrodynamics and Integrability in Related Dynamical Systems*, eds. M. Tabor and Y. Treve, AIP Conference Proceedings **88**, La Jolla, 1982, p. 13.

[3] R. Salmon, Ann. Rev. Fluid Mech. **20**, 225 (1988).

---

[*] N. Jacobson, *Lie Algebras* (Wiley Interscience, New York, 1962).
[*] See e.g. P. J. Morrison (1982), *Ref. III A.*

B.   Examples

Below the noncanonical Poisson brackets for several systems are presented. The first, the free rigid body, is a finite dimensional system, the others are infinite dimensional. We present the brackets here and refer the reader to the references[†] for a discussion 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:

$$\dot{\ell}_1 = \ell_2 \ell_3 \left(\frac{1}{I_3} - \frac{1}{I_2} \right),$$

$$\dot{\ell}_2 = \ell_3 \ell_1 \left(\frac{1}{I_1} - \frac{1}{I_3} \right),$$

$$\dot{\ell}_3 = \ell_1 \ell_2 \left(\frac{1}{I_2} - \frac{1}{I_1} \right), \tag{III.41}$$

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

$$H = \tfrac{1}{2} \sum_{i=1}^{3} \frac{\ell_i^2}{I_i}. \tag{III.42}$$

The function H is easily seen to be a constant of motion upon differentiating with respect to time and making use of (III.41). The Poisson bracket for this system is of Lie-Poisson type

$$[f, g] = -\epsilon_{ijk} \ell_k \frac{\partial f}{\partial \ell_i} \frac{\partial g}{\partial \ell_j}. \tag{III.43}$$

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 (III.34)—something that is not difficult to verify directly. It is evident upon substituting (III.42) into (III.43), that

$$\dot{\ell}_i = [\ell_i, H], \tag{III.44}$$

is equivalent to (III.41). This system possesses the Casimir invariant

$$C = \tfrac{1}{2} \sum_{i=1}^{3} \ell_i^2, \tag{III.45}$$

which satisfies

$$[C, f] = 0, \qquad \forall f. \tag{III.46}$$

[†]See e.g. Morrison (1982), *l.c.*

Thus the global picture of the phase space $\mathcal{Z}$, which here corresponds to Figure 1, 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-deVries Equation

We write the famous Korteweg-deVries (KdV) equation*, which describes long wavelength water waves and ion-acoustic waves in plasmas, in the following form:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0. \tag{III.47}$$

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:

$$H = \int_D \left[\tfrac{1}{6}u^3 - \tfrac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2\right] dx. \tag{III.48}$$

The noncanonical Poisson bracket, due to Gardner,[†] is given by

$$\{F, G\} = -\int_D \frac{\delta F}{\delta u}\frac{\partial}{\partial x}\frac{\delta G}{\delta u}dx, \tag{III.49}$$

from which it is seen that the cosymplectic operator is

$$\mathfrak{J} = -\frac{\partial}{\partial x}. \tag{III.50}$$

The skew-symmetry of (III.49) 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 (III.48),

$$\frac{\delta H}{\delta u} = \left(\tfrac{1}{2}u^2 + \frac{\partial^2 u}{\partial x^2}\right) \tag{III.51}$$

into (III.49), yields

$$\frac{\partial u}{\partial t} = -\frac{\partial}{\partial x}\left(\tfrac{1}{2}u^2 + \frac{\partial^2 u}{\partial x^2}\right) = -u\frac{\partial u}{\partial x} - \frac{\partial^3 u}{\partial x^3}. \tag{III.52}$$

This bracket possesses one Casimir invariant,

$$C[u] = \int_D u\, dx. \tag{III.53}$$

---

*See e.g. G. Whitham, *Linear and Nonlinear Waves* (Wiley-Interscience, New York, 1974).
[†]C. S. Gardner, J. Math. Phys. **12**, 1548 (1971).

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 has the Lie-Poisson form, when approximations are made this form can change.

3 1-D Pressureless Fluid

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

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0 \,. \tag{III.54}$$

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,

$$H[u] = \int_D \tfrac{1}{2} u^2 dx \,, \tag{III.55}$$

there being no internal energy. The following bracket, with the above Hamiltonian, produces (III.54):

$$\{F, G\} = -\tfrac{1}{3} \int_D u \left[\frac{\delta F}{\delta u} \frac{\partial}{\partial x} \frac{\delta G}{\delta u} - \frac{\delta G}{\delta u} \frac{\partial}{\partial x} \frac{\delta F}{\delta u} \right] dx \,; \tag{III.56}$$

that is

$$\frac{\partial u}{\partial t} = \{u, H\} = -\tfrac{1}{3} \left(u\frac{\partial u}{\partial x} + \frac{\partial(u^2)}{\partial x} \right) = -u\frac{\partial u}{\partial x} \,. \tag{III.57}$$

The cosymplectic operator is evidently given by

$$\mathfrak{J} = -\tfrac{1}{3} \left(u\frac{\partial}{\partial x} + \frac{\partial}{\partial x} u \right) = -\tfrac{1}{3} \left(2u\frac{\partial}{\partial x} + \frac{\partial u}{\partial x} \right) \,. \tag{III.58}$$

The following Casimir invariant is easily obtained by solving $\{C, F\} = 0$ for all functionals F; i.e. by searching for null eigenvectors of (III.58) and undoing the functional derivative:

$$C = \int_D |u|^{1/2} \, dx \,. \tag{III.59}$$

It is evident that the following Hamiltonian:

$$H[u] = \tfrac{1}{6} \int_D u^3 dx \,, \tag{III.60}$$

44

together with the bracket (III.49) will also produce (III.54). Thus it is possible for a system to have two Hamiltonian 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*.

4 1-D 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:

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

$$\frac{\partial \rho}{\partial t} = -\frac{\partial(\rho u)}{\partial x}. \tag{III.61}$$

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

$$H[\rho, u] = \int_D \left[\tfrac{1}{2}\rho u^2 + \rho U(\rho) \right] dx, \tag{III.62}$$

and the Poisson bracket is given by

$$\{F, G\} = -\int_D \left[\frac{\delta F}{\delta \rho}\frac{\partial}{\partial x}\frac{\delta G}{\delta u} - \frac{\delta G}{\delta \rho}\frac{\partial}{\partial x}\frac{\delta F}{\delta u} \right] dx. \tag{III.63}$$

The cosymplectic operator

$$\left(\mathfrak{z}^{ij} \right) = \begin{pmatrix} 0 & \frac{\partial}{\partial x} \\ \frac{\partial}{\partial x} & 0 \end{pmatrix}, \tag{III.64}$$

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 won't 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

$$\frac{\partial}{\partial x}\frac{\delta C}{\delta \rho} = 0, \qquad \frac{\partial}{\partial x}\frac{\delta C}{\delta u} = 0, \tag{III.65}$$

from which we obtain the following Casimir invariants:

$$C_1[u] = \int_D u\, dx, \qquad C_2[\rho] = \int_D \rho\, dx. \tag{III.66}$$

---

*See F. Magri, J. Math. Phys. **19**, 1156 (1978).

Using

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

$$\frac{\delta H}{\delta \rho} = \tfrac{1}{2}u^2 + h(\rho) \,, \tag{III.67}$$

where $h(\rho) := \rho U_\rho + U$ is the enthalpy (note that $\delta H / \delta \rho = $ constant is Bernoulli's law), in (III.63) produces

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

$$\frac{\partial \rho}{\partial t} = \{\rho, H\} = -\frac{\partial}{\partial x}(\rho u) \,. \tag{III.68}$$

These equations are seen to be equivalent to (III.61) upon making use of $h_x = p_x/\rho$ (recall $p = \rho^2 U_\rho$).

5   2-D Euler Scalar Vortex Dynamics

The vortex dynamics we consider here, unlike the examples above, has two spatial variables, $r := (x, y) \in D$, in addition to time; that is, it is a $2 + 1$ theory. The noncanonical Poisson bracket possessed by this system[*] is the prototype of $2 + 1$ theories, it being shared by the 1-D Vlasov-Poisson equation[†], quasigeostrophy or the Hasegawa-Mima equation, and others.

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

$$\omega(r, t) := \hat{z} \cdot \nabla \times v \,, \tag{III.69}$$

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

$$v = \left(-\frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x}\right) \,, \tag{III.70}$$

which is related to the vorticity through

$$\omega = \nabla^2 \psi \,. \tag{III.71}$$

The equation of motion for this system is

$$\frac{\partial \omega}{\partial t} = -v \cdot \nabla \omega = -\,[\psi, \omega] \,, \tag{III.72}$$

---

[*]P. J. Morrison, "Hamiltonian field description of two-dimensional vortex fluids and guiding center plasmas," Princeton University Plasma Physics Laboratory Report, PPPL-1783 (1981) (Available as American Institute of Physics Document No. PAPS-PFBPE-04-771, AIP Auxiliary Publication Service, 335 East 45th Street, New York, NY 10017) and P. J. Olver, J. Math. Anal. Appl. **89**, 233 (1982).

[†]P. J. Morrison, Phys. Lett. **80A**, 383 (1980).

where

$$[f,g] = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}. \tag{III.73}$$

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 $\psi =$ constant on ∂D. Since ω is the dynamical variable one might expect the boundary condition to be $\omega =$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 is correct for the Vlasov-Poisson equation, it is unphysical for the ideal fluid where the vorticity at a point on the boundary need not be 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. However, all these complications can be avoided by choosing the domain D to be a finite box and impose 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

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

$$= \tfrac{1}{2}\int_D\int_D \omega(r)\, K(r|r')\, \omega(r')\, d^2r\, d^2r' = -\tfrac{1}{2}\int_D \omega\,\psi\, d^2r, \tag{III.74}$$

where K is defined by

$$\psi(r) = -\int_D K(r|r')\omega(r')\, d^2r'. \tag{III.75}$$

Observe that in the case where $D = \mathbb{R}^2$ the last equality of (III.74) requires the elimination of the logarithmic singularity that comes from integration by parts. The noncanonical Poisson bracket for this system is given by

$$\{F,G\} = \int_D \omega\left[\frac{\delta F}{\delta\omega},\frac{\delta G}{\delta\omega}\right] d^2r, \tag{III.76}$$

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

$$\mathfrak{J} = -[\omega,\cdot]. \tag{III.77}$$

Skew-symmetry follows from

$$\int_D f[g,h]\, d^2r = -\int_D g[f,h]\, d^2r, \tag{III.78}$$

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 generally the case for Lie-Poisson brackets. The Casimir invariant for the bracket of (III.76) is given by

$$C[\omega] = \int \mathcal{C}(\omega)\, d^2r, \tag{III.79}$$

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*. We mention that even though there are an infinity of constants this is insufficient for the 2-D Euler equations to be integrable. In order to obtain the equations of motion we require

$$\frac{\delta H}{\delta \omega} = -\psi \,. \tag{III.80}$$

Evidently,

$$\frac{\partial \omega}{\partial t} = \{\omega, H\} = -[\omega, \frac{\delta H}{\delta \omega}] = [\omega, \psi] \,, \tag{III.81}$$

which is equivalent to (III.72).

6   3-D 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 Cartesian coordinates $r := (x, y, z) =: (x_1, x_2, x_3) \in D$. The dynamical variables used are the same as those discussed in Lecture II: 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

$$\frac{\partial v}{\partial t} = -v \cdot \nabla v - \frac{1}{\rho} \nabla p \,, \tag{III.82}$$

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho v) \,, \tag{III.83}$$

$$\frac{\partial s}{\partial t} = -v \cdot \nabla s \,. \tag{III.84}$$

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

$$H[v, \rho, s] = \int_D \left(\tfrac{1}{2} \rho v^2 + \rho U(\rho, s) \right) d^3r \,, \tag{III.85}$$

and the noncanonical Poisson bracket* is

$$\{F, G\} = -\int_D \left[\left(\frac{\delta F}{\delta \rho} \nabla \cdot \frac{\delta G}{\delta v} - \frac{\delta G}{\delta \rho} \nabla \cdot \frac{\delta F}{\delta v} \right) + \left(\frac{\nabla \times v}{\rho} \cdot \frac{\delta F}{\delta v} \times \frac{\delta G}{\delta v} \right) \right.$$
$$\left. + \frac{\nabla s}{\rho} \cdot \left(\frac{\delta F}{\delta s} \frac{\delta G}{\delta v} - \frac{\delta G}{\delta s} \frac{\delta F}{\delta v} \right) \right] d^3r \,. \tag{III.86}$$

*For a physical explanation see P. J. Morrison, Zeitschrift für Naturforschung **42a**, 1115 (1987).

*P. J. Morrison and J. M. Greene, Phys. Rev. Lett. **45**, 790 (1980); *ibid.* **48**, 569 (1982).

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 $\delta F/\delta v = \nabla \times \delta F/\delta \omega$, the second term is seen to be a generalization of that for the $2+1$ scalar vortex dynamics given above[†]. 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

$$\frac{\delta H}{\delta v} = \rho v\,, \qquad \frac{\delta H}{\delta \rho} = \tfrac{1}{2}v^2 + (\rho U)_\rho\,, \qquad \frac{\delta H}{\delta s} = \rho U_s\,, \qquad \text{(III.87)}$$

Eqs. (III.82)-(III.84) are seen to be equivalent to

$$\frac{\partial v}{\partial t} = \{v, H\}\,, \qquad \frac{\partial \rho}{\partial t} = \{\rho, H\}\,, \qquad \frac{\partial s}{\partial t} = \{s, H\}\,. \qquad \text{(III.88)}$$

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. In the case where D is a finite domain there is a complication with the vanishing of these terms, as in the case for the 2-D 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. However, when D is a finite domain there is a technicality. 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. A 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 (III.86) is not of Lie-Poisson from. However, if a transformation form the variables v, ρ, and s to the conserved variables $M := \rho v$, ρ, and $\sigma := \rho s$, which were introduced in Lecture I (and alluded to above), is made, then the bracket becomes[*]

$$\{F, G\} = -\int_D \left[M_i \left(\frac{\delta F}{\delta M_j} \frac{\partial}{\partial x_j} \frac{\delta G}{\delta M_i} - \frac{\delta G}{\delta M_j} \frac{\partial}{\partial x_j} \frac{\delta F}{\delta M_i} \right) + \rho \left(\frac{\delta F}{\delta M} \cdot \nabla \frac{\delta G}{\delta \rho} - \frac{\delta G}{\delta M} \cdot \nabla \frac{\delta F}{\delta \rho} \right) \right.$$

$$\left. + \sigma \left(\frac{\delta F}{\delta M} \cdot \nabla \frac{\delta G}{\delta \sigma} - \frac{\delta G}{\delta M} \cdot \nabla \frac{\delta F}{\delta \sigma} \right) \right] d^3 r\,. \qquad \text{(III.89)}$$

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

$$\left. \frac{\delta F}{\delta \rho} \right|_{v,s} = \left. \frac{\delta F}{\delta \rho} \right|_{M,s} + \frac{M}{\rho} \cdot \frac{\delta F}{\delta M} + \frac{\sigma}{\rho} \frac{\delta F}{\delta \sigma}\,. \qquad \text{(III.90)}$$

[†]See, for three-dimensional vortex dynamics, E. A. Kuznetsov and A. Mikhailov, Phys. Lett. **77A**, 37 (1980).

[*]Morrison and Greene, *l.c.*; I. E. Dzyaloshinskii and G. E. Volovick, Ann. Phys. **125**, 67 (1980).

It is straightforward to show that (III.89) together with the Hamiltonian

$$H[M,\rho,\sigma] = \int_D \left(\tfrac{1}{2}\frac{M^2}{\rho} + \rho U(\rho,\sigma/\rho) \right) d^3r, \qquad \text{(III.91)}$$

produces the fluid equations of motion in conservation form as follows:

$$\frac{\partial M}{\partial t} = \{M,H\}, \qquad \frac{\partial \rho}{\partial t} = \{\rho,H\}, \qquad \frac{\partial \sigma}{\partial t} = \{\sigma,H\}. \qquad \text{(III.92)}$$

Now consider the condition for the Casimir invariants, $\{F,C\} = 0$ for all F. From (III.86) it is seen that this implies

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

$$\nabla \frac{\delta C}{\delta \rho} + \frac{(\nabla \times v)}{\rho} \times \frac{\delta C}{\delta v} - \frac{\nabla s}{\rho}\frac{\delta C}{\delta s} = 0. \qquad \text{(III.93)}$$

One solution of these equations is

$$C_1[\rho,s] = \int_D \rho f(s)\, d^3r, \qquad \text{(III.94)}$$

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

$$C_2[v] = \int_D v \cdot \nabla \times v\, d^3r. \qquad \text{(III.95)}$$

It will be left to you to investigate the general solution.

7 General Comments

Above we 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 variables (v,ρ,s) or (M,ρ,σ). In closing this lecture we make some brief comments about the classification of the various brackets.

Consider the cases where 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. An invertible transformation of this kind is given by

$$\tilde{u} = \tfrac{1}{6}u^2\, \text{sgn}\,(u), \qquad \text{(III.96)}$$

with the inverse

$$u = \text{sgn}\,(\tilde{u})\sqrt{6\tilde{u}}. \qquad \text{(III.97)}$$

Inserting

$$\frac{\delta F}{\delta u} = 2 \, \text{sgn} \, (\tilde{u}) \sqrt{6\tilde{u}} \, \frac{\delta F}{\delta \tilde{u}} \tag{III.98}$$

into the KdV bracket yields

$$\{F, G\} = -\int_D \frac{\delta F}{\delta u} \frac{\partial}{\partial x} \frac{\delta G}{\delta u} \, dx = -\tfrac{1}{3} \int_D \tilde{u} \left[\frac{\delta F}{\delta \tilde{u}} \frac{\partial}{\partial x} \frac{\delta G}{\delta \tilde{u}} - \frac{\delta G}{\delta \tilde{u}} \frac{\partial}{\partial x} \frac{\delta F}{\delta \tilde{u}} \right] \, dx \, . \tag{III.99}$$

Now it is evident from (III.99) and from above, where we changed fluid variables form (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 will not address here. However, since typically for fluid theories this can be done, this suggests a classification of such theories by their Lie-Poisson bracket, which in turn are classified by the Lie group corresponding to the structure operators. Thus theories can be classified by a Lie group[*] 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 semidirect product. Treatment of this area is beyond the scope of these lectures, although we will briefly comment on this in the context of Clebsch variables in Lecture IV.

---

[*]This idea is an old one. It in essence was developed in the work of Sudarshan (1963), *Ref. IV B*. See also Sudarshan and Mukunda (1974), *Ref. III A*. Further development in the geometrical setting was given by V. I. Arnold, Ann. Inst. Four. **16**, 319 (1966) and Usp. Mat. Nauk. **24**, 225 (1969), although unlike here the (cumbersome) Lagrange bracket is emphasized.

IV. Tutorial on Lie Groups and Algebras, Reduction—Realization, and Clebsch Variables

A. Tutorial on Lie Groups and Lie Algebras

This section, which was in fact a lecture, was given after all the others. However in retrospect it appears for continuity better placed here. It can be skipped by the cognoscenti.

A Lie group $\mathfrak{G}$ is both a group and a differentiable manifold. The elements of the group, which are uncountably infinite in number, correspond to points of the manifold. To be concrete we will consider a realization where elements of $\mathfrak{G}$ correspond to functions that define transformations (coordinate changes) on some manifold $\mathcal{Z}$.

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

$$z'^i = f^i(z, a) \qquad i = 1, 2, \ldots, M, \qquad (IV.1)$$

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 parameterization of the family. For each value of a the functions f constitute a one-one, onto transformation of $\mathcal{Z}$ to 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 form a group under composition of functions.

It is important to distinguish between the M-dimensional manifold $\mathcal{Z}$ 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 is dynamics of a sort; a point of $\mathcal{Z}$ is mapped into 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

$$T_c z = T_b T_a z \qquad (IV.2)$$

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 will see how this goes, but not work it out in detail. In terms of the function of (IV.1) closure can be stated as follows:

$$f(z, c) = f(f(z, a), b) = f(z, \phi(b, a)). \qquad (IV.3)$$

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

$$\begin{bmatrix} z'^1 \\ z'^2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} z^1 \\ z^2 \end{bmatrix}, \qquad (IV.4)$$

or equivalently

$$z' = T_\theta z . \tag{IV.5}$$

This 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 must be equivalent to a rotation through some angle χ:

$$T_\psi T_\theta = \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos(\psi+\theta) & \sin(\psi+\theta) \\ -\sin(\psi+\theta) & \cos(\psi+\theta) \end{bmatrix} =: T_\chi . \tag{IV.6}$$

Evidently, the analogue 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

$$T_a(T_b T_c) = (T_a T_b) T_c . \tag{IV.7}$$

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

$$(T_a T_b) T_c = f(f(f(z, c), b), a) = f(f(z, c), \phi(a, b)) = f(z, \phi(\phi(a, b), c)) . \tag{IV.8}$$

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

$$T_a(T_b T_c) = T_{\phi(a, \phi(b,c))} . \tag{IV.9}$$

Upon comparing (IV.8) and (IV.9) we see that associativity implies

$$\phi(a, \phi(b, c)) = \phi(\phi(a, b), c) . \tag{IV.10}$$

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

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

$$T_0 T_a = T_a T_0 = T_a \tag{IV.11}$$

or

$$f(f(z, a), 0) = f(f(z, 0), a) = f(z, a) . \tag{IV.12}$$

Therefore,

$$\phi(0, a) = \phi(a, 0) = a . \tag{IV.13}$$

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

$$T_a T_{a^{-1}} = T_{a^{-1}} T_a = T_0 . \tag{IV.14}$$

Evidently,

$$\phi(a, a^{-1}) = \phi(a^{-1}, a) = 0 \tag{IV.15}$$

53

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

$$\det\left(\frac{\partial\phi(a,b)}{\partial a}\right) \neq 0\,; \qquad \det\left(\frac{\partial\phi(a,b)}{\partial b}\right) \neq 0\,. \qquad (\text{IV}.16)$$

It is easy to verify the above properties for the example of $S0(2)$; it is recommended that you do this.

A *Lie algebra*, $\mathfrak{g}$, arises in studying the group manifold in a neighborhood of the identity. Such a study yields ordinary differential equations for ϕ.

Suppose δa is small and consider

$$T_{\tilde{a}} := T_a T_{\delta a} \qquad (\text{IV}.17)$$

or

$$\tilde{a} = \phi(a, \delta a)\,. \qquad (\text{IV}.18)$$

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

$$\tilde{a} = a + da = \phi(a, \delta a) \qquad (\text{IV}.19)$$

or in terms of the transformations

$$T_{a+da}z = T_a T_{\delta a} z\,. \qquad (\text{IV}.20)$$

This is depicted in Figure 1.
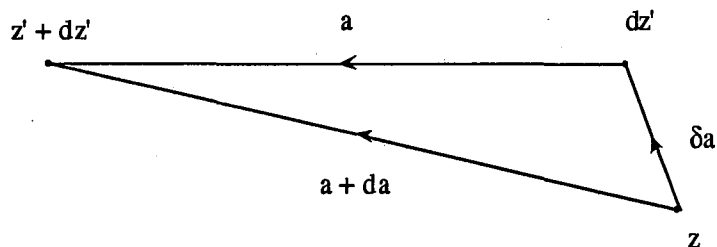
Figure 1:

Taylor expanding (IV.18) about $a = 0$ and $b = \delta a = 0$ yields

$$a^\alpha + da^\alpha = \phi^\alpha(a, 0) + \left.\frac{\partial\phi^\alpha(a,b)}{\partial b^\beta}\right|_{b=0}\delta a^\beta + \dots\,, \qquad (\text{IV}.21)$$

where the greek indices α, β etc., which we will use to denote coordinates of the group manifold, run over $1, 2, \dots, N$. From (IV.21),

$$da^\alpha = L^\alpha_\beta(a)\delta a^\beta \qquad (\text{IV}.22)$$

54

where

$$L_\beta^\alpha(a) := \left.\frac{\partial \phi^\alpha(a,b)}{\partial b^\beta}\right|_{b=0}. \tag{IV.23}$$

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

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

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

$$=: \delta a^\beta X_\beta F(a). \tag{IV.24}$$

The quantities X_β defined by

$$X_\beta := L_\beta^\alpha(a)\frac{\partial}{\partial a^\alpha}, \tag{IV.25}$$

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 quantities X_β are to be thought of as vectors with components $\{L_\beta^\alpha\}$ and basis vectors $\{\partial/\partial a^\alpha\}$. Evidently, if we choose $F(a) = a$, (IV.24) implies

$$da^\gamma = \delta a^\beta X_\beta a^\gamma = L_\beta^\gamma \delta a^\beta. \tag{IV.26}$$

We will use this later.

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

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

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

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

$$+ \mathcal{O}(4), \tag{IV.27}$$

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

$$\phi(a,0) = \phi(0,a) = a \tag{IV.28}$$

for all a, it is clear that $\phi(0,0) = 0$, and upon differentiating (IV.28)

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

$$\frac{\partial \phi^\nu(0,0)}{\partial b^\kappa} = L_\kappa^\nu(0) = \delta_\kappa^\nu. \tag{IV.29}$$

55

Differentiating (IV.28) twice and then thrice in the nonzero argument implies

$$\frac{\partial^2 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda} = \frac{\partial^2 \phi^\nu(0,0)}{\partial b^\kappa \partial b^\lambda} = \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda \partial a^\mu} = \frac{\partial^3 \phi^\nu(0,0)}{\partial b^\kappa \partial b^\lambda \partial b^\mu} = 0 \,. \qquad \text{(IV.30)}$$

However, (IV.28) does not contain information about mixed derivatives; viz.

$$\frac{\partial^2 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda} \,, \qquad \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda \partial b^\mu} \,, \qquad \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda \partial b^\mu} \,. \qquad \text{(IV.31)}$$

Thus far we have reduced (IV.27) to

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

$$+ \frac{1}{2} \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial a^\lambda \partial b^\mu} a^\kappa a^\lambda b^\mu + \frac{1}{2} \frac{\partial^3 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda \partial b^\mu} a^\kappa b^\lambda b^\mu + \mathcal{O}(4) \,, \qquad \text{(IV.32)}$$

To go farther the associativity condition (IV.10) is imposed. If you expand 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

$$c^\nu_{\kappa\lambda} := \frac{\partial^2 \phi^\nu(0,0)}{\partial a^\kappa \partial b^\lambda} - \frac{\partial^2 \phi^\nu(0,0)}{\partial a^\lambda \partial b^\kappa} \,, \qquad \text{(IV.33)}$$

which obviously satisfies

$$c^\nu_{\kappa\lambda} = -c^\nu_{\lambda\kappa} \,, \qquad \text{(IV.34)}$$

the condition obtained is

$$c^\nu_{\kappa\lambda} c^\lambda_{\delta\gamma} + c^\nu_{\gamma\lambda} c^\lambda_{\kappa\delta} + c^\nu_{\delta\lambda} c^\lambda_{\gamma\kappa} = 0 \,. \qquad \text{(IV.35)}$$

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 (IV.34) and (IV.35) are enough to determine ϕ—the structure of the group (that is connected to the identity).

Now we will obtain a differential equation for the group and then discuss briefly some important theorems proved by Lie. Recall Eq. (IV.26)

$$da^\gamma = \delta a^\beta X_\beta a^\gamma = L^\gamma_\beta \delta a^\beta \,, \qquad \text{(IV.36)}$$

which we derived by expanding $T_{a+da} = T_a T_{\delta a}$. Since $\phi(a, a^{-1})$ must have a solution, this implies $L^\alpha_\beta(a)$ has an inverse for all a. We call this $L^{-1\beta}_\gamma$; i.e.

$$L^\alpha_\beta L^{-1\beta}_\gamma = \delta^\alpha_\gamma \qquad \text{(IV.37)}$$
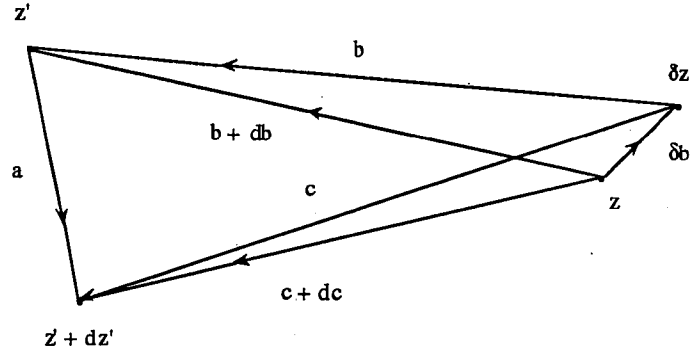
Figure 2:

and (IV.26) can be inverted,

$$\delta a^\beta = L_\alpha^{-1\beta} da^\alpha .$$

(IV.38)

Now suppose

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

(IV.39)

which is depicted below in Figure 2.

Equation (IV.39) implies

$$c + dc = \phi(a, b + db) = \phi(a, \phi(b, \delta b)) .$$

(IV.40)

If $dc = db = \delta b = 0$, then $c = \phi(a, b)$; thus (IV.40) becomes by associativity,

$$c + dc = \phi(\phi(a, b), \delta b) = \phi(c, \delta b) .$$

(IV.41)

Therefore,

$$c^\alpha + dc^\alpha = \phi^\alpha(c, \delta b) = \phi^\alpha(c, 0) + L_\beta^\alpha(c) \, \delta b^\beta + \dots ,$$

(IV.42)

and

$$dc^\alpha = L_\beta^\alpha(c) \, \delta b^\beta = L_\beta^\alpha(c) L_\gamma^{-1\beta}(b) \, db^\gamma ,$$

(IV.43)

where the second equality follows from (IV.38). Evidently,

$$\frac{\partial c^\alpha}{\partial b^\gamma} = L_\beta^\alpha(c) L_\gamma^{-1\beta}(b) ,$$

(IV.44)

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

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

(IV.45)

Equation (IV.45) is a system of partial differential equations of Mayer-Lie type. Here $\phi(a, b)$ is the unknown and a is a fixed parameter. 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, they must satisfy an integrability condition, *viz*

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

(IV.46)

57

which implies

$$\frac{\partial}{\partial b^{\mu}}\left[L_{\beta}^{\alpha}(\phi(a,b))L_{\gamma}^{-1\beta}(b)\right] = \frac{\partial}{\partial b^{\gamma}}\left[L_{\beta}^{\alpha}(\phi(a,b))L_{\mu}^{-1\beta}(b)\right] . \tag{IV.47}$$

Performing the differentiation in (IV.47)

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

$$= \frac{\partial L_{\beta}^{\alpha}(c)}{\partial c^{\nu}}\frac{\partial \phi^{\nu}(a,b)}{\partial b^{\gamma}}L_{\mu}^{-1\beta}(b) + L_{\beta}^{\alpha}(c)\frac{\partial L_{\mu}^{-1\beta}(b)}{\partial b^{\gamma}}, \tag{IV.48}$$

and then using (IV.45) yields

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

$$= L_{\beta}^{\alpha}(c)\left[\frac{\partial L_{\mu}^{-1\beta}(b)}{\partial b^{\gamma}} - \frac{\partial L_{\gamma}^{-1\beta}(b)}{\partial b^{\mu}}\right], \tag{IV.49}$$

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

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

$$= L_{\delta}^{\nu}(b)L_{\beta}^{\alpha}(b)\left[\frac{\partial L_{\nu}^{-1\gamma}(b)}{\partial b^{\alpha}} - \frac{\partial L_{\alpha}^{-1\gamma}(b)}{\partial b^{\nu}}\right], \tag{IV.50}$$

Since the points b and c were arbitrary, the two sides of (IV.50) must equal the same constant, which is determined by setting $c = 0$. Using

$$L_{\delta}^{\nu}(0) = L_{\delta}^{-1\nu}(0) = \delta_{\delta}^{\nu} \tag{IV.51}$$

and

$$\frac{\partial L_{\beta}^{\alpha}(0)}{\partial c^{\nu}} = \frac{\partial^2 \phi^{\alpha}(0,0)}{\partial c^{\nu}\partial b^{\beta}} \tag{IV.52}$$

yields for the two sides

$$\frac{\partial L_{\beta}^{\alpha}(c)}{\partial c^{\nu}}L_{\gamma}^{\nu}(c) - \frac{\partial L_{\gamma}^{\alpha}(c)}{\partial c^{\nu}}L_{\beta}^{\nu}(c) = c_{\gamma\beta}^{\delta}\, L_{\delta}^{\alpha}(c) , \tag{IV.53}$$

$$\frac{\partial L_{\nu}^{-1\delta}(b)}{\partial b^{\gamma}} - \frac{\partial L_{\gamma}^{-1\delta}(b)}{\partial b^{\nu}} = c_{\alpha\beta}^{\delta}\, L_{\nu}^{-1\alpha}(b)L_{\gamma}^{-1\beta}(b) . \tag{IV.54}$$

Equation (IV.54) 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

58

provided the structure constants satisfy (IV.35). Therefore (IV.54) can be solved for $L^{-1}(b)$. With this value of L^{-1} (IV.44) 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 a remarkable theorem: given the Lie algebra of generators

$$[X_\alpha, X_\beta] = c_{\alpha\beta}^\gamma X_\gamma \qquad\qquad (IV.55)$$

where the structure constants $c_{\alpha\beta}^\gamma$ satisfy (IV.34) and (IV.35), or equivalently

$$[X_\alpha, X_\beta] = -[X_\beta, X_\alpha], \qquad\qquad (IV.56)$$

$$\big[X_\alpha, [X_\beta, X_\gamma]\big] + [X_\beta, [X_\gamma, X_\alpha]] + [X_\gamma, [X_\alpha, X_\beta]] = 0, \qquad\qquad (IV.57)$$

then 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 semi-simple algebra and a solvable algebra. It is not possible to pursue this within the confines of a single lecture like this.

References

[1] L. P. Eisenhart, *Continuous Groups of Transformations*, (Dover, New York, 1961).

[2] F. Gürsey, "Introduction to Group Theory," in *Relativity, Groups and Topology*, eds. C. DeWitt and B. DeWitt (Gordon-Breach, New York, 1963) pp. 91-161.

[3] M. Hammermesh, *Group Theory and its Applications to Physical Problems*, (Addison-Wesley, Reading, MA, 1962).

[4] C. Loewner, *Theory of Continuous Groups*, notes by H. Flanders and M. Protter, (MIT Press, Cambridge, MA, 1971).

[5] R. Salmon, "Proceedings of the 1990 Geophysical Fluid Dynamics Summer School," Woods Hole Oceanographic Institution, WHOI report (1990).

B. Reduction—Realization

Reduction is a procedure for obtaining from a given Hamiltonian system one of smaller dimension. The idea dates to Poincaré and Cartan. It is an example 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 collection of functions defined on phase space.[*] There are two parts to reduction: kinematics and dynamics. The kinematic part is concerned with the use of special variables that have a certain closure property, while the dynamic part refers to a symmetry of the Hamiltonian, viz. that the Hamiltonian be expressible in terms of the special variables. The symmetry 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 practice 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 section we discuss reduction in general terms for finite systems, and then consider a reduction that we term *standard reduction*, where the new variables are linear in the momenta. This is followed by two examples, the free rigid body and the ideal fluid, both of which are standard reductions. In the next section 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

$$[f,g] = \frac{\partial f}{\partial z^i} J_c^{ij} \frac{\partial g}{\partial z^j} \qquad i,j = 1, 2, \cdots, 2N, \tag{IV.58}$$

where recall

$$(J_c^{ij}) = \begin{bmatrix} 0_N & I_N \\ -I_N & 0_N \end{bmatrix}, \tag{IV.59}$$

and

$$z = (q, p), \tag{IV.60}$$

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

$$f(z) = \overline{f}(w(z)). \tag{IV.61}$$

Differentiation of (IV.61) yields,

$$\frac{\partial f}{\partial z^i} = \frac{\partial \overline{f}}{\partial w^\alpha} \frac{\partial w^\alpha}{\partial z^i}, \tag{IV.62}$$

---

[*]See again Sudarshan (1963) and Sudarshan and Mukunda (1974), *l.c.*

60

which upon insertion into (IV.58) gives

$$[f, g] = \frac{\partial \bar{f}}{\partial w^\alpha} \frac{\partial g}{\partial w^\beta} \left(\frac{\partial w^\alpha}{\partial z^i} J_c^{ij} \frac{\partial w^\beta}{\partial z^j} \right) , \tag{IV.63}$$

where we have dropped the "overbar." The quantity

$$J^{\alpha\beta} := \frac{\partial w^\alpha}{\partial z^i} J_c^{ij} \frac{\partial w^\beta}{\partial z^j} \tag{IV.64}$$

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 a Lie algebra realization composed of the functions w and the Poisson bracket.

In order for functions of w together with the bracket

$$[f, g] = \frac{\partial f}{\partial w^\alpha} \frac{\partial g}{\partial w^\beta} J^{\alpha\beta}(w) , \tag{IV.65}$$

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

$$S^{\alpha\beta\gamma}(w) := J^{\alpha\delta} \frac{\partial J^{\beta\gamma}}{\partial w^\delta} + J^{\gamma\delta} \frac{\partial J^{\alpha\beta}}{\partial w^\delta} + J^{\beta\delta} \frac{\partial J^{\gamma\alpha}}{\partial w^\delta} = 0 . \tag{IV.66}$$

(Recall Lecture III.) Substituting (IV.64) into (IV.66) gives

$$S^{\alpha\beta\gamma}(w) = \frac{\partial w^\alpha}{\partial z^i} J^{ij} \frac{\partial w^\delta}{\partial z^j} \frac{\partial}{\partial w^\delta} \left(\frac{\partial w^\beta}{\partial z^k} J^{kl} \frac{\partial w^\gamma}{\partial z^l} \right) + \cdots \tag{IV.67}$$

$$= \frac{\partial w^\alpha}{\partial z^i} J^{ij} \frac{\partial}{\partial z^j} \left(\frac{\partial w^\beta}{\partial z^k} J^{kl} \frac{\partial w^\gamma}{\partial z^l} \right) + \cdots \tag{IV.68}$$

$$= [w^\alpha, [w^\beta, w^\gamma]] + [w^\beta, [w^\gamma, w^\alpha]] + [w^\gamma, [w^\alpha, w^\beta]] = 0 , \tag{IV.69}$$

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 complete, 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 $\overline{H}(w)$ such that

$$H(z) = \overline{H}(w) . \tag{IV.70}$$

Equation (IV.70) is in fact a statement of symmetry. 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 following special form:

$$w^\alpha = A^{\alpha i}(q) p_i \,. \tag{IV.71}$$

Writing out (IV.64)

$$J^{\alpha\beta} = \frac{\partial w^\alpha}{\partial p^i}\frac{\partial w^\beta}{\partial q^i} - \frac{\partial w^\beta}{\partial p^i}\frac{\partial w^\alpha}{\partial q^i} \,, \tag{IV.72}$$

and inserting (IV.72) into (IV.71) yields

$$J^{\alpha\beta} = A^{\alpha i}\frac{\partial A^{\beta j}}{\partial q^i}p_j - A^{\beta i}\frac{\partial A^{\alpha j}}{\partial q^i}p_j$$

$$= \left(A^{\alpha i}\frac{\partial A^{\beta j}}{\partial q^i} - A^{\beta i}\frac{\partial A^{\alpha j}}{\partial q^i}\right)p_j \,. \tag{IV.73}$$

Closure occurs if constant numbers $c^{\alpha\beta}_\gamma$ can be found such that

$$\left(A^{\alpha i}\frac{\partial A^{\beta j}}{\partial q^i} - A^{\beta i}\frac{\partial A^{\alpha j}}{\partial q^i}\right)p_j = c^{\alpha\beta}_\gamma A^{\gamma j}p_j = c^{\alpha\beta}_\gamma w^\gamma \,. \tag{IV.74}$$

The form of (IV.74) may ring a bell. Recall the discussion in Section A where we talked about integrability and obtained the Maurer-Cartan equation. From Eq. (IV.53) it is clear that if the A's are chosen to be the components of the infinitesimal generators of some Lie algebra, then Eq. (IV.74) holds, with the constant numbers $c^{\alpha\beta}_\gamma$ being the structure constants.

You may have noticed that above the structure constants have two covariant indices and one contravariant index, which is the opposite of that of Section A. Technically, above we have considered the dual of the algebra—the algebra of linear functionals. Evidently there is more to this story than we are telling you. For now, we emphasize that the important thing is that (IV.53) have a solution.

Since reduction involves a symmetry and symmetries are related to constants of motion, it should come as no surprise that a general expression for constants of motion, which are, of course the Casimir invariants, comes along with the reduction framework. A clean way of seeing this is afforded by *triple bracket* dynamics.*

This construction begins by considering a semisimple Lie algebra with structure constants c^k_{ij} and metric tensor[†] g_{ij} which is given by

$$g_{ij} = -c^k_{il}\,c^l_{jk} \,. \tag{IV.75}$$

This quantity can be used to raise and lower indices. (Note the minus sign is introduced here to make g_{ij} positive for e.g. the rotation group.)

---

*Bialynicki-Birula and Morrison, (1991), *Ref. IV C* Triple bracket dynamics is a generalization of a formalism due to Y. Nambu, Phys. Rev. D **7**, 2405 (1973).

†This is also called either the trace form or Killing form. See e.g. Jacobson, *l.c.*

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,

$$c^{ijk} = g^{im} g^{jn} c^k_{mn} \,,$$

(IV.76)

from which the following triple bracket is constructed:

$$[A, B, C] = -c^{ijk} \frac{\partial A}{\partial w^i} \frac{\partial B}{\partial w^j} \frac{\partial C}{\partial w^k} \,.$$

(IV.77)

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

$$C := \frac{1}{2} g_{ij} w^i w^j \,,$$

(IV.78)

into one of the slots of the triple bracket, i.e.

$$[f, g, h] = [f, g] \,,$$

(IV.79)

Due to this relationship time evolution can be represented as follows:

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

(IV.80)

where f is an arbitrary dynamical variable. In this formulation the dynamics is determined by two generating functions, the Hamiltonian H and the Casimir C and 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 dimensional 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 $\chi = (\chi_1, \chi_2, \chi_3)$, which are defined by Figure 3 below.

Evidently, the rotation matrix, $R(\chi)$, 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 important formulae relating the angular velocities, relative to a set of
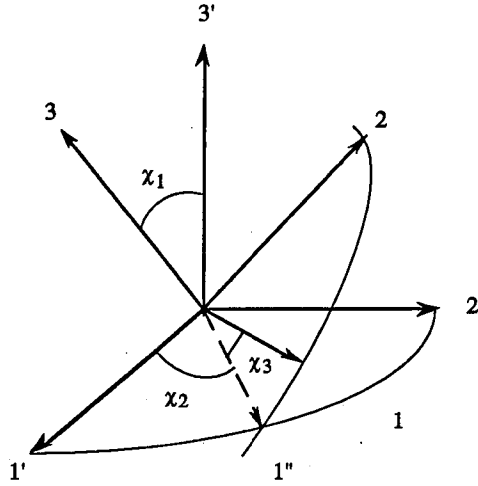
Figure 3:

Cartesian coordinates fixed in the body, to the time rate of change of the Euler angles:

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

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

$$\omega_3 = \dot{\chi}_3 + \dot{\chi}_2 \cos \chi_2 \,. \tag{IV.81}$$

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 inertial tensor is diagonal, the so-called principal axes. In these coordinates the Lagrangian is deceptively simple,

$$L(\chi, \dot{\chi}) = \tfrac{1}{2} \left(I_1 \omega_1^2 + I_2 \omega_2^2 + I_3 \omega_3^2 \right) , \tag{IV.82}$$

it being merely the kinetic energy since there are no external forces. Note however, upon insertion of (IV.81) 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

$$p_i = \frac{\partial L}{\partial \dot{\chi}_i} = \frac{\partial L}{\partial \omega_j} \frac{\partial \omega_j}{\partial \dot{\chi}_i} = \ell_j \frac{\partial \omega_j}{\partial \dot{\chi}_i} , \tag{IV.83}$$

i.e.

$$p_i = A_{ij}^{-1}(\chi) \, \ell_j , \tag{IV.84}$$

where

$$(A^{-1}) = \begin{pmatrix} \cos \chi_3 & -\sin \chi_3 & 0 \\ \sin \chi_1 \sin \chi_3 & \sin \chi_1 \cos \chi_3 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \tag{IV.85}$$

64

and the angular momentum $\ell_i := I_i \omega_i$ (not summed). Comparing (IV.81) with (IV.84) reveals that

$$\omega_i = D_{ij} \dot{\chi}_j, \qquad (IV.86)$$

where $D^T = A^{-1}$ with "T" indicating transpose. If Ω is defined to be the antisymmetric matrix composed of the three components of ω, then an important way to write (IV.85) is as follows:

$$\Omega = R^T \dot{R}. \qquad (IV.87)$$

This form is analogous to that of the map from Lagrange to Euler variables. I will leave it for you to work this out.

The inverse of (IV.84) is given by

$$\ell_j = A_{ji}(\chi) \, p_i, \qquad (IV.88)$$

where

$$(A) = \frac{1}{\sin \chi_1} \begin{pmatrix} \sin \chi_1 \cos \chi_3 & \sin \chi_3 & -\sin \chi_3 \cos \chi_1 \\ -\sin \chi_1 \sin \chi_3 & \cos \chi_3 & -\cos \chi_3 \cos \chi_1 \\ 0 & 0 & \sin \chi_1 \end{pmatrix}. \qquad (IV.89)$$

This is the standard reduction formula of the form of (IV.71).

Upon effecting the Legendre transform, the Hamiltonian is obtained:

$$H(p, \chi) = p_i \dot{\chi}_i - L = \tfrac{1}{2} \ell_k \omega_k = \tfrac{1}{2} \sum_k \frac{\ell_k^2}{I_k} = \tfrac{1}{2} \sum_k \frac{1}{I_k} A_{ki} A_{kj} \, p_i p_j, \qquad (IV.90)$$

which obviously possesses the necessary symmetry of (IV.70).

It remains to show whether or not the variables ℓ_i allow a reduction. To see if this is the case consider $[\ell_i, \ell_j]$, which upon insertion of (IV.88) becomes

$$[\ell_i, \ell_j] = A_{rs}^{-1} \ell_s \left(\frac{\partial A_{ir}}{\partial \chi_k} A_{jk} - \frac{\partial A_{jr}}{\partial \chi_k} A_{ik} \right), \qquad (IV.91)$$

as expected from the results of the previous subsection. Since the left-hand side of (IV.91) is difficult to evaluate, we make use of

$$\frac{\partial A_{ir}}{\partial \chi_k} A_{rs}^{-1} = -A_{ir} \frac{\partial A_{rs}^{-1}}{\partial \chi_k}, \qquad (IV.92)$$

which follows upon differentiating

$$A_{ir} A_{rs}^{-1} = \delta_{is}, \qquad (IV.93)$$

to obtain

$$[\ell_i, \ell_j] = \left(\ell_s \frac{\partial A_{rs}^{-1}}{\partial \chi_k} \right) (A_{jr} A_{ik} - A_{ir} A_{jk}). \qquad (IV.94)$$

The matrix in the first parentheses is not too difficult to calculate. The evaluation of the second parenthesis 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 and expected result:

$$[\ell_i, \ell_j] = -\epsilon_{ijk} \ell_k. \qquad (IV.95)$$

65

4 Reduction for the Ideal Fluid: Lagrangian → Eulerian Variables

Now consider reduction for the ideal fluid, which amounts to the transformation from Lagrangian to Eulerian variables. In the Lagrangian variable description of Lecture II we had the Hamiltonian

$$H[\pi, q] = \int_D \left[\frac{\pi^2}{2\rho_0} + \rho_0 U(s_0, \rho_0/\mathcal{J}) \right] d^3a \qquad \text{(IV.96)}$$

which together with the canonical Poisson bracket,

$$[F, G] = \int_D \left[\frac{\delta F}{\delta q} \cdot \frac{\delta G}{\delta \pi} - \frac{\delta G}{\delta q} \cdot \frac{\delta F}{\delta \pi} \right] d^3a. \qquad \text{(IV.97)}$$

produces the ideal fluid equations of motion. For the moment, let us forget about the Hamiltonian structure and just consider the change from (q, π), the Lagrangian canonically conjugate pair, to (ρ, σ, M), the Eulerian non-canonical variables. Recall from Lecture II that

$$\rho(r, t) = \int_D \rho_0(a)\, \delta(r - q(a, t))\, d^3a,$$

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

$$M(r, t) = \int_D \pi\, \delta(r - q(a, t))\, d^3a. \qquad \text{(IV.98)}$$

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 we need it to calculate variations of

$$F[q, \pi] = \overline{F}[\rho, \sigma, M]. \qquad \text{(IV.99)}$$

In (IV.99) we are supposing that F obtains its q and π dependence through some functional $\overline{F}$ of (ρ, σ, M). The functionals F and $\overline{F}$ are defined on different functions, which are themselves defined on different (space-like) domains, a and x, respectively.

Consider the variation of F,

$$\delta F = \int_D \left[\frac{\delta F}{\delta q} \cdot \delta q + \frac{\delta F}{\delta \pi} \cdot \delta \pi \right] d^3a$$

$$= \int_D \left[\frac{\delta \overline{F}}{\delta \rho} \delta\rho + \frac{\delta \overline{F}}{\delta \sigma} \delta\sigma + \frac{\delta \overline{F}}{\delta M} \cdot \delta M \right] d^3r. \qquad \text{(IV.100)}$$

Note that the two domains of integration coincide, although the variables of integration have different names. We will now try to find the functional derivatives with respect to the Lagrangian fields in terms of the Eulerian fields. This will allow us to express the bracket in Eulerian fields. The variations of the Eulerian fields induced by a variation of the Lagrangian

fields are

$$\delta \rho = - \int_D \rho_0(a) \, \nabla \delta(r - q) \cdot \delta q \, d^3 a \,,$$

$$\delta \sigma = - \int_D \sigma_0(a) \, \nabla \delta(r - q) \cdot \delta q \, d^3 a \,,$$

$$\delta M = \int_D [\delta \pi \, \delta(r - q) - \pi \nabla \delta(r - q) \cdot \delta q] \, d^3 a \,. \qquad \text{(IV.101)}$$

Above (and below) the ∇-operator operates on the r-dependence. Inserting (IV.101) into (IV.100), interchanging the order of integration, and equating the coefficients of δq and $\delta \pi$ implies

$$\frac{\delta F}{\delta q} = - \int_D \left[\rho_0 \frac{\delta \overline{F}}{\delta \rho} + \sigma_0 \frac{\delta \overline{F}}{\delta \sigma} + \pi \cdot \frac{\delta \overline{F}}{\delta M} \right] \nabla \delta(r - q) \, d^3 r$$

$$= \int_D \left[\rho_0 \nabla \frac{\delta \overline{F}}{\delta \rho} + \sigma_0 \nabla \frac{\delta \overline{F}}{\delta \sigma} + \pi_i \nabla \frac{\delta \overline{F}}{\delta M_i} \right] \delta(r - q) \, d^3 r \qquad \text{(IV.102)}$$

$$\frac{\delta F}{\delta \pi} = \int_D \frac{\delta \overline{F}}{\delta M} \delta(r - q) \, d^3 r = \left. \frac{\delta \overline{F}}{\delta M} \right|_{r=q} d^3 r =: \frac{\delta \overline{F}}{\delta M'} \,, \qquad \text{(IV.103)}$$

where the second formula is obtained after integration by parts, assuming the boundary terms vanish. Inserting (IV.102) and (IV.103), for both F and G, into (IV.98), yields

$$\{\overline{F}, \overline{G}\} = \int_D \delta(r - q) \left\{ \rho_0 \nabla \frac{\delta \overline{F}}{\delta \rho} \cdot \frac{\delta \overline{G}}{\delta M'} + \sigma_0 \nabla \frac{\delta \overline{F}}{\delta \sigma} \cdot \frac{\delta \overline{G}}{\delta M'} + \frac{\delta \overline{G}}{\delta M'_j} \pi_i \frac{\partial}{\partial x_j} \frac{\delta \overline{F}}{\delta M_i} \right.$$

$$\left. - \rho_0 \nabla \frac{\delta \overline{G}}{\delta \rho} \cdot \frac{\delta \overline{F}}{\delta M'} - \sigma_0 \nabla \frac{\delta \overline{G}}{\delta \sigma} \cdot \frac{\delta \overline{F}}{\delta M'} - \frac{\delta \overline{F}}{\delta M_j} \pi_i \frac{\partial}{\partial_j} \frac{\delta \overline{G}}{\delta M_i} \right\} d^3 a \,. \qquad \text{(IV.104)}$$

After interchanging the order of integration, the integral over $d^3 a$ can be carried out,

$$\{F, G\} = - \int_D \left[M_i \left(\frac{\delta F}{\delta M_j} \frac{\partial}{\partial x_j} \frac{\delta G}{\delta M_i} - \frac{\delta G}{\delta M_j} \frac{\partial}{\partial x_j} \frac{\delta F}{\delta M_i} \right) + \rho \left(\frac{\delta F}{\delta M} \cdot \nabla \frac{\delta G}{\delta \rho} - \frac{\delta G}{\delta M} \cdot \nabla \frac{\delta F}{\delta \rho} \right) \right.$$

$$\left. + \sigma \left(\frac{\delta F}{\delta M} \cdot \nabla \frac{\delta G}{\delta \sigma} - \frac{\delta G}{\delta M} \cdot \nabla \frac{\delta F}{\delta \sigma} \right) \right] d^3 r \,. \qquad \text{(IV.105)}$$

Equation (IV.105) is the noncanonical bracket that was given in Lecture III. It is a bracket expression in terms of Eulerian functionals, that is ones that depend on Eulerian fields, integrated over the Eulerian spatial domain.

Above we have considered the transformation of variables only. This can be viewed as kinematics. To complete the Hamiltonian description in terms of Eulerian variables we must

obtain the Hamiltonian in terms of ρ, σ, and M. The reduction we have performed can only yield dynamics if we can find a Hamiltonian, $\overline{H}$ that satisfies

$$H[q, \pi] = \overline{H}[\rho, \sigma, M], \qquad \text{(IV.106)}$$

upon substitution of Eqs. (IV.98). In general this is not possible, but for the ideal fluid the one that does the trick is of course

$$\overline{H}[\rho, \sigma M] = \int_D \left[\frac{1}{2} \frac{M^2}{\rho} + \rho U(\rho, \frac{\sigma}{\rho}) \right] d^3 r . \qquad \text{(IV.107)}$$

Note that while the reduction of the bracket only depends upon the definitions of ρ, σ, and M, the corresponding reduction of the Hamiltonian involves a symmetry, namely the independence of the Hamiltonian under fluid particle relabelling.

References

[1] E. Sudarshan, "Principles of Classical Dynamics," Report NYO 10250, University of Rochester, Department of Physics and Astronomy (1963).

[2] E. Sudarshan and N. Mukunda, *Classical Dynamics: a Modern Perspective* (John Wiley, New York, 1974).

[3] P. J. Morrison, "Hamiltonian field description of two-dimensional vortex fluids an guiding center plasmas," Princeton University Plasma Physics Laboratory Report, PPPL-1783 (1981). Available as American Institute of Physics Document No. PAPS-PFBPE-04-771, AIP Auxiliary Publication Service, 335 East 45th Street, New York, NY 10017.

[4] J. E. Marsden, A. Weinstein, T. Ratiu, R. Schmid, and R. Spencer, Proc. IUTAM-IS1MM Symposium on Modern Developments in Analytical Mechanics, Actti della Acad. della Sc. di Torino **117**, 289 (1983).

[5] J. E. Marsden and A. Weinstein, Rep. Math. Phys. **5**, 121-130 (1974).

[6] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt, *Dynamics III: Mathematical Aspects of Classical and Celestial Mechanics* (Springer-Verlag, Berlin, 1980).

C. Clebsch Variables

In this section we consider *Clebsch variables*. These are canonical variables that reduce to noncanonical variables where, as mentioned above, the noncanonical variables are bilinear in the momenta and configuration space coordinates. We will 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.[*]

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 will present the Clebsch reduction from an historical, if not logical, point of view. Suppose we have a noncanonical Lie-Poisson bracket of the following form:

$$[f,g] = -w^k c_k^{ij} \frac{\partial f}{\partial w^i} \frac{\partial f}{\partial w^j} , \tag{IV.108}$$

where c_k^{ij} 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 (IV.71), reduces to this form. Now we turn things around and ask the question, can we inflate (IV.108) and obtain other canonical descriptions. Here we have used the word inflation, since we are not talking about the canonical description on the symplectic leaves of Lecture III, 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:

$$w^i = c_k^{ij} q^k p_j , \tag{IV.109}$$

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,

$$\{f,g\} = \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} , \tag{IV.110}$$

---

[*]A. Clebsch, J. Reine Angew. Math. **54**, 293 (1857); *ibid.* **56**, 1 (1859).

the bracket in terms of w is obtained by a reduction. This is seen upon substituting

$$\frac{\partial f}{\partial p_i} = \frac{\partial f}{\partial w^j} c_k^{ji} q^k$$

$$\frac{\partial f}{\partial q^i} = \frac{\partial f}{\partial w^j} c_i^{jk} p_k \,,$$

(IV.111)

into (IV.110)

$$\frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} = p_t\, q^r \left(c_k^{it} c_r^{jk} - c_k^{jt} c_r^{ik} \right) \frac{\partial f}{\partial w^i}\frac{\partial g}{\partial w^j}$$

$$= -w^k c_k^{ij} \frac{\partial f}{\partial w^i}\frac{\partial g}{\partial w^j} \,,$$

(IV.112)

where the last equality follows upon making use of the Jacobi identity for the structure constants, (IV.35).

Given any noncanonical system in terms of the w's one can obtain a canonical system of equations in terms of the Clebsch q and p; when these are solved for $q(t)$ and $p(t)$ then the w constructed according to (IV.109) is a solution of the noncanonical system.

2 Clebsch Variables for Infinite Systems

Here we will be a bit formal and define things in somewhat general terms. First we will denote by $\langle\,,\,\rangle$ a pairing between a vector space and its dual. We will, for now, leave the particular form of this unspecified, but we have in the back of our mind an integration like that in (III.36). The first slot of $\langle\,,\,\rangle$ can be thought of as an infinite dimensional analogue of the finite dimensional "up" indices, while the second slot is the analogue of the "down" indices. We will refer to elements of the first slot as belonging to $\mathfrak{g}$ and those of the second slot, the dual, as belonging to $\mathfrak{g}^*$. In general the pairing is not symmetric.

In terms of the pairing, noncanonical Lie-Poisson brackets have the following compact form:

$$\{F,G\} = - \langle \chi, [F\chi, G\chi]\rangle \,,$$

(IV.113)

where $[\,,\,]$ is a Lie algebra product, which takes $\mathfrak{g}^* \times \mathfrak{g}^* \to \mathfrak{g}^*$, and we have introduced the shorthand

$$F\chi := \frac{\delta F}{\delta \chi}, \qquad G\chi := \frac{\delta G}{\delta \chi}, $$

(IV.114)

which 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:

$$\langle \chi, [f,g]\rangle =: \left\langle [\chi,g]^\dagger, f \right\rangle \,,$$

(IV.115)

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 (IV.109) is given by

$$\chi = [Q,\Pi]^\dagger \,.$$

(IV.116)

70

In order to effect the reduction, consider a variation of (IV.116),

$$\delta \mathcal{X} = [\delta Q, \Pi]^\dagger + [Q, \delta \Pi]^\dagger , \tag{IV.117}$$

which is used to relate functional derivatives as follows:

$$\begin{aligned}
\delta F &= \langle \delta \mathcal{X}, F_\mathcal{X} \rangle \\
&= \left\langle [\delta Q, \Pi]^\dagger + [Q, \delta \Pi]^\dagger, F_\mathcal{X} \right\rangle \\
&= \langle \delta Q, F_Q \rangle + \langle F_\Pi, \delta \Pi \rangle .
\end{aligned} \tag{IV.118}$$

Manipulation of the second equality of (IV.118) yields

$$\begin{aligned}
\delta F &= \langle \delta Q, [F_\mathcal{X}, \Pi] \rangle + \langle Q, [F_\mathcal{X}, \delta \Pi] \rangle \\
&= \langle \delta Q, [F_\mathcal{X}, \Pi] \rangle - \langle Q, [\delta \Pi, F_\mathcal{X}] \rangle \\
&= \langle \delta Q, [F_\mathcal{X}, \Pi] \rangle - \left\langle [Q, F_\mathcal{X}]^\dagger, \delta \Pi \right\rangle ,
\end{aligned} \tag{IV.119}$$

where the antisymmetry of $[\,,\,]$ and the definition of $[\,,\,]^\dagger$ have been used. Upon comparing the last equality of (IV.119) with the last equality of (IV.118) we obtain

$$F_Q = [F_\mathcal{X}, \Pi] , \qquad F_\Pi = -[Q, F_\mathcal{X}]^\dagger . \tag{IV.120}$$

The canonical bracket in terms of Q and Π can be written as

$$\{F, G\} = \langle G_\Pi, F_Q \rangle - \langle F_\Pi, G_Q \rangle . \tag{IV.121}$$

Inserting (IV.120), produces

$$\begin{aligned}
\{F, G\} &= - \left\langle [Q, G_\mathcal{X}]^\dagger, [F_\mathcal{X}, \Pi] \right\rangle + \left\langle [Q, F_\mathcal{X}]^\dagger, [G_\mathcal{X}, \Pi] \right\rangle \\
&= \left\langle Q, \big[[G_\mathcal{X} \Pi], F_\mathcal{X} \big] + \big[[\Pi, F_\mathcal{X}], G_\mathcal{X} \big] \right\rangle \\
&= \left\langle Q, \big[[G_\mathcal{X}, F_\mathcal{X}], \Pi \big] \right\rangle = - \left\langle [Q, \Pi]^\dagger, [F_\mathcal{X}, G_\mathcal{X}] \right\rangle \\
&= - \langle \mathcal{X}, [F_\mathcal{X}, G_\mathcal{X}] \rangle ,
\end{aligned} \tag{IV.122}$$

where use has been made of the Jacobi identity of $[\,,\,]$.

3 Fluid Examples

Now consider two examples from fluid mechanics: the first is the two-dimensional Euler equation, while the second is related to the three-dimensional ideal fluid.

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 2-D Euler noncanonical bracket, which was given in Lecture III, shares this property. This is clear from the "fgh" identity of (III.78), from which we also observe that

$$[f,g]^{\dagger} = -[f,g].\tag{IV.123}$$

Here no distinction is made between the vector space and its dual. For this case

$$[f,g] = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x},\tag{IV.124}$$

and

$$\langle\,,\,\rangle = \int_{D} d^2 r.\tag{IV.125}$$

The Clebsch variables $Q(r,t)$ and $\Pi(r,t)$ are related to the scalar vorticity via

$$\omega(r,t) = [\Pi,Q],\tag{IV.126}$$

and the reduction from these canonical variables to the 2-D 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 (IV.126) into the Hamiltonian $H[\omega]$ of (III.74) and then calculate

$$\frac{\partial Q}{\partial t} = \frac{\delta H}{\delta \Pi},\qquad\qquad \frac{\partial \Pi}{\partial t} = -\frac{\delta H}{\delta Q}.\tag{IV.127}$$

The other way is to insert (IV.126) directly into the equation of motion for ω, viz.

$$\frac{\partial \omega}{\partial t} = -[\psi,\omega],\tag{IV.128}$$

[cf. (III.72)] and then manipulate as follows:

$$\frac{\partial \omega}{\partial t} = \left[\frac{\partial \Pi}{\partial t},Q\right] + \left[\Pi,\frac{\partial Q}{\partial t}\right]$$

$$= -\left[\psi,[\Pi,Q]\right] = \left[\Pi,[Q,\psi] + [Q,[\psi,\Pi]\right],\tag{IV.129}$$

where the Jacobi identity was used to obtain the last equality. From (IV.129) we obtain

$$\left[\frac{\partial \Pi}{\partial t} + [\psi,\Pi],Q\right] + \left[\Pi,\frac{\partial Q}{\partial t} + [\psi,Q]\right] = 0,\tag{IV.130}$$

---

*The careful reader will notice a sign discrepancy. There is a story that goes with this sign, but unfortunately we are not able to tell it here.

which is satisfied if

$$\frac{\partial \Pi}{\partial t} = -[\psi, \Pi] + \frac{\partial \Upsilon}{\partial Q}$$

$$\frac{\partial Q}{\partial t} = -[\psi, Q] - \frac{\partial \Upsilon}{\partial \Pi}, \qquad \text{(IV.131)}$$

where the terms involving Υ point to the gauge ambiguity present in (IV.126), something that will not be discussed further here. If $Q(r,t)$ and $\Pi(r,t)$ are solutions of (IV.131), then the $\omega = [\Pi, Q]$ constructed from these solutions are solutions of (IV.128).

Turn now to the following bracket, which is a portion of the noncanonical bracket for the ideal fluid, [cf. (III.89)]:

$$\{F, G\} = -\int_D M_i \left(\frac{\delta F}{\delta M_j} \frac{\partial}{\partial x_j} \frac{\delta G}{\delta M_i} - \frac{\delta G}{\delta M_j} \frac{\partial}{\partial x_j} \frac{\delta F}{\delta M_i} \right) d^3 r$$

$$=: -\langle M, [F_M, G_M] \rangle . \qquad \text{(IV.132)}$$

It is obvious that this bracket will satisfy the Jacobi identity if (III.89) does. The inner bracket in this case is given by

$$[f, g]_i = f_j \frac{\partial g_i}{\partial x_j} - g_j \frac{\partial f_i}{\partial x_j}, \qquad \text{(IV.133)}$$

where, evidently, f and g now have three components. Integration by parts and neglect of surface terms yields

$$[\chi, g]_j^\dagger = \chi_i \frac{\partial g_i}{\partial x_j} + \frac{\partial (\chi_j g_i)}{\partial x_i}, \qquad \text{(IV.134)}$$

whence the Clebsch variables are seen to be related to M by

$$M_j = Q_i \frac{\partial \Pi_i}{\partial x_j} + \frac{\partial (Q_j \Pi_i)}{\partial x_i}. \qquad \text{(IV.135)}$$

In reality the decomposition above is not quite that due to Clebsch, whose transformation did not have the second term of (IV.135). However, it is closely related to that introduced for MHD.* In fact the reduction occurs without this last term; it also occurs with the last term with opposite sign. Also, it is not important that Q and Π have three components. Some of this will be discussed 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 cannot discuss this is any kind of detail here so the

---

*P. J. Morrison and J. M. Greene, Phys. Rev. Lett. **48**, 569 (1982) and Morrison (1982), *Ref. IV C.*

interested reader is referred to the references.[†] However, this notion makes its way up to Lie algebras and thus to Lie-Poisson brackets, a case of which we will discuss (briefly) here.

Suppose the functional F in (IV.118), in addition to its $\mathcal{X}$ dependence, depends upon Q, i.e. $\overline{F}[\mathcal{X}, Q] = F[Q, \Pi]$. (We have included the overbar now, as in Lecture II, to avoid confusion.) Effecting the chain rule with this additional dependence yields

$$F_Q = [\overline{F}_\mathcal{X}, \Pi] + \overline{F}_Q \,, \tag{IV.136}$$

which upon substitution into (IV.121) produces instead of (IV.122), the following:

$$\{\overline{F}, \overline{G}\} = -\left\langle \mathcal{X}, [\overline{F}_\mathcal{X}, \overline{G}_\mathcal{X}]\right\rangle - \left\langle [Q, \overline{G}_\mathcal{X}]^\dagger, \overline{F}_Q\right\rangle + \left\langle [Q, \overline{F}_\mathcal{X}]^\dagger, \overline{G}_Q\right\rangle$$

$$= -\left\langle \mathcal{X}, [\overline{F}_\mathcal{X}, \overline{G}_\mathcal{X}]\right\rangle + \left\langle Q, [\overline{G}_\mathcal{X}, \overline{F}_Q] - [\overline{F}_\mathcal{X}, \overline{G}_Q]\right\rangle \,, \tag{IV.137}$$

where the second equality follows from manipulations similar to those performed above.

Many systems possess brackets of this (and similar) form(s). The rigid body in a gravitational field is an example of finite dimension. An example of infinite dimension, which was first given in the context of reduced MHD,[*] but also occurs in fluid mechanics, is the semidirect product extension of the noncanonical bracket for the 2-D Euler fluid. For this example one simply interprets (IV.137) using (IV.124) and (IV.125).

5 Other Clebsch Reductions: That for the Ideal Fluid

In this final subsection we present some other forms of Clebsch reductions. The first is another way to reduce to the reduced MHD bracket of above. This emphasizes the fact that reductions are not unique. Following this we show another way to reduce to the portion of the ideal fluid bracket, also treated above. Finally we reduce to the complete ideal fluid noncanonical bracket. This final transformation is the one actually due to Clebsch.

Suppose we have a system with canonical variables $(Q_i(r, t), \Pi_i(r, t))$, where $i = 1, 2$ and $r = (x, y)$. The canonical Poisson bracket is then

$$\{F, G\} = \int_D (F_Q \cdot G_\Pi - G_Q \cdot F_\Pi) \, d^2r \,. \tag{IV.138}$$

The following transformation is a reduction:

$$\mathcal{X} = [Q_1, \Pi_1] + [Q_2, \Pi_2]$$

$$\psi = [Q_1, Q_2] \,, \tag{IV.139}$$

where $[\,,\,]$ is given by (IV.124). We leave it as an exercise to show via the chain rule that with (IV.139), (IV.138) reduces to a bracket of the form of (IV.137).

Now consider the portion of the fluid bracket discussed above in (IV.132), but now instead of (IV.135) we let

$$M = Q_i \nabla \Pi_i \,. \tag{IV.140}$$

[†]See e.g. Sudarshan and Mukunda (1974), *Ref. IV C*, J. E. Marsden and P. J. Morrison, Contemp. Math. **28**, 133 (1984).

[*]Morrison and Hazeltine (1984), *Ref. IV C*.

74

where $i = 1, 2, ...N$ and N is arbitrary. We also leave it as an exercise to show via the chain rule that with (IV.140), 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 (IV.132).

Finally, suppose in addition to (IV-14-3) that

$$\rho = Q_1, \qquad \sigma = Q_2. \qquad (IV.141)$$

We leave it as a *last* exercise to show via the chain rule that with (IV.140) and (IV.141), a canonical bracket in terms of $(Q_i(r,t), \Pi_i(r,t))$, reduces to the ideal 3-D fluid bracket of (III.89). One can choose N large enough to describe the velocity field of interest.

References

[1] H. Lamb, *Hydrodynamics* (Dover, New York, 1945) art. 167.

[2] P. J. Morrison, "Hamiltonian Field Description of Two-dimensional Vortex Fluids and Guiding Center Plasmas," Princeton University Plasma Physics Laboratory Report, PPPL-1783 (1981). Available as American Institute of Physics Document No. PAPS-PFBPE-04-771, AIP Auxiliary Pub. Service, 335 East 45th St., New York, NY 10017.

[3] P. J. Morrison, in *Mathematical Methods in Hydrodynamics and Integrability in Related Dynamical Systems*, eds. M. Tabor and Y. Treve, AIP Conference Proceedings 88, La Jolla, 1982, p. 13.

[4] D. D. Holm and B. Kupershmidt, Physica D **6**, 347 (1983).

[5] P. J. Morrison and R. D. Hazeltine, Physics of Fluids **27**, 886 (1984).

[6] I. Bialynicki-Birula and P. J. Morrison, Phys. Lett. A **158**, 453 (1991).

[7] E. Sudarshan and N. Mukunda, *Classical Dynamics: a Modern Perspective* (John Wiley, New York, 1974).

V. Stability and Hamiltonian Systems

This Lecture concerns notions of stability in Hamiltonian systems. In Section 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 that illustrates properties of NEM's is given, in which context simple Hamiltonian bifurcation theory is reviewed. Finally in this section, these ideas are applied to the ideal fluid in the Lagrangian variable description. Section B is concerned with stability in noncanonical Hamiltonian systems. The energy-Casimir method is described and two examples are given: a charged rigid body in an external magnetic field and the 2-D Euler equation. The examples exhibit a pathology related to the rank changing behavior of the cosymplectic form, that is discussed. In Section 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

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

where, as in Lecture III, we will not get into what is required of $V(z)$ for existence and uniqueness of solutions, but just assume everything is alright. An *equilibrium point*, z_e, is a type of solution of (V.1) 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 by the following:

> *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 don't we just say, if it starts in a set and stays in the set, then it is stable? The answer to this is illustrated in Figure 1, 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 rectangular region in which we have marked an initial condition by the symbol "×." Since trajectories move round and round on the circles of constant H, it is clear that in a short time the trajectory starting at × 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 = $ constant surface contained in N, then the trajectory will remain in N for all time. Thus, $H = $ constant surfaces serve as a handy means of defining subneighborhoods.
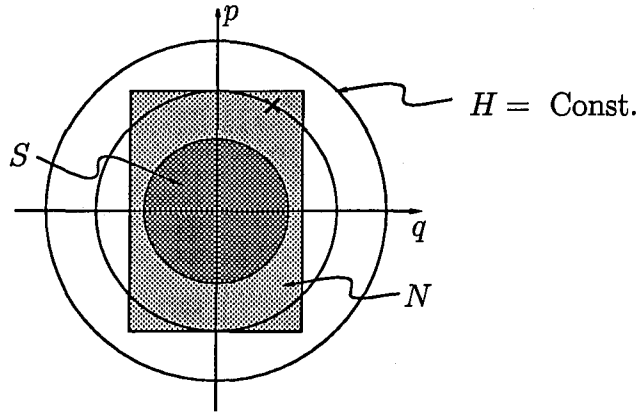
Figure 1:

Observe that the neighborhood N can be chosen to be *any* neighborhood N of z_e. We can make them 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,

$$\delta \dot{z}^i = \frac{\partial V^i}{\partial z^j} (z_e) \, \delta z^j \,, \qquad (\text{V.2})$$

where now $z(t) := z_e + \delta z$, and this dynamics is stable according to the above definition, we say that (V.2) or z_e is *linearly stable*.

One might think, since N can be made as small as we like, that these types of 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*, where 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 (V.2) 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. In the case where 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

77

a counterexample from the book of A. Wintner[*] shows this not to be the case. Consider

$$V(q) = \begin{cases} e^{-1/q^2} \cos(1/q) & q \neq 0 \\ 0 & q = 0. \end{cases} \qquad \text{(V.3)}$$

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 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 where 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-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 you can see that this is not true. The one degree-of-freedom system with potential

$$V(q) = \frac{q^4}{4} \qquad \text{(V.4)}$$

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 satisfy

$$\delta \dot{p} = 0, \qquad \delta \dot{q} = \delta p \qquad \text{(V.5)}$$

and thus

$$\delta p = \text{constant}, \qquad \delta q = \delta q_0 + \delta p\, t. \qquad \text{(V.6)}$$

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 deal of 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.[*] 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 have been used

---

[*]A. Wintner, *The Analytical Foundations of Celestial Mechanics* (Princeton University, Princeton, New Jersey, 1947).

[*]I. Bernstein, E. A. Frieman, M. D. Kruskal, and R. M. Kulsrud, Proc. Roy. Soc. A **244**, 17-40 (1958); Von K. Hain, R. Lüst, and A. Schlüter, Zeitschrift für Naturforschung A **12**, 833-841 (1957); G. Laval, C. Mercier, and R. Pellat, Nucl. Fusion **5**, 156-158 (1965).

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. It should be no surprise to you now that if in the vicinity of an equilibrium point surfaces of $H =$ constant 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 quantity is definite, i.e. there are no zero eigenvalues and they all 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[*] 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

$$H = \tfrac{1}{2}\omega_2(p_2^2 + q_2^2) - \tfrac{1}{2}\omega_1(p_1^2 + q_1^2) + \tfrac{1}{2}\alpha \left[2q_1 p_1 p_2 - q_2(q_1^2 - p_1^2)\right],\qquad\text{(V.7)}$$

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[†]. Also, with the inclusion of nonlinearity NEM's can be driven unstable. The example of Cherry demonstrates this; assuming $\alpha \neq 0$ and $\omega_2 = 2\omega_1$, (V.7) possesses a solution[‡] of the form

$$q_1 = \frac{\sqrt{2}}{\epsilon - \alpha t}\sin(\omega_1 t + \gamma), \qquad p_1 = -\frac{\sqrt{2}}{\epsilon - \alpha t}\cos(\omega_1 t + \gamma)$$

$$q_2 = -\frac{\sqrt{2}}{\epsilon - \alpha t}\sin(2\omega_1 t + 2\gamma), \qquad p_2 = -\frac{\sqrt{2}}{\epsilon - \alpha t}\cos(2\omega_1 t + 2\gamma). \qquad\text{(V.8)}$$

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

[*]T. M. Cherry, Trans. Cambridge Philos. Soc. **23**, 199 (1925)

[†]This is a fairly old idea that is sometimes called the Kelvin-Tait theorem. See W. Thompson and P. G. Tait, *Treatise on Natural Philosophy* (Cambridge University Press, Cambridge, 1921), part 1, p. 388.

[‡]See E. T. Whittaker, *Analytical Dynamics* (Cambridge University Press, London, 1937), Sec. 136, p. 101, but be careful because there are typographical errors.

explosive growth and is characteristic of systems that possess both NEM's and resonance. Another example is the well-known "three-wave" problem.[§] 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. Thus explosive behavior is to be expected when there is 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.

One might think that systems with NEM's are artifacts or unphysical, purely mathematical, oddities; this, however, is not the case. They occur in fluid and plasma systems[*] 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

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + G(\dot{y}x - \dot{x}y) + \tfrac{1}{2}k(x^2 + y^2)\,, \qquad (V.9)$$

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:

$$H = \tfrac{1}{2}(p_1^2 + p_2^2) + \omega_G(q_2 p_1 - q_1 p_2) + \tfrac{1}{2}(\omega_G^2 - \omega_k^2)(q_1^2 + q_2^2)\,, \qquad (V.10)$$

where the two time scales of the problem are determined by the frequencies

$$\omega_G := \frac{G}{m}\,, \qquad \omega_k := \sqrt{\frac{k}{m}}\,. \qquad (V.11)$$

Assuming $q_{1,2}$, $p_{1,2} \sim e^{i\omega t}$, it is easy to solve for eigenvalues,

$$\omega = \pm\omega_k(\sqrt{\varepsilon - 1} \pm \sqrt{\varepsilon})\,, \qquad (V.12)$$

where $\varepsilon := \omega_G^2/\omega_k^2$. 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

[§]See e.g. C. Kueny, "Nonlinear Instability and Chaos in Plasma Wave-Wave Interactions," Ph.D. Thesis, University of Texas at Austin (1993) and many references cited therein; See also D. Pfirsch, Phys. Rev. D **48**, 1428 (1993).

[*]In the context of MHD see J. M. Greene and B. Coppi, Phys. Fluids **8**, 1745 (1965); of fluids see R. A. Cairns, J. Fluid Mech. **92** 1 (1979), R. S. MacKay and P. G. Saffman, Proc. Roy. Soc. A **406**, 115 (1986), P. Ripa, Geophys. Astrophys. Fluid Dyn. **70**, 85 (1993); and of Vlasov theory see P. J. Morrison and D. Pfirsch, Phys. Rev. A **40**, 3998 (1989), Phys. Fluids B **2**, 1105 (1990) and *ibid.* **4**, 3038 (1992).

In Hamiltonian systems eigenvalues occur in doublets or quartets. Case (1) is the only stable case. It occurs in the example when $\varepsilon = \omega_G^2/\omega_k^2 > 1$, 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 $\varepsilon > 1$ there exists a canonical transformation $(q, p) \to (Q, P)$ that takes H into

$$H(Q, P) = -\tfrac{1}{2}\omega_s(P_s^2 + Q_s^2) + \tfrac{1}{2}\omega_f(P_f^2 + Q_f^2), \qquad (V.13)$$

which is the linear part of Cherry's Hamiltonian. The canonical transformation is effected by the following mixed variable generating function:

$$F_2(q_1, q_2, P_f, P_s) = c(q_1 P_s + q_2 P_f) + P_f P_s + \tfrac{1}{2}c^2 q_1 q_2, \qquad (V.14)$$

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 $\varepsilon < 1$. 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 is 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 Figure 2. Here the stable pair is indicated by $\times$ while the
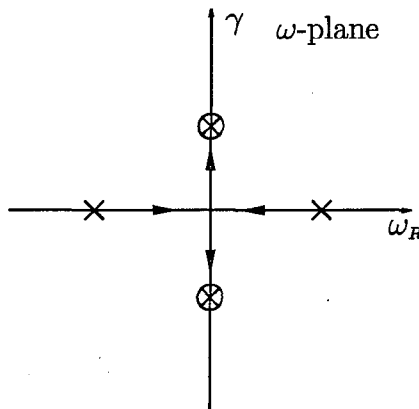

Figure 2:

unstable pair by the $\otimes$. This bifurcation generally occurs in systems where the Hamiltonian is separable, i.e. $H = p^2/2 + V(q)$, i.e. those for which Lagrange's theorem applies. It occurs in one degree-of-freedom systems where the potential goes from concave up to concave down. The arrows of the figure correspond to this case. For the system of (V.10) it occurs when $G = 0$ and $\omega_k^2 \to -\omega_k^2$.

The other bifurcation, which is something called a *Krein crash*, is illustrated in Figure 3. The arrows indicate the path followed by the eigenvalues of system (V.10) 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
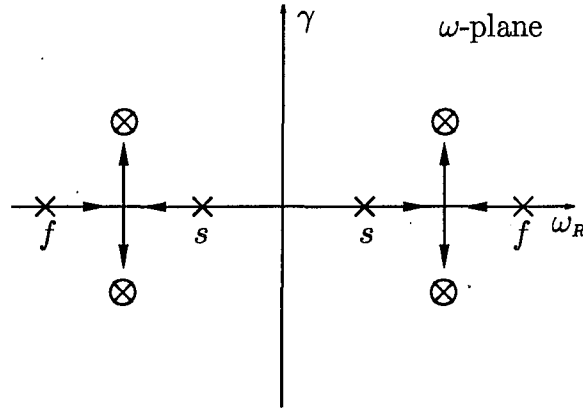
Figure 3:

real axis or they can migrate off the real axis as shown in the figure. *Krein's theorem*[*] states that a necessary condition for this latter case is that the signature of the colliding modes must be different, i.e. one of them must be an NEM. The proof of Krein's theorem is not difficult; it relies on the fact that definite Hamiltonians cannot have instabilities.

Krein's theorem provides a means for detecting the occurrence of NEM's. 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?

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 Lecture II. 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 = $ constant, 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 orbit, which we denote by

$$q_e = q_e(a,t), \qquad \pi_e = \pi_e(a,t). \qquad\qquad (V.15)$$

The functions above are particular in that they have the properties

$$\left.\frac{\rho_0(a)}{\mathcal{J}(a,t)}\right|_{a=q_e^{-1}(r,t)} = \rho_e(r) \qquad\qquad (V.16)$$

$$s_0(a)|_{a=q_e^{-1}(r,t)} = s_e(r) \qquad\qquad (V.17)$$

$$\left.\frac{\pi_e(a,t)}{\rho_0}\right|_{a=q_e^{-1}(r,t)} = \dot{q}_e(a,t)|_{a=q_e^{-1}(r,t)} = v_e(r)\,, \qquad\qquad (V.18)$$

---

[*]Moser (1958) and (1968), *Ref. V A.*

82

where we emphasize that, upon doing the substitutions indicated on the right-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 could linearize. We could set

$$q(a,t) = q_e(a,t) + \xi(a,t), \qquad \pi = \pi_e(a,t) + p(a,t) \qquad (\text{V.19})$$

and expand (II.88); however, the resulting equation would 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 by an old trick. To see this we turn to the action principle of (II.72), insert (V.19), and expand

$$S[q] = S[q_e] + \delta S[q_e; \xi] + \delta^2 S[q_e; \xi] + \dots. \qquad (\text{V.20})$$

The first term of (V.20) is merely a number, while the second term vanishes since the reference trajectory q_e is assumed to be a solution and is thus an extremal point. The third term, upon variation with respect to ξ, generates the linear dynamics relative to the reference state q_e. It is given by

$$\delta^2 S[q_e; \xi] = \int_{t_0}^{t_1} dt \int_D d^3a \left(\tfrac{1}{2}\rho_0 \dot{\xi}^2 - \left[\frac{\rho_0^2 U_\rho}{2\mathcal{J}}\right]_{q_e} \left[(\xi_{i,i})^2 + \xi_{j,i}\,\xi_{i,j}\right] - \left[\frac{\rho_0^3}{2\mathcal{J}^2}U_{\rho\rho}\right]_{q_e}(\xi_{i,i})^2 \right). \qquad (\text{V.21})$$

It is important to observe that in (V.21) the term involving U_ρ and $U_{\rho\rho}$ possesses the explicit time dependence arising from $q_e(a,t)$. The old trick 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

$$\eta(r,t) := \xi(a,t)|_{a=q_e^{-1}(r,t)}. \qquad (\text{V.22})$$

The quantity $\eta(r,t)$ is a sort of Eulerian field for the Lagrangian displacement variable. A time derivative of (V.22) yields

$$\dot{\xi}(a,t) = \frac{\partial \eta(r,t)}{\partial t} + \frac{\partial \eta(r,t)}{\partial r} \cdot \dot{q}_e|_{a=q_e^{-1}(r,t)} \qquad (\text{V.23})$$

or in light of (V.18)

$$\dot{\xi}(a,t) = \frac{\partial \eta(r,t)}{\partial t} + v_e(r) \cdot \nabla \eta(r,t). \qquad (\text{V.24})$$

Note that we have used "$\cdot$" for time derivatives at constant a and $\partial/\partial t$ for time derivatives at constant r. Since in (V.24) $v_e(r)$, the equilibrium velocity, is time independent, no explicit time dependence is introduced by this transformation.

It is interesting and revealing to compare (V.24) with the transformation for time derivatives when going into a rotating frame of reference

$$\left.\frac{\partial}{\partial t}\right|_{\text{fixed}} = \left.\frac{\partial}{\partial t}\right|_{\text{rot}} + \Omega \times . \qquad (\text{V.25})$$

Just as the second term of (V.25) gives rise to noninertial (or fictional) forces, notably the Coriolis force that gives rise to the gyroscopic term in the Hamiltonian of (V.10), the second term of (V.24) will give rise to a noninertial type force in the fluid Hamiltonian. Transforming (V.21), using (V.22) and (V.24) yields

$$\delta^2 S[\eta] = \tfrac{1}{2} \int_{t_0}^{t_1} dt \int_D d^3r \left(\rho_e \, |\dot\eta + v_e \cdot \nabla\eta|^2 - \eta \cdot \mathfrak{W}_e \cdot \eta \right) \tag{V.26}$$

where $\mathfrak{W}_e$ 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, the potential energy, can be written as

$$\delta^2 W := \tfrac{1}{2} \int_D d^3r \, \eta \cdot \mathfrak{W}_e \cdot \eta$$

$$= \tfrac{1}{2} \int_D d^3r \left((\nabla \cdot \eta)^2 \, \rho_e \frac{\partial p_e}{\partial \rho_e} + (\nabla \cdot \eta)(\eta \cdot \nabla p_e) \right) , \tag{V.27}$$

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) Hamiltonian by Legendre transformation. The canonical momentum is given by

$$p = \frac{\delta L}{\delta \dot\eta} = \rho_e(\dot\eta + v_e \cdot \nabla\eta) , \tag{V.28}$$

whence the Hamiltonian is seen to be

$$\delta^2 H[p, \eta] = \tfrac{1}{2} \int_D d^3r \left(\frac{p^2}{\rho_e} - 2p \cdot (v_e \cdot \nabla\eta) + \eta \cdot \mathfrak{W}_e \cdot \eta \right) , \tag{V.29}$$

which has the "noninertial" term $-p_i \, v_{ej} \, \partial\eta_i/\partial r_j$ that is reminiscent of the gyroscopic term of (V.10).

Now, it should come as no surprise that ideal fluids 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, is not possible since the Hamiltonian is not of the separable form.

References

[1] N. G. Van Kampen and B. U. Felderhof, *Theoretical Methods in Plasma Physics* (North Holland, Amsterdam, 1967).

[2] W. A. Newcomb, "Lagrangian and Hamiltonian Methods in Magnetohydrodynamics," Nuclear Fusion: Supplement, pt. 2, 451-463 (1962).

84

[3] J. K. Moser, Comm. Pure Appl. Math. **11**, 81 (1958); Mem. Am. Math. Soc. **81**, 1 (1968).

[4] P. J. Morrison and M. Kotschenreuther, "The Free Energy Principle, Negative Energy Modes, and Stability," in *Nonlinear World: IV International Workshop on Nonlinear and Turbulent Processes in Physics* eds. V.G. Bar'yakhtar, V. M. Chernousenko, N. S. Erokhin, A. B. Sitenko, and V. E. Zakharov (World Scientific, Singapore, 1990).

[5] D. Pfirsch and R. N. Sudan, Phys. Fluids B **5**, 2052 (1993).

B.   Stability and Noncanonical Hamiltonian Systems

In noncanonical Hamiltonian systems it is still the case that equilibria occur at extremal points of the Hamiltonian,

$$\dot{z}^i = J^{ij}\frac{\partial H}{\partial z^j} = [z^i, H] = 0\,, \tag{V.30}$$

but the situation is more complicated. To see that something is amiss, consider the variation of the energy for a barotropic fluid, where

$$H[\rho, v] = \int_D \left[\tfrac{1}{2}\rho v^2 + \rho U(\rho)\right] d^3r\,; \tag{V.31}$$

namely,

$$\frac{\delta H}{\delta v} = \rho v$$

$$\frac{\delta H}{\delta \rho} = \frac{v^2}{2} + U(\rho) + \rho U(\rho)\,. \tag{V.32}$$

Setting the right-hand side of (V.32) to zero results in the trivial equilibrium state with $v = 0$ and $\rho = $ constant (which is generally zero). 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?

To answer these questions, compare (V.30) with its counterpart for the canonical case:

$$\dot{z}^i = J_c^{ij}\frac{\partial H}{\partial z^j} = 0\,. \tag{V.33}$$

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. 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 corank of J. In this case the general solution to (V.33) is given by

$$\left.\frac{\partial F}{\partial z^i}\right|_{z_e} = \left.\frac{\partial H}{\partial z^i}\right|_{z_e} + \lambda_\alpha \left.\frac{\partial C^\alpha}{\partial z^i}\right|_{z_e} = 0\,. \tag{V.34}$$

Here λ_α are Lagrange multipliers, which are determined by choosing the values of the constants of motion C^α. Thus (V.34) gives those equilibria that lie on the symplectic leaf with the chosen values.

Not surprisingly, the linear dynamics obtained by setting $z = z_e + \delta z$ and expanding to first order, exhibits behavior arising from $\det J = 0$, namely, the existence of zero frequency modes. The equation for the linear dynamics is easily seen to be

$$\delta \dot{z} = A^i_k(z_e) \delta z^k \,, \tag{V.35}$$

where

$$A^i_k(z_e) := J^{ij}(z_e) \frac{\partial^2 F(z_e)}{\partial z^j \partial z^k} =: J^{ij}_e F_{,jk} \,. \tag{V.36}$$

[Note, this linear dynamics has a Hamiltonian structure with the Poisson bracket defined by J_e (which is constant) and the Hamiltonian given by $\delta^2 F := \frac{1}{2} F_{,jk} \, \delta z^j \delta z^k$.] Assuming $\delta z \sim e^{i\omega t}$ yields an eigenvalue problem with a characteristic equation given by

$$\det(i\omega I - A) = 0 \,, \tag{V.37}$$

where zero frequency modes satisfy

$$\det A = 0 \,. \tag{V.38}$$

In the canonical case, A is given by

$$A^i_{c\,k} = J^{ij}_c H_{,jk} \tag{V.39}$$

and

$$\det(A^i_{c\,k}) = \det(J^{ij}_c) \det(H_{,jk}) = \det(H_{,jk}) \,. \tag{V.40}$$

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

$$\mathrm{Rank}(A^i_k) \leq \min\{\mathrm{Rank}(J^{ij}), \mathrm{Rank}(F_{,jk})\} \,. \tag{V.41}$$

Thus for every Casimir there exists a null eigenvector, δz^k_0. To avoid complication suppose $\det(F_{,jk}) \neq 0$, i.e. that there are no local troughs in F, then all the null eigenvectors come from degeneracy in the bracket and they are given by

$$\delta z^k_0 = (F^{-1})^{kj} \frac{\partial C(z_e)}{\partial z^j} \,, \tag{V.42}$$

where $(F^{-1})^{kj} F_{,jl} = \delta^k_l$. Evidently, with δz_0 given by (V.42),

$$A^j_k \delta z^k_0 = J^{ij} F_{,jk} (F^{-1})^{kl} \frac{\partial C}{\partial z_l} = J^{ij} \frac{\partial C}{\partial z_j} = 0 \,. \tag{V.43}$$

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 can be used to define the
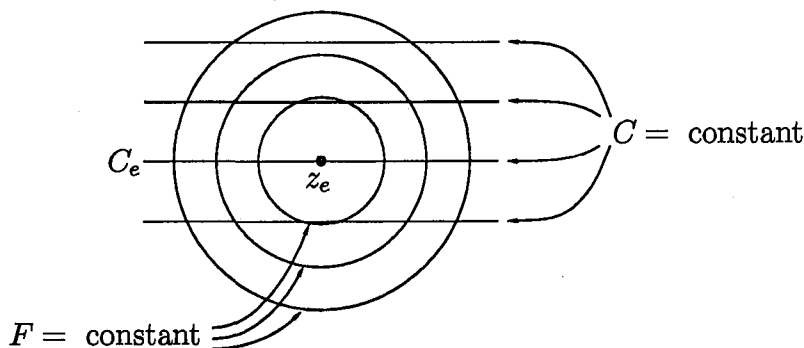
C_e

z_e

$C = $ constant

$F = $ constant

Figure 4:

subneighborhoods in the definition of stability given above, provided $\delta^2 F = \frac{1}{2}F_{,jk}\delta z^j \delta z^k = $ constant defines compact (as depicted in Figure 4) surfaces in the vicinity of z_e. 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 Figure 4). We will see in the next section that sometimes it is useful to take advantage of this information.

Although the picture described above for 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 (V.34) for all choices of λ_α. When the rank decreases on an open set, there is no problem in obtaining new Casimirs whose gradients span the null space of J. However, when the rank changes at (for example) a point, a new null eigenvector of J appears, an eigenvector that cannot be written as a gradient in the normal way.

The above pathology is perhaps best illustrated by an example. Consider the free rigid body of Lecture III, but modified so that the Hamiltonian has the form

$$H = \sum_{i=1}^{3}\left(\frac{\ell_i^2}{2I_i} + B_i\ell_i\right).$$
(V.44)

Here, we have added the linear term with B_i constant and nonzero for $i = 1, 2, 3$. This Hamiltonian is a sort of mixture between that of a spin system and a free rigid body. This form serves our purpose and we won't dwell on the physics, although it isn't hard to imagine a physical system where Hamiltonians of this form might arise. The equations of motion are now

$$\dot{\ell}_i = -\epsilon_{ijk}\ell_k\frac{\partial H}{\partial \ell_j} = -\epsilon_{ijk}\ell_k\left(\frac{\ell_j}{I_j} + B_j\right)$$
(V.45)

and it is clear that equilibria must satisfy

$$\ell_1(I_2^{-1}\ell_2 + B_2) - \ell_2(I_1^{-1}\ell_1 + B_1) = 0$$

$$\ell_1(I_3^{-1}\ell_3 + B_3) - \ell_3(I_1^{-1}\ell_1 + B_1) = 0$$

$$\ell_2(I_3^{-1}\ell_3 + B_3) - \ell_3(I_2^{-1}\ell_2 + B_2) = 0.\tag{V.46}$$

From (V.46) it is 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 (III.45). (Note the Casimirs remain the same as in Lecture III 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

$$\ell_1(I_1^{-1} + \lambda) = -B_1$$

$$\ell_2(I_2^{-1} + \lambda) = -B_2$$

$$\ell_3(I_3^{-1} + \lambda) = -B_3.\tag{V.47}$$

It is evident from (V.47) that there exists no choice of λ for which the equilibrium point

$$\ell_1 = \ell_2 = \ell_3 = 0\tag{V.48}$$

extremizes F. Observe, also, that the inequivalence of (V.46) and (V.47) occurs for an equilibrium, namely (V.48), that corresponds to a point where $J^{ij} = -\epsilon_{ijk}\ell_k$ changes from rank 2 to rank 0.

Another example* where $\delta F = 0$ does not yield all equilibria, is that of the 2-D Euler's equations for fluid motion (cf. Lecture III). Here the equation of motion yields the equilibrium relation

$$\frac{\partial \omega}{\partial t} = [\omega, \psi] = 0,\tag{V.49}$$

which is satisfied if ω and ψ are functionally dependent. Suppose $S = S(x, y)$ defines a locus of points, 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 the graph of ψ_e and vice versa. Thus we can write, e.g.

$$\nabla^2 \psi_e = G(\psi_e),\tag{V.50}$$

where $G(\psi_e)$ is an arbitrary function of ψ_e.

Let us contrast this with the equation obtained upon varying the functional $F = H + C$, which for the 2-D Euler equations, is given by

$$F[\omega] = -\frac{1}{2}\int_D \psi\,\omega\,d^2r + \int_D \mathcal{C}(\omega)\,d^2r.\tag{V.51}$$

The functional derivative $\delta F / \delta \omega = 0$ implies

$$\psi_e = \mathcal{C}'(\omega_e).\tag{V.52}$$

---

*This example is credited to V. Arnold, Izv. Vyssh. Uchebn. Zaved. Mat. **5** (54), 3 (1966) and (1966), *l.c.* Lecture III, which is the origin of the popular terminology "Arnold's method" or "Arnold's theorem" for the application of these ideas to other situations. This terminology is erroneous since the method was used in earlier papers: R. Fjortoft, Geofy. Pub. **17**, 1 (1950), W. Newcomb, in Appendix of I. Bernstein, Phys. Rev. **109**, 10 (1958), M. D. Kruskal and C. Oberman, Phys. Fluids **1**, 275 (1958), C. S. Gardner quoted in K. Fowler, J. Math. Phys. **4**, 559 (1963) and Phys. Fluids **6**, 839 (1963), and K. Fowler, *ibid.*

Assuming $C'(\omega)$ is monotonic we can solve for ω as follows:

$$\omega_e = \nabla^2 \psi_e = C'^{-1}(\psi_e).$$ (V.53)

Thus here, in contrast to (V.50), the vorticity must be a monotonic function of the stream function—if it is not, then it does not satisfy (V.52) and hence is not extremal. (Suppose $\psi_0 \neq \psi_1$ and $\omega(\psi_0) = \omega(\psi_1) = \omega_*$. Then (V.52) implies $\psi_0 = C'(\omega_*) = \psi_1$, which is a contradiction.)

In stability analyses it is advantageous for the equilibrium to be extremal. When this is the case, as for the monotonic equilibria above, one can calculate the second variation

$$\delta^2 F[\omega_e; \delta\omega] = \tfrac{1}{2} \int_D \left(|\nabla\psi|^2 + C''(\delta\omega)^2 \right) d^2r$$

$$= \tfrac{1}{2} \int_D \left(|\nabla\psi|^2 + (\delta\omega)^2 \left(\frac{\partial\omega_e(\psi_e)}{\partial\psi_e} \right)^{-1} \right) d^2r,$$ (V.54)

where the second equality follows upon differentiation of (V.52) 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 a "norm" defined by $\delta^2 F$. This would also be the case if $\partial\omega_e(\psi_e)/\partial\psi_e < 0$ and the second term of (V.54) could be shown to always dominate the first when $\delta\omega$ is in some space. 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 in Lecture II 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 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 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 that the functional is not differentiable at the equilibrium of interest but still can be proven to be stable by obtaining appropriate bounds.* Another technique is to restrict the class of variations so that they lie within symplectic leaves. In the next section we will see how this removes problems related to the rank changing behavior of J.

References

[1] D. D. Holm, J. E. Marsden, T. Ratiu, A. Weinstein, "Nonlinear Stability of Fluid and Plasma Equilibria," Physics Reports **123**, 1-116 (1985).

[2] P. J. Morrison and S. Eliezer, "Spontaneous Symmetry Breaking and Neutral Stability in the Noncanonical Hamiltonian Formalism," Physical Rev. A **33**, 4205-4214 (1986).

*G. Rein, preprint (1993).

[3]P. J. Morrison and M. Kotschenreuther, "The Free Energy Principle, Negative Energy Modes, and Stability," in *Nonlinear World: IV International Workshop on Nonlinear and Turbulent Processes in Physics* eds. V. G. Bar'yakhtar, V. M. Chernousenko, N. S. Erokhin, A. B. Sitenko, and V. E. Zakharov (World Scientific, Singapore, 1990).

[4]M. E. McIntyre and T. G. Shepherd, J. Fluid Mech. **181**, 527-565 (1987).

[5]T. G. Shepherd, "Extremal Properties and Hamiltonian Structure of the Euler Equations," in, *Topological Aspects of the Dynamics of Fluids and Plasmas*, eds. H. K. Moffatt *et al.* (Kluwer Academic, 1992), 275-283.

C. Dynamical Accessibility

Dynamically accessible perturbations are ones for which *all* the Casimir invariants are unchanged. As depicted in Figure 5, these perturbations lie in the surfaces defined by C^α = constant for all α. 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. The values of these invariants are determined by the Lagrange multipliers (and vice versa). In contrast, dynamically accessible perturbations are "direct" variations that automatically satisfy the constraints without choosing their particular values. The particular constraint surface is selected after the fact by the equilibrium point, rather than by Lagrange multipliers. Since the cosymplectic form, J^{ij}, projects (co)vectors onto the symplectic leaves, it is natural to consider a first order variation of the form

$$\delta z^i_{\mathrm{da}} = [\mathcal{G}, z^i] = J^{ji}(z)g_j \,, \tag{V.55}$$

where $\mathcal{G} := z^i g_i$. Here the arbitrariness in the variation is embodied in the arbitrariness in the generating function g_j, but because of the presence of J^{ij} the variation $\delta^{(1)} z_{\mathrm{da}}$ is arbitrary only within the symplectic leaf. Observe that J^{ij} is evaluated at any point z, in practice this will be a candidate equilibrium point that is determined after setting the first variation to zero.

Whether or not one wants to restrict to dynamically accessible variations, as described above, is a question of physics that likely must be determined on a case by case basis. In some systems the constraint is quite robust, while in others it is not. However, we will make the comment that if there exist mechanisms for creating perturbations that are not dynamically accessible, then it would seem appropriate to reexamine the model equation to see if such a mechanism should be incorporated into the dynamics.

Before considering equilibria and stability with this kind of variation, let us show explicitly that $\delta^{(1)} z_{\mathrm{da}}$ preserves the constraints to first order:

$$\delta C(z) = \frac{\partial C}{\partial z^i} \, \delta^{(1)} z^i = \frac{\partial C}{\partial z^i} \, J^{ji} g_j^{(1)} = 0 \,. \tag{V.56}$$

An expression that preserves the constraint to second order is given by

$$\delta^{(2)} z^i_{\mathrm{da}} = J^{ij} g_j^{(2)} + \tfrac{1}{2} J^{jl} \frac{\partial J^{ti}}{\partial z^l} \, g_t^{(1)} g_j^{(1)} \,. \tag{V.57}$$
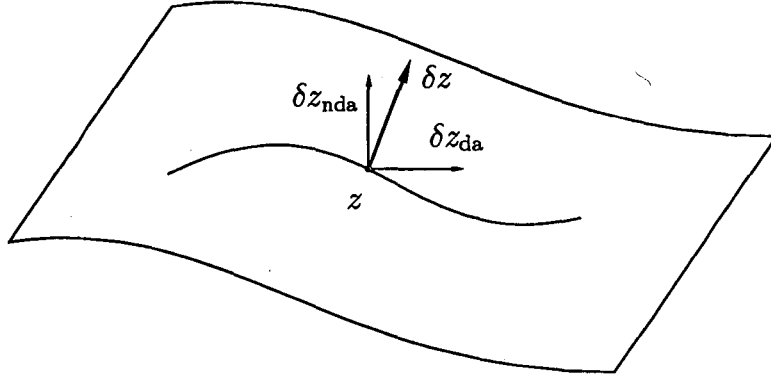
Figure 5:

Here we have added the superscripts (1) and (2) to distinguish the first order from the second order generating functions. Inserting (V.57) into $\delta^2 C$ and using the fact that $J^{ij}\partial C/\partial z^i = 0$ (in at least an open set) verifies the assertion.

In the case where $J^{ij} = c_k^{ij}z^k$, the first and second order variations have the form

$$\delta^{(1)}z_{\text{da}}^i = = c_k^{ji}z^k g_j^{(1)}$$

$$\delta^{(2)}z_{\text{da}}^i = c_k^{ji}z^k g_j^{(2)} + \tfrac{1}{2}c_l^{ti}c_k^{jl}z^k g_t^{(1)} g_j^{(1)}\,. \tag{V.58}$$

A convenient form to all orders is given by

$$\hat{z}^i = e^{g_j c_k^{ji}}z^k\,. \tag{V.59}$$

where $\Delta z := \hat{z} - z$ is a finite variation. The infinite dimensional analogue of (V.59) can be used to construct finite leaf variations, which are important for proving convexity in infinite dimensional systems. Expanding $g = g^{(1)} + g^{(2)} + \ldots$ and the exponential of (V.59), yields Eq. (V.58) to second order.

Return now to the example of the rigid body with the modified Hamiltonian. Using

$$\delta\ell_i^{\text{da}} = \epsilon_{ijk}\ell_k g_j\,, \tag{V.60}$$

we obtain

$$\delta F = \frac{\partial F}{\partial \ell_i}\,\delta\ell_i^{\text{da}} = \frac{\partial H}{\partial \ell_i}\delta\ell_i^{\text{da}} = (I_i^{-1}\ell_i + B_i)\epsilon_{ijk}\ell_k g_j^{(1)} = 0\,, \tag{V.61}$$

for the extremal equilibrium condition. Equation (V.61) yields a result that is identical to (V.46), the equilibrium condition obtained upon setting $\dot{\ell}_i = 0$ in the equation of motion.

In the case of the 2-D Euler fluid

$$\delta\omega_{\text{da}} = \{\mathcal{G},\omega\} = -[g,\omega]\,, \tag{V.62}$$

91

where $\mathcal{G} := \int_D \omega g \, d^2r$ with g arbitrary, and

$$\delta F_{\mathrm{da}} := \delta F[\omega; \delta\omega_{\mathrm{da}}] = -\int_D \psi \, \delta\omega_{\mathrm{da}} \, d^2r$$

$$= \int_D \psi[g, \omega] \, d^2r = -\int_D g[\psi, \omega] \, d^2r = 0 \,, \tag{V.63}$$

which implies $[\psi, \omega] = 0$—the condition obtained upon setting $\partial\omega/\partial t = 0$ in (V.49).

Proceeding now to the second variation, it is clear that stability can depend upon the class of variations allowed. Decomposing a general perturbation as

$$\delta z = \delta z_{\mathrm{da}} + \delta z_{\mathrm{nda}} \tag{V.64}$$

and inserting into $\delta^2 F$ yields

$$\delta^2 F = \delta^2 F_{\mathrm{da}} + \delta^2 F_{\mathrm{nda}} \,, \tag{V.65}$$

where

$$\delta^2 F_{\mathrm{da}} = \tfrac{1}{2} \left(\frac{\partial^2 H(z_e)}{\partial z^i \partial z^j} + \lambda_\alpha \frac{\partial^2 C^\alpha(z_e)}{\partial z^i \partial z^j} \right) J^{li}(z_e) g_l^{(1)} J^{kj}(z_e) g_k^{(1)} \,. \tag{V.66}$$

Note, it is always the case that $\delta^2 F_{\mathrm{nda}}$ depends only on the first order g's. It is evident that $\delta^2 F$ can be indefinite because of the presence of $\delta^2 F_{\mathrm{nda}}$, even if $\delta^2 F_{\mathrm{da}}$, which involves only perturbations of the form $J^{ij}(z_e) g_i^{(1)}$, is of definite sign. An example is given by the free rigid body with the equilibrium

$$\ell_1^e = -\frac{BI_1}{\lambda I_1 + 1}, \qquad \ell_2^e = \ell_3^e = 0 \,, \tag{V.67}$$

where we set $B_1 = B$ and $B_2 = B_3 = 0$. In this case

$$\delta^2 F = -\frac{B}{2\ell_1^e}(\delta\ell_1)^2 + \frac{1}{2}\left(\frac{1}{I_2} - \frac{1}{I_1}\right)(\delta\ell_2)^2 + \frac{1}{2}\left(\frac{1}{I_3} - \frac{1}{I_1}\right)(\delta\ell_3)^2 \,. \tag{V.68}$$

If $I_1 < I_2 < I_3$, the last two terms are positive; however, the first term can have either sign. Dynamically accessible perturbations satisfy

$$\delta\ell_i^{\mathrm{da}} = \epsilon_{ijk}\ell_j^e g_k = \epsilon_{ijk}\ell_j^e g_k \,; \tag{V.69}$$

hence $\delta\ell_1^{\mathrm{da}} = 0$. Therefore, $\delta^2 F_{\mathrm{da}}$ is definite, even though $\delta^2 F$ need not be. Observe that the nondynamically accessible perturbation corresponds to the null eigenvector described above.

In this example, and above, we substituted the first order dynamically accessible variation into the second order quantity $\delta^2 F$. To some of you it may not be clear that $\delta^2 F_{\mathrm{da}}$ is identical to $\delta^2 H_{\mathrm{da}}$, which is obtained by expanding H to second order and then inserting (V.55) and (V.57). It is, however, straightforward to show that these are in fact identical. Expanding some Casimir C^α to second order about the equilibrium yields

$$\Delta^{(2)} C^\alpha = \frac{\partial C^\alpha}{\partial z^i}\delta^{(2)}z^i + \tfrac{1}{2}\frac{\partial^2 C^\alpha}{\partial z^i \partial z^j}\delta^{(1)}z^i \delta^{(1)}z^j \,, \tag{V.70}$$

92

but, when restricted to the constraint surface, (V.70) reduces to

$$\Delta^{(2)}C^\alpha \mathrm{da} = \frac{\partial C^\alpha}{\partial z^i} J_e^{li} g_l^{(1)} + \frac{\partial C^\alpha}{\partial z^i} J_e^{li} g_l^{(2)}$$

$$+ \frac{1}{2}\frac{\partial C^\alpha}{\partial z^i}\frac{\partial J_e^{ti}}{\partial z_0^l} J_e^{jl} g_t^{(1)} g_j^{(1)} + \frac{1}{2}\frac{\partial^2 C^\alpha}{\partial z^i \partial z^j} J_e^{li} g_l^{(1)} J_e^{kj} g_k^{(1)} \,. \qquad (V.71)$$

The first and second terms in (V.71) clearly vanish because

$$J^{ij}\frac{\partial C}{\partial z^j} = 0 \,. \qquad (V.72)$$

However, by exploiting the local nature of the constraint surface, it is also possible to show that the last two terms cancel, so that, to second order, $\Delta^{(2)}C_{\mathrm{da}}^\alpha$ vanishes identically. Indeed, one can realize (V.72) 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

$$0 = J^{ij}\frac{\partial C}{\partial z^j} = J_e^{ij}\frac{\partial C}{\partial z_e^j} + \delta z^l \left(\frac{\partial J_e^{ij}}{\partial z_e^l}\frac{\partial C^\alpha}{\partial z_e^j} + J_e^{ij}\frac{\partial^2 C^\alpha}{\partial z_e^l \partial z_e^j}\right) + \dots \qquad (V.73)$$

must vanish identically. The first term in (V.73) is clearly zero, while the vanishing of the second term, the one linear in δz^l, yields the desired relation

$$\frac{\partial J_e^{ij}}{\partial z_e^l}\frac{\partial C^\alpha}{\partial z_e^j} = -J_e^{ij}\frac{\partial^2 C^\alpha}{\partial z_e^l \partial z_e^j} \,, \qquad (V.74)$$

between the first and second partial derivatives of C^α. It follows immediately that the second variation $\Delta^{(2)}C_{\mathrm{da}} = 0$.

Similarly, expanding H to second order yields

$$\Delta^{(2)}H = \frac{\partial H}{\partial z^i}\delta^{(1)}z^i + \frac{\partial H}{\partial z^i}\delta^{(2)}z^i + \frac{1}{2}\frac{\partial^2 H}{\partial z^i \partial z^j}\delta^{(1)}z^i \delta^{(1)}z^j \,, \qquad (V.75)$$

which, when restricted to lie within the constraint surface, takes the form

$$\Delta^{(2)}H_{\mathrm{da}} = \delta^{(2)}H_{\mathrm{da}} = \frac{1}{2}\frac{\partial^2 H}{\partial z^i \partial z^j} J_e^{li} g_l^{(1)} J_e^{kj} g_k^{(1)} + \frac{1}{2}\frac{\partial H}{\partial z^i}\frac{\partial J_e^{ti}}{\partial z_e^l} J_e^{jl} g_t^{(1)} g_j^{(1)} \,. \qquad (V.76)$$

It is evident that the first term of (V.76) is the same as the first term of the free energy $\delta^{(2)}F_{\mathrm{da}}$, but in order to compare the second terms in these relations, one must again use (V.74) and the equilibrium condition (V.34) involving the Lagrange multipliers. Indeed, by summing (V.74) over λ_α and then exploiting (V.34), one concludes that

$$\lambda_\alpha J_e^{ij}\frac{\partial^2 C^\alpha}{\partial z_e^l \partial z_e^j} = -\lambda_\alpha\frac{\partial J_e^{ij}}{\partial z_e^l}\frac{\partial C^\alpha}{\partial z_e^j} = \frac{\partial J_e^{ij}}{\partial z_e^l}\frac{\partial H}{\partial z_e^j} \,. \qquad (V.77)$$

It thus follows that, as was asserted, the constrained variation $\delta^{(2)}H_{\mathrm{da}} = \delta^{(2)}F_{\mathrm{da}}$.

Now we show how dynamically accessible variations are related to Lagrangian variations, in the context of the ideal fluid. In particular we will relate the Lagrangian and Eulerian potential energy functional.

To obtain dynamically accessible variations for the fluid, the following functional:

$$\mathcal{G} := \int_D (M \cdot \eta + h\rho + k\sigma) \ d^3r \,, \tag{V.78}$$

can be inserted into the bracket of (III.89). Here the arbitrariness of variation within the symplectic leaf is described by the free functions of r: η, h, and k. We will only need the expressions for the first and second variations of the density and entropy per unit mass

$$\delta^{(1)}\rho_{\text{da}} = \{\mathcal{G}, \rho\} = \nabla \cdot \left(\rho \frac{\delta \mathcal{G}}{\delta M}\right) = \nabla \cdot (\rho\eta)$$

$$\delta^{(1)}\sigma_{\text{da}} = \{\mathcal{G}, \sigma\} = \nabla \cdot \left(\sigma \frac{\delta \mathcal{G}}{\delta M}\right) = \nabla \cdot (\sigma\eta) \tag{V.79}$$

$$\delta^{(2)}\rho_{\text{da}} = \tfrac{1}{2}\{\mathcal{G}, \{\mathcal{G}, \rho\}\} = \tfrac{1}{2}\nabla \cdot [\eta\nabla \cdot (\rho\eta)]$$

$$\delta^{(2)}\sigma_{\text{da}} = \tfrac{1}{2}\{\{\mathcal{G}, \mathcal{G}, \sigma\}\} = \tfrac{1}{2}\nabla \cdot [\eta\nabla \cdot (\sigma\eta)] \,. \tag{V.80}$$

(Note we are not expanding $\mathcal{G}$ since we already know only the first order part contributes.) Observe that the variations of (V.79) are compatible with those of (IV.101), which are induced by variation of the Lagrangian coordinates.

The potential energy functional for the ideal fluid is

$$W[\rho, \sigma] = \int_D \rho\tilde{U}(\rho, \sigma) \ d^3r \,, \tag{V.81}$$

where recall $\tilde{U}$ is the internal energy per unit mass and $\sigma = \rho s$. In terms of the function $\tilde{U}(\rho, \sigma)$ the equation of state for the pressure is given by

$$\tilde{p}(\rho, \sigma) = \rho^2 \left(\frac{\partial \tilde{U}}{\partial \rho} + \frac{\sigma}{\rho}\frac{\partial \tilde{U}}{\partial \sigma}\right) \,. \tag{V.82}$$

Here we have used the tilde to indicate that the dependence is upon ρ and σ instead of ρ and s. Upon Taylor expansion, the second order potential energy functional is seen to be

$$\delta^2 W = \tfrac{1}{2}\int_D \left(\left(\delta^{(1)}\rho\right)^2 \left(\rho\tilde{U}_{\rho\rho} + 2\tilde{U}_\rho\right) + \left(\delta^{(1)}\sigma\right)^2 \left(\rho\tilde{U}_{\sigma\sigma}\right) + 2\left(\delta^{(1)}\sigma \, \delta^{(1)}\rho\right)\left(\rho\tilde{U}_{\rho\sigma} + \tilde{U}_\sigma\right)\right.$$

$$\left. + 2\left(\delta^{(2)}\rho\right)\left(\rho\tilde{U}_\rho + \tilde{U}\right) + 2\left(\delta^{(2)}\sigma\right)\left(\rho\tilde{U}_\sigma\right)\right) d^3r \,, \tag{V.83}$$

where subscripts denote partial differentiation. Inserting (V.79) and (V.80) into (V.83) creates a relatively complicated formula, one with terms that are similar but with no immediate

simplification. What we have is in reality a sort of integration by parts puzzle. We will not give all the details here of a calculation that gets us to the desired end, but only a few "landmarks." The first move is to integrate the second order variations by parts. Next, the terms are grouped as follows:

$$\delta^2 W_{\mathrm{da}} = \tfrac{1}{2} \int_D \bigg((\nabla \cdot \eta)^2 \left(\rho^3 \tilde{U}_{\rho\rho} + 2\rho^2 \tilde{U}_\rho + \sigma^2 \rho \tilde{U}_{\sigma\sigma} + 2\sigma\rho^2 \tilde{U}_{\sigma\rho} + 2\sigma\rho \tilde{U}_\sigma \right)$$

$$+ (\nabla \cdot \eta) \left(\eta \cdot \nabla \sigma \right) \left(\sigma\rho \tilde{U}_{\sigma\sigma} + \rho^2 \tilde{U}_{\rho\sigma} + \rho \tilde{U}_\rho \right)$$

$$+ (\nabla \cdot \eta) \left(\eta \cdot \nabla \rho \right) \left(\rho^2 \tilde{U}_{\rho\rho} + 2\rho \tilde{U}_\rho + \sigma\rho \tilde{U}_{\rho\sigma} + \sigma \tilde{U}_\sigma \right) \bigg) \, d^3 r \,, \qquad (V.84)$$

which upon making use of (V.82) can be put into the form

$$\delta^2 W_{\mathrm{da}} = \tfrac{1}{2} \int_D \left((\nabla \cdot \eta)^2 \left(\rho \tilde{p}_\rho + \sigma \tilde{p}_\sigma \right) + (\nabla \cdot \eta) \left(\eta \cdot \nabla \tilde{p} \right) \right) \, d^3 r \,. \qquad (V.85)$$

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 (V.85) yields, finally,

$$\delta^2 W_{\mathrm{da}} = \tfrac{1}{2} \int_D \left((\nabla \cdot \eta)^2 \left(\rho p_\rho \right) + (\nabla \cdot \eta) \left(\eta \cdot \nabla \tilde{p} \right) \right) \, d^3 r \,. \qquad (V.86)$$

This expression, when evaluated on $\rho = \rho_e$ and $\sigma = \sigma_e$, is precisely that of (V.27), which was obtained in the strictly Lagrangian variable context. We have thus, in a sense, gone full circle!

References

[1] V. I. Arnold, "The Hamiltonian Nature of the Euler Equations in the Dynamics of a Rigid Body and of an Ideal Fluid," Usp. Mat. Nauk. **24**, 225 (1969).

[2] H. E. Kandrup and P. J. Morrison, "Hamiltonian Structure of the Vlasov-Einstein System and the Problem of Stability for Spherical Relativistic Star Clusters," Annals of Phys. **225**, 114-166 (1993).

[3] P. J. Morrison and M. Kotschenreuther, "The Free Energy Principle, Negative Energy Modes, and Stability," in *Nonlinear World: IV International Workshop on Nonlinear and Turbulent Processes in Physics* eds. V. G. Bar'yakhtar, V. M. Chernousenko, N. S. Erokhin, A. B. Sitenko, and V. E. Zakharov (World Scientific, Singapore, 1990).