1
Introduction and Overview

1.1 Lagrangian and Hamiltonian Formalisms

Mechanics deals with the dynamics of particles, rigid bodies, continuous media (fluid, plasma, and elastic materials), and field theories such as electromagnetism and gravity. This theory plays a crucial role in quantum mechanics, control theory, and other areas of physics, engineering, and even chemistry and biology. Clearly, mechanics is a large subject that plays a fundamental role in science. Mechanics also played a key part in the development of mathematics. Starting with the creation of calculus stimulated by Newton's mechanics, it continues today with exciting developments in group representations, geometry, and topology; these mathematical developments in turn are being applied to interesting problems in physics and engineering.

Symmetry plays an important role in mechanics, from fundamental formulations of basic principles to concrete applications, such as stability criteria for rotating structures. The theme of this book is to emphasize the role of symmetry in various aspects of mechanics.

This introduction treats a collection of topics fairly rapidly. The student should not expect to understand everything perfectly at this stage. *We will return to many of the topics in subsequent chapters.*

Lagrangian and Hamiltonian Mechanics. Mechanics has two main points of view, ***Lagrangian mechanics*** and ***Hamiltonian mechanics***. In one sense, Lagrangian mechanics is more fundamental, since it is based on variational principles and it is what generalizes most directly to the gen-

eral relativistic context. In another sense, Hamiltonian mechanics is more fundamental, since it is based directly on the energy concept and it is what is more closely tied to quantum mechanics. Fortunately, in many cases these branches are equivalent, as we shall see in detail in Chapter 7. Needless to say, the merger of quantum mechanics and general relativity remains one of the main outstanding problems of mechanics. In fact, the methods of mechanics and symmetry are important ingredients in the developments of string theory, which has attempted this merger.

Lagrangian Mechanics. The Lagrangian formulation of mechanics is based on the observation that there are variational principles behind the fundamental laws of force balance as given by Newton's law $\mathbf{F} = m\mathbf{a}$. One chooses a configuration space Q with coordinates q^i, $i = 1, \ldots, n$, that describe the ***configuration*** of the system under study. Then one introduces the ***Lagrangian*** $L(q^i, \dot{q}^i, t)$, which is shorthand notation for $L(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n, t)$. Usually, L is the kinetic *minus* the potential energy of the system, and one takes $\dot{q}^i = dq^i/dt$ to be the system velocity. The ***variational principle of Hamilton*** states

$$\delta \int_a^b L(q^i, \dot{q}^i, t)\, dt = 0. \tag{1.1.1}$$

In this principle, we choose curves $q^i(t)$ joining two fixed points in Q over a fixed time interval $[a, b]$ and calculate the integral regarded as a function of this curve. Hamilton's principle states that this function has a critical point at a solution within the space of curves. If we let δq^i be a variation, that is, the derivative of a family of curves with respect to a parameter, then by the chain rule, (1.1.1) is equivalent to

$$\sum_{i=1}^{n} \int_a^b \left(\frac{\partial L}{\partial q^i} \delta q^i + \frac{\partial L}{\partial \dot{q}^i} \delta \dot{q}^i \right) dt = 0 \tag{1.1.2}$$

for all variations δq^i.

Using equality of mixed partials, one finds that

$$\delta \dot{q}^i = \frac{d}{dt} \delta q^i.$$

Using this, integrating the second term of (1.1.2) by parts, and employing the boundary conditions $\delta q^i = 0$ at $t = a$ and b, (1.1.2) becomes

$$\sum_{i=1}^{n} \int_a^b \left[\frac{\partial L}{\partial q^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) \right] \delta q^i \, dt = 0. \tag{1.1.3}$$

Since δq^i is arbitrary (apart from being zero at the endpoints), (1.1.2) is equivalent to the ***Euler–Lagrange equations***

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0, \quad i = 1, \ldots, n. \tag{1.1.4}$$

As Hamilton [1834] realized, one can gain valuable information by *not* imposing the fixed endpoint conditions. We will have a deeper look at such issues in Chapters 7 and 8.

For a system of N particles moving in Euclidean 3-space, we choose the configuration space to be $Q = \mathbb{R}^{3N} = \mathbb{R}^3 \times \cdots \times \mathbb{R}^3$ (N times), and L often has the form of kinetic minus potential energy:

$$L(\mathbf{q}_i, \dot{\mathbf{q}}_i, t) = \frac{1}{2}\sum_{i=1}^{N} m_i \|\dot{\mathbf{q}}_i\|^2 - V(\mathbf{q}_i), \tag{1.1.5}$$

where we write points in Q as $\mathbf{q}_1, \ldots, \mathbf{q}_N$, where $\mathbf{q}_i \in \mathbb{R}^3$. In this case the Euler–Lagrange equations (1.1.4) reduce to ***Newton's second law***

$$\frac{d}{dt}(m_i \dot{\mathbf{q}}_i) = -\frac{\partial V}{\partial \mathbf{q}_i}, \quad i = 1, \ldots, N, \tag{1.1.6}$$

that is, $\mathbf{F} = m\mathbf{a}$ for the motion of particles in the potential V. As we shall see later, in many examples more general Lagrangians are needed.

Generally, in Lagrangian mechanics, one identifies a configuration space Q (with coordinates $(q^1, \ldots, q^n)$) and then forms the ***velocity phase space*** TQ, also called the ***tangent bundle*** of Q. Coordinates on TQ are denoted by

$$(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n),$$

and the Lagrangian is regarded as a function $L : TQ \to \mathbb{R}$.

Already at this stage, interesting links with geometry are possible. If $g_{ij}(q)$ is a given metric tensor or ***mass matrix*** (for now, just think of this as a q-dependent positive definite symmetric $n \times n$ matrix) and we consider the kinetic energy Lagrangian

$$L(q^i, \dot{q}^i) = \frac{1}{2}\sum_{i,j=1}^{n} g_{ij}(q)\dot{q}^i\dot{q}^j, \tag{1.1.7}$$

then *the Euler–Lagrange equations are equivalent to the equations of geodesic motion*, as can be directly verified (see §7.5 for details). Conservation laws that are a result of symmetry in a mechanical context can then be applied to yield interesting geometric facts. For instance, theorems about geodesics on surfaces of revolution can be readily proved this way.

The Lagrangian formalism can be extended to the infinite-dimensional case. One view (but not the only one) is to replace the q^i by *fields* $\varphi^1, \ldots, \varphi^m$ that are, for example, functions of spatial points x^i and time. Then L is a function of $\varphi^1, \ldots, \varphi^m, \dot{\varphi}^1, \ldots, \dot{\varphi}^m$ and the spatial derivatives of the fields. We shall deal with various examples of this later, but we emphasize that properly interpreted, the variational principle and the Euler–Lagrange equations remain intact. One replaces the partial derivatives in the Euler–Lagrange equations by *functional derivatives* defined below.

Hamiltonian Mechanics. To pass to the Hamiltonian formalism, introduce the ***conjugate momenta***

$$p_i = \frac{\partial L}{\partial \dot{q}^i}, \qquad i = 1, \dots, n, \tag{1.1.8}$$

make the change of variables $(q^i, \dot{q}^i) \mapsto (q^i, p_i)$, and introduce the ***Hamiltonian***

$$H(q^i, p_i, t) = \sum_{j=1}^{n} p_j \dot{q}^j - L(q^i, \dot{q}^i, t). \tag{1.1.9}$$

Remembering the change of variables, we make the following computations using the chain rule:

$$\frac{\partial H}{\partial p_i} = \dot{q}^i + \sum_{j=1}^{n} \left(p_j \frac{\partial \dot{q}^j}{\partial p_i} - \frac{\partial L}{\partial \dot{q}^j} \frac{\partial \dot{q}^j}{\partial p_i} \right) = \dot{q}^i \tag{1.1.10}$$

and

$$\frac{\partial H}{\partial q^i} = \sum_{j=1}^{n} p_j \frac{\partial \dot{q}^j}{\partial q^i} - \frac{\partial L}{\partial q^i} - \sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}^j} \frac{\partial \dot{q}^j}{\partial q^i} = -\frac{\partial L}{\partial q^i}, \tag{1.1.11}$$

where (1.1.8) has been used twice. Using (1.1.4) and (1.1.8), we see that (1.1.11) is equivalent to

$$\frac{\partial H}{\partial q^i} = -\frac{d}{dt} p_i. \tag{1.1.12}$$

Thus, *the Euler–Lagrange equations are equivalent to* ***Hamilton's equations***

$$\begin{aligned} \frac{dq^i}{dt} &= \frac{\partial H}{\partial p_i}, \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial q^i}, \end{aligned} \tag{1.1.13}$$

where $i = 1, \dots, n$. The analogous Hamiltonian partial differential equations for time-dependent *fields* $\varphi^1, \dots, \varphi^m$ and their conjugate momenta $\pi_1, \dots, \pi_m$ are

$$\begin{aligned} \frac{\partial \varphi^a}{\partial t} &= \frac{\delta H}{\delta \pi_a}, \\ \frac{\partial \pi_a}{\partial t} &= -\frac{\delta H}{\delta \varphi^a}, \end{aligned} \tag{1.1.14}$$

where $a = 1, \ldots, m$, H is a functional of the fields φ^a and π_a, and the ***variational***, or ***functional***, ***derivatives*** are defined by the equation

$$\int_{\mathbb{R}^n} \frac{\delta H}{\delta \varphi^1} \delta\varphi^1 \, d^n x = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [H(\varphi^1 + \varepsilon\delta\varphi^1, \varphi^2, \ldots, \varphi^m, \pi_1, \ldots, \pi_m) - H(\varphi^1, \varphi^2, \ldots, \varphi^m, \pi_1, \ldots, \pi_m)], \quad (1.1.15)$$

and similarly for $\delta H/\delta\varphi^2, \ldots, \delta H/\delta\pi_m$. Equations (1.1.13) and (1.1.14) can be recast in ***Poisson bracket form***:

$$\dot{F} = \{F, H\}, \quad (1.1.16)$$

where the brackets in the respective cases are given by

$$\{F, G\} = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q^i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q^i} \right) \quad (1.1.17)$$

and

$$\{F, G\} = \sum_{a=1}^{m} \int_{\mathbb{R}^n} \left(\frac{\delta F}{\delta \varphi^a} \frac{\delta G}{\delta \pi_a} - \frac{\delta F}{\delta \pi_a} \frac{\delta G}{\delta \varphi^a} \right) d^n x. \quad (1.1.18)$$

Associated to any configuration space Q (coordinatized by $(q^1, \ldots, q^n)$) is a phase space T^*Q called the ***cotangent bundle*** of Q, which has coordinates $(q^1, \ldots, q^n, p_1, \ldots, p_n)$. On this space, the canonical bracket (1.1.17) is intrinsically defined in the sense that the value of $\{F, G\}$ is independent of the choice of coordinates. Because the Poisson bracket satisfies $\{F, G\} = -\{G, F\}$ and in particular $\{H, H\} = 0$, we see from (1.1.16) that $\dot{H} = 0$; that is, *energy is conserved*. This is the most elementary of many deep and beautiful *conservation properties* of mechanical systems.

There is also a variational principle on the Hamiltonian side. For the Euler–Lagrange equations, we deal with curves in q-space (configuration space), whereas for Hamilton's equations we deal with curves in (q, p)-space (momentum phase space). The principle is

$$\delta \int_a^b \left[\sum_{i=1}^{n} p_i \dot{q}^i - H(q^j, p_j) \right] dt = 0, \quad (1.1.19)$$

as is readily verified; one requires $p_i \delta q^i = 0$ at the endpoints.

This formalism is the basis for the analysis of many important systems in particle dynamics and field theory, as described in standard texts such as Whittaker [1927], Goldstein [1980], Arnold [1989], Thirring [1978], and Abraham and Marsden [1978]. The underlying geometric structures that are important for this formalism are those of *symplectic* and *Poisson geometry*. How these structures are related to the Euler–Lagrange equations and variational principles via the Legendre transformation is an essential ingredient

of the story. Furthermore, in the infinite-dimensional case it is fairly well understood how to deal rigorously with many of the functional analytic difficulties that arise; see, for example, Chernoff and Marsden [1974] and Marsden and Hughes [1983].

Exercises

◇ **1.1-1.** Show by *direct calculation* that the classical Poisson bracket satisfies the ***Jacobi identity***. That is, if F and K are both functions of the $2n$ variables $(q^1, q^2, \ldots, q^n, p_1, p_2, \ldots, p_n)$ and we define

$$\{F, K\} = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q^i} \frac{\partial K}{\partial p_i} - \frac{\partial K}{\partial q^i} \frac{\partial F}{\partial p_i} \right),$$

then the identity $\{L, \{F, K\}\} + \{K, \{L, F\}\} + \{F, \{K, L\}\} = 0$ holds.

1.2 The Rigid Body

It was already clear in the 19th century that certain mechanical systems resist the canonical formalism outlined in §1.1. For example, to obtain a Hamiltonian description for fluids, Clebsch [1857, 1859] found it necessary to introduce certain nonphysical potentials.[1] We will discuss fluids in §1.4 below.

Euler's Rigid-Body Equations. In the absence of external forces, the Euler equations for the rotational dynamics of a rigid body about its center of mass are usually written as follows, as we shall derive in detail in Chapter 15:

$$\begin{aligned} I_1 \dot{\Omega}_1 &= (I_2 - I_3)\Omega_2 \Omega_3, \\ I_2 \dot{\Omega}_2 &= (I_3 - I_1)\Omega_3 \Omega_1, \\ I_3 \dot{\Omega}_3 &= (I_1 - I_2)\Omega_1 \Omega_2, \end{aligned} \tag{1.2.1}$$

where $\mathbf{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$ is the body angular velocity vector (the angular velocity of the rigid body as seen from a frame fixed in the body) and I_1, I_2, I_3 are constants depending on the shape and mass distribution of the body—the principal moments of inertia of the rigid body.

Are equations (1.2.1) Lagrangian or Hamiltonian in any sense? Since there is an *odd* number of equations, they obviously cannot be put in canonical Hamiltonian form in the sense of equations (1.1.13).

---

[1] For a geometric account of Clebsch potentials and further references, see Marsden and Weinstein [1983], Marsden, Ratiu, and Weinstein [1984a, 1984b], Cendra and Marsden [1987], and Cendra, Ibort, and Marsden [1987].

A classical way to see the Lagrangian (or Hamiltonian) structure of the rigid-body equations is to use a description of the orientation of the body in terms of three Euler angles denoted by θ, φ, ψ and their velocities $\dot{\theta}, \dot{\varphi}, \dot{\psi}$ (or conjugate momenta $p_\theta, p_\varphi, p_\psi$), relative to which the equations *are* in Euler–Lagrange (or canonical Hamiltonian) form. However, this procedure requires using *six equations*, while many questions are easier to study using the *three equations* (1.2.1).

Lagrangian Form. To see the sense in which (1.2.1) are Lagrangian, introduce the Lagrangian

$$L(\mathbf{\Omega}) = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2), \tag{1.2.2}$$

which, as we will see in detail in Chapter 15, is the (rotational) kinetic energy of the rigid body. We then write (1.2.1) as

$$\frac{d}{dt}\frac{\partial L}{\partial \mathbf{\Omega}} = \frac{\partial L}{\partial \mathbf{\Omega}} \times \mathbf{\Omega}. \tag{1.2.3}$$

These equations appear explicitly in Lagrange [1788, Volume 2, p. 212] and were generalized to arbitrary Lie algebras by Poincaré [1901b]. We will discuss these general ***Euler–Poincaré equations*** in Chapter 13. We can also write a variational principle for (1.2.3) that is analogous to that for the Euler–Lagrange equations but is written *directly* in terms of $\mathbf{\Omega}$. Namely, (1.2.3) is equivalent to

$$\delta \int_a^b L\, dt = 0, \tag{1.2.4}$$

where variations of $\mathbf{\Omega}$ are restricted to be of the form

$$\delta\mathbf{\Omega} = \dot{\mathbf{\Sigma}} + \mathbf{\Omega} \times \mathbf{\Sigma}, \tag{1.2.5}$$

where $\mathbf{\Sigma}$ is a curve in $\mathbb{R}^3$ that vanishes at the endpoints. This may be proved in the same way as we proved that the variational principle (1.1.1) is equivalent to the Euler–Lagrange equations (1.1.4); see Exercise 1.2-2. In fact, later on, in Chapter 13, we shall see how to *derive* this variational principle from the more "primitive" one (1.1.1).

Hamiltonian Form. If instead of variational principles we concentrate on Poisson brackets and drop the requirement that they be in the canonical form (1.1.17), then there is also a simple and beautiful Hamiltonian structure for the rigid-body equations. To state it, introduce the ***angular momenta***

$$\Pi_i = I_i\Omega_i = \frac{\partial L}{\partial \Omega_i}, \quad i = 1, 2, 3, \tag{1.2.6}$$

so that the Euler equations become

$$\begin{aligned}\dot{\Pi}_1 &= \frac{I_2 - I_3}{I_2 I_3}\Pi_2\Pi_3,\\ \dot{\Pi}_2 &= \frac{I_3 - I_1}{I_3 I_1}\Pi_3\Pi_1,\\ \dot{\Pi}_3 &= \frac{I_1 - I_2}{I_1 I_2}\Pi_1\Pi_2,\end{aligned} \tag{1.2.7}$$

that is,

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbf{\Omega}. \tag{1.2.8}$$

Introduce the ***rigid-body Poisson bracket*** on functions of the $\mathbf{\Pi}$'s,

$$\{F, G\}(\mathbf{\Pi}) = -\mathbf{\Pi} \cdot (\nabla F \times \nabla G), \tag{1.2.9}$$

and the Hamiltonian

$$H = \frac{1}{2}\left(\frac{\Pi_1^2}{I_1} + \frac{\Pi_2^2}{I_2} + \frac{\Pi_3^2}{I_3}\right). \tag{1.2.10}$$

One checks (Exercise 1.2-3) that Euler's equations (1.2.7) are equivalent to[2]

$$\dot{F} = \{F, H\}. \tag{1.2.11}$$

For *any* equation of the form (1.2.11), conservation of total angular momentum holds regardless of the Hamiltonian; indeed, with

$$C(\mathbf{\Pi}) = \frac{1}{2}(\Pi_1^2 + \Pi_2^2 + \Pi_3^2),$$

we have $\nabla C(\mathbf{\Pi}) = \mathbf{\Pi}$, and so

$$\frac{d}{dt}\frac{1}{2}(\Pi_1^2 + \Pi_2^2 + \Pi_3^2) = \{C, H\}(\mathbf{\Pi}) \tag{1.2.12}$$

$$= -\mathbf{\Pi} \cdot (\nabla C \times \nabla H) \tag{1.2.13}$$

$$= -\mathbf{\Pi} \cdot (\mathbf{\Pi} \times \nabla H) = 0. \tag{1.2.14}$$

The same calculation shows that $\{C, F\} = 0$ for any F. Functions such as these that ***Poisson commute*** with *every* function are called ***Casimir functions***; they play an important role in the study of *stability*, as we shall see later.[3]

---

[2]This Hamiltonian formulation of rigid body mechanics is implicit in many works, such as Arnold [1966a, 1969], and is given explicitly in this Poisson bracket form in Sudarshan and Mukunda [1974]. (Some preliminary versions were given by Pauli [1953], Martin [1959], and Nambu [1973].) On the other hand, the variational form (1.2.4) appears implicitly in Poincaré [1901b] and Hamel [1904]. It is given explicitly for fluids in Newcomb [1962] and Bretherton [1970] and in the general case in Marsden and Scheurle [1993a, 1993b].

[3]H. B. G. Casimir was a student of P. Ehrenfest and wrote a brilliant thesis on the quantum mechanics of the rigid body, a problem that has not been adequately

Exercises

◇ **1.2-1.** Show by direct calculation that the rigid-body Poisson bracket satisfies the Jacobi identity. That is, if F and K are both functions of (Π_1, Π_2, Π_3) and we define

$$\{F, K\}(\mathbf{\Pi}) = -\mathbf{\Pi} \cdot (\nabla F \times \nabla K),$$

then the identity $\{L, \{F, K\}\} + \{K, \{L, F\}\} + \{F, \{K, L\}\} = 0$ holds.

◇ **1.2-2.** Verify directly that the Euler equations for a rigid body are equivalent to

$$\delta \int L \, dt = 0$$

for variations of the form $\delta\mathbf{\Omega} = \dot{\mathbf{\Sigma}} + \mathbf{\Omega} \times \mathbf{\Sigma}$, where $\mathbf{\Sigma}$ vanishes at the endpoints.

◇ **1.2-3.** Verify directly that the Euler equations for a rigid body are equivalent to the equations

$$\frac{d}{dt}F = \{F, H\},$$

where $\{\,,\}$ is the rigid-body Poisson bracket and H is the rigid-body Hamiltonian.

◇ **1.2-4.**

(a) Show that the rotation group SO(3) can be identified with the ***Poincaré sphere***, that is, the ***unit circle bundle*** of the two-sphere S^2, defined to be the set of unit tangent vectors to the two-sphere in $\mathbb{R}^3$.

(b) Using the known fact from basic topology that any (continuous) vector field on S^2 must vanish somewhere, show that SO(3) *cannot* be written as $S^2 \times S^1$.

1.3 Lie–Poisson Brackets, Poisson Manifolds, Momentum Maps

The rigid-body variational principle and the rigid-body Poisson bracket are special cases of general constructions associated to any ***Lie algebra***

---

addressed in the detail that would be desirable, even today. Ehrenfest in turn wrote his thesis under Boltzmann around 1900 on variational principles in fluid dynamics and was one of the first to study fluids from this point of view in material, rather than Clebsch, representation. Curiously, Ehrenfest used the Gauss–Hertz principle of least curvature rather than the more elementary Hamilton principle. This is a seed for many important ideas in this book.

$\mathfrak{g}$, that is, a vector space together with a bilinear, antisymmetric bracket $[\xi, \eta]$ satisfying ***Jacobi's identity***:

$$[[\xi, \eta], \zeta] + [[\zeta, \xi], \eta] + [[\eta, \zeta], \xi] = 0 \tag{1.3.1}$$

for all $\xi, \eta, \zeta \in \mathfrak{g}$. For example, the Lie algebra associated to the rotation group is $\mathfrak{g} = \mathbb{R}^3$ with bracket $[\xi, \eta] = \xi \times \eta$, the ordinary vector cross product.

The Euler–Poincaré Equations. The construction of a variational principle on $\mathfrak{g}$ replaces

$$\delta\boldsymbol{\Omega} = \dot{\boldsymbol{\Sigma}} + \boldsymbol{\Omega} \times \boldsymbol{\Sigma} \quad \text{by} \quad \delta\xi = \dot{\eta} + [\eta, \xi].$$

The resulting general equations on $\mathfrak{g}$, which we will study in detail in Chapter 13, are called the ***Euler–Poincaré equations***. These equations are valid for either finite- or infinite-dimensional Lie algebras. To state them in the finite-dimensional case, we use the following notation. Choosing a basis $e_1, \ldots, e_r$ of $\mathfrak{g}$ (so $\dim \mathfrak{g} = r$), the ***structure constants*** C_{ab}^d are defined by the equation

$$[e_a, e_b] = \sum_{d=1}^{r} C_{ab}^d e_d, \tag{1.3.2}$$

where a, b run from 1 to r. If ξ is an element of the Lie algebra, its components relative to this basis are denoted by ξ^a so that $\xi = \sum_{a=1}^r \xi^a e_a$. If $e^1, \ldots, e^r$ is the corresponding dual basis, then the components of the differential of the Lagrangian L are the partial derivatives $\partial L/\partial \xi^a$. Then the Euler–Poincaré equations are

$$\frac{d}{dt}\frac{\partial L}{\partial \xi^d} = \sum_{a,b=1}^{r} C_{ad}^b \frac{\partial L}{\partial \xi^b} \xi^a. \tag{1.3.3}$$

The coordinate-free version reads

$$\frac{d}{dt}\frac{\partial L}{\partial \xi} = \mathrm{ad}_\xi^* \frac{\partial L}{\partial \xi},$$

where $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$ is the linear map $\eta \mapsto [\xi, \eta]$, and $\mathrm{ad}_\xi^* : \mathfrak{g}^* \to \mathfrak{g}^*$ is its dual. For example, for $L : \mathbb{R}^3 \to \mathbb{R}$, the Euler–Poincaré equations become

$$\frac{d}{dt}\frac{\partial L}{\partial \boldsymbol{\Omega}} = \frac{\partial L}{\partial \boldsymbol{\Omega}} \times \boldsymbol{\Omega},$$

which generalize the Euler equations for rigid-body motion. As we mentioned earlier, these equations were written down for a fairly general class

of L by Lagrange [1788, Volume 2, equation A, p. 212], while it was Poincaré [1901b] who generalized them to any Lie algebra.

The generalization of the rigid-body variational principle states that the Euler–Poincaré equations are equivalent to

$$\delta \int L\, dt = 0 \tag{1.3.4}$$

for all variations of the form $\delta\xi = \dot{\eta} + [\xi, \eta]$ for some curve η in $\mathfrak{g}$ that vanishes at the endpoints.

The Lie–Poisson Equations. We can also generalize the rigid-body Poisson bracket as follows: Let F, G be defined on the dual space $\mathfrak{g}^*$. Denoting elements of $\mathfrak{g}^*$ by μ, let the ***functional derivative*** of F at μ be the unique element $\delta F/\delta\mu$ of $\mathfrak{g}$ defined by

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}[F(\mu + \varepsilon\delta\mu) - F(\mu)] = \left\langle \delta\mu, \frac{\delta F}{\delta \mu} \right\rangle, \tag{1.3.5}$$

for all $\delta\mu \in \mathfrak{g}^*$, where $\langle\, , \rangle$ denotes the pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$. This definition (1.3.5) is consistent with the definition of $\delta F/\delta\varphi$ given in (1.1.15) when $\mathfrak{g}$ and $\mathfrak{g}^*$ are chosen to be appropriate spaces of fields. Define the ($\pm$) ***Lie–Poisson brackets*** by

$$\{F, G\}_\pm(\mu) = \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta G}{\delta \mu}\right] \right\rangle. \tag{1.3.6}$$

Using the coordinate notation introduced above, the ($\pm$) Lie–Poisson brackets become

$$\{F, G\}_\pm(\mu) = \pm \sum_{a,b,d=1}^{r} C^d_{ab}\mu_d \frac{\partial F}{\partial \mu_a}\frac{\partial G}{\partial \mu_b}, \tag{1.3.7}$$

where $\mu = \mu_a e^a$.

Poisson Manifolds. The Lie–Poisson bracket and the canonical brackets from the last section have four simple but crucial properties:

PB1	$\{F, G\}$ is real bilinear in F and G.	
PB2	$\{F, G\} = -\{G, F\}$,	antisymmetry.
PB3	$\{\{F, G\}, H\} + \{\{H, F\}, G\} + \{\{G, H\}, F\} = 0$,	Jacobi identity.
PB4	$\{FG, H\} = F\{G, H\} + \{F, H\}G$,	Leibniz identity.

A manifold (that is, an n–dimensional "smooth surface") P together with a bracket operation on $\mathcal{F}(P)$, the space of smooth functions on P, and satisfying properties **PB1**–**PB4**, is called a ***Poisson manifold***. In

particular, $\mathfrak{g}^*$ *is a Poisson manifold*. In Chapter 10 we will study the general concept of a Poisson manifold.

For example, if we choose $\mathfrak{g} = \mathbb{R}^3$ with the bracket taken to be the cross product $[\mathbf{x}, \mathbf{y}] = \mathbf{x} \times \mathbf{y}$, and identify $\mathfrak{g}^*$ with $\mathfrak{g}$ using the dot product on $\mathbb{R}^3$ (so $\langle \mathbf{\Pi}, \mathbf{x} \rangle = \mathbf{\Pi} \cdot \mathbf{x}$ is the usual dot product), then the $(-)$ Lie–Poisson bracket becomes the rigid-body bracket.

Hamiltonian Vector Fields. On a Poisson manifold $(P, \{\cdot\,,\cdot\})$, associated to any function H there is a vector field, denoted by X_H, which has the property that for any smooth function $F : P \to \mathbb{R}$ we have the identity

$$\langle \mathbf{d}F, X_H \rangle = \mathbf{d}F \cdot X_H = \{F, H\},$$

where $\mathbf{d}F$ is the differential of F and $\mathbf{d}F \cdot X_H$ denotes the derivative of F in the direction X_H. We say that the vector field X_H is ***generated*** by the function H, or that X_H is the ***Hamiltonian vector field*** associated with H. We also define the associated ***dynamical system*** whose points z in phase space evolve in time by the differential equation

$$\dot{z} = X_H(z). \tag{1.3.8}$$

This definition is consistent with the equations in Poisson bracket form (1.1.16). The function H may have the interpretation of the energy of the system, but of course the definition (1.3.8) makes sense for *any* function. For canonical systems with the Poisson bracket given by (1.1.17), X_H is given by the formula

$$X_H(q^i, p_i) = \left(\frac{\partial H}{\partial p_i}, -\frac{\partial H}{\partial q^i} \right), \tag{1.3.9}$$

whereas for the rigid-body bracket given on $\mathbb{R}^3$ by (1.2.9),

$$X_H(\mathbf{\Pi}) = \mathbf{\Pi} \times \nabla H(\mathbf{\Pi}). \tag{1.3.10}$$

The general Lie–Poisson equations, determined by $\dot{F} = \{F, H\}$, read

$$\dot{\mu}_a = \mp \sum_{b,c=1}^{r} \mu_d C_{ab}^d \frac{\partial H}{\partial \mu_b},$$

or intrinsically,

$$\dot{\mu} = \mp \operatorname{ad}^*_{\delta H/\delta \mu} \mu. \tag{1.3.11}$$

Reduction. There is an important feature of the rigid-body bracket that also carries over to more general Lie algebras, namely, *Lie–Poisson brackets arise from canonical brackets on the cotangent bundle* (phase space) T^*G associated with a Lie group G that has $\mathfrak{g}$ as its associated Lie algebra. (The

general theory of Lie groups is presented in Chapter 9.) Specifically, there is a general construction underlying the association

$$(\theta, \varphi, \psi, p_\theta, p_\varphi, p_\psi) \mapsto (\Pi_1, \Pi_2, \Pi_3) \tag{1.3.12}$$

defined by

$$\begin{aligned} \Pi_1 &= \frac{1}{\sin\theta}[(p_\varphi - p_\psi \cos\theta)\sin\psi + p_\theta \sin\theta \cos\psi], \\ \Pi_2 &= \frac{1}{\sin\theta}[(p_\varphi - p_\psi \cos\theta)\cos\psi - p_\theta \sin\theta \sin\psi], \\ \Pi_3 &= p_\psi. \end{aligned} \tag{1.3.13}$$

This rigid-body map takes the canonical bracket in the variables (θ, φ, ψ) and their conjugate momenta $(p_\theta, p_\varphi, p_\psi)$ to the $(-)$ Lie–Poisson bracket in the following sense. If F and K are functions of Π_1, Π_2, Π_3, they determine functions of $(\theta, \varphi, \psi, p_\theta, p_\varphi, p_\psi)$ by substituting (1.3.13). Then a (tedious but straightforward) exercise using the chain rule shows that

$$\{F, K\}_{(-)\{\text{Lie–Poisson}\}} = \{F, K\}_{\text{canonical}}. \tag{1.3.14}$$

We say that the map defined by (1.3.13) is a ***canonical map*** or a ***Poisson map*** and that the $(-)$ Lie–Poisson bracket has been obtained from the canonical bracket by ***reduction***.

For a rigid body free to rotate about its center of mass, G is the (proper) rotation group SO(3), and the Euler angles and their conjugate momenta are coordinates for T^*G. The choice of T^*G as the primitive phase space is made according to the classical procedures of mechanics: The configuration space SO(3) is chosen, since each element $A \in \mathrm{SO}(3)$ describes the orientation of the rigid body relative to a reference configuration, that is, the rotation A maps the reference configuration to the current configuration. For the description using Lagrangian mechanics, one forms the velocity–phase space $T\,\mathrm{SO}(3)$ with coordinates $(\theta, \varphi, \psi, \dot\theta, \dot\varphi, \dot\psi)$. The Hamiltonian description is obtained as in §1.1 by using the Legendre transform that maps TG to T^*G.

The passage from T^*G to the space of $\mathbf{\Pi}$'s (body angular momentum space) given by (1.3.13) turns out to be determined by *left* translation on the group. This mapping is an example of a *momentum map*, that is, a mapping whose components are the "Noether quantities" associated with a symmetry group. That the map (1.3.13) is a Poisson (canonical) map (see equation (1.3.14)) *is a general fact about momentum maps* proved in §12.6. To get to *space coordinates* one would use *right* translations and the $(+)$ bracket. This is what is done to get the standard description of fluid dynamics.

Momentum Maps and Coadjoint Orbits. From the general rigid-body equations, $\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \nabla H$, we see that

$$\frac{d}{dt}\|\mathbf{\Pi}\|^2 = 0.$$

In other words, Lie–Poisson systems on $\mathbb{R}^3$ conserve the total angular momenta, that is, they leave the spheres in $\mathbf{\Pi}$-space invariant. The generalization of these objects associated to arbitrary Lie algebras are called ***coadjoint orbits***.

Coadjoint orbits are submanifolds of $\mathfrak{g}^*$ with the property that any Lie–Poisson system $\dot{F} = \{F, H\}$ leaves them invariant. We shall also see how these spaces are Poisson manifolds in their own right and are related to the right (+) or left (−) invariance of the system regarded on T^*G, and the corresponding conserved Noether quantities.

On a general Poisson manifold $(P, \{\cdot\,,\cdot\})$, the definition of a momentum map is as follows. We assume that a Lie group G with Lie algebra $\mathfrak{g}$ acts on P by canonical transformations. As we shall review later (see Chapter 9), the infinitesimal way of specifying the action is to associate to each Lie algebra element $\xi \in \mathfrak{g}$ a vector field ξ_P on P. A ***momentum map*** is a map $\mathbf{J} : P \to \mathfrak{g}^*$ with the property that for every $\xi \in \mathfrak{g}$, the function $\langle \mathbf{J}, \xi \rangle$ (the pairing of the $\mathfrak{g}^*$-valued function $\mathbf{J}$ with the vector ξ) generates the vector field ξ_P; that is,

$$X_{\langle \mathbf{J}, \xi \rangle} = \xi_P.$$

As we shall see later, this definition generalizes the usual notions of linear and angular momentum. The rigid body shows that the notion has much wider interest. A fundamental fact about momentum maps is that if the Hamiltonian H is invariant under the action of the group G, then the vector-valued function $\mathbf{J}$ is a constant of the motion for the dynamics of the Hamiltonian vector field X_H associated to H.

One of the important notions related to momentum maps is that of ***infinitesimal equivariance***, or the ***classical commutation*** relations, which state that

$$\{\langle \mathbf{J}, \xi \rangle, \langle \mathbf{J}, \eta \rangle\} = \langle \mathbf{J}, [\xi, \eta] \rangle \tag{1.3.15}$$

for all Lie algebra elements ξ and η. Relations like this are well known for the angular momentum and can be directly checked using the Lie algebra of the rotation group. Later, in Chapter 12, we shall see that the relations (1.3.15) hold for a large important class of momentum maps that are given by computable formulas. Remarkably, it is the condition (1.3.15) that is exactly what is needed to prove that $\mathbf{J}$ *is, in fact, a Poisson map.* It is via this route that one gets an intellectually satisfying generalization of the fact that the map defined by equations (1.3.13) is a Poisson map; that is, equation (1.3.14) holds.

Some History. The Lie–Poisson bracket was discovered by Sophus Lie (Lie [1890, Vol. II, p. 237]). However, Lie's bracket and his related work was not given much attention until the work of Kirillov, Kostant, and Souriau (and others) revived it in the mid-1960s. Meanwhile, it was noticed by Pauli and Martin around 1950 that the rigid-body equations are in Hamiltonian

form using the rigid-body bracket, but they were apparently unaware of the underlying Lie theory. Meanwhile, the generalization of the Euler equations to any Lie algebra $\mathfrak{g}$ by Poincaré [1901b] (and picked up by Hamel [1904]) proceeded as well, but without much contact with Lie's work until recently. The symplectic structure on coadjoint orbits also has a complicated history and itself goes back to Lie (Lie [1890, Ch. 20]).

The general notion of a Poisson manifold also goes back to Lie. However, the four defining properties of the Poisson bracket have been isolated by many authors such as Dirac [1964, p. 10]. The term "Poisson manifold" was coined by Lichnerowicz [1977]. We shall give more historical information on Poisson manifolds in §10.3.

The notion of the momentum map (the English translation of the French words "application moment") also has roots going back to the work of Lie.[4]

Momentum maps have found an astounding array of applications beyond those already mentioned. For instance, they are used in the study of the space of all solutions of a relativistic field theory (see Arms, Marsden, and Moncrief [1982]) and in the study of singularities in algebraic geometry (see Atiyah [1983] and Kirwan [1984]). They also enter into convex analysis in many interesting ways, such as the Schur–Horn theorem (Schur [1923], Horn [1954]) and its generalizations (Kostant [1973]) and in the theory of integrable systems (Bloch, Brockett, and Ratiu [1990, 1992] and Bloch, Flaschka, and Ratiu [1990, 1993]). It turns out that the image of the momentum map has remarkable convexity properties: see Atiyah [1982], Guillemin and Sternberg [1982, 1984], Kirwan [1984], Delzant [1988], and Lu and Ratiu [1991].

Exercises

◇ **1.3-1.** A linear operator D on the space of smooth functions on $\mathbb{R}^n$ is called a ***derivation*** if it satisfies the Leibniz identity: $D(FG) = (DF)G + F(DG)$. Accept the fact from the theory of manifolds (see Chapter 4) that in local coordinates the expression of DF takes the form

$$(DF)(x) = \sum_{i=1}^{n} a^i(x) \frac{\partial F}{\partial x^i}(x)$$

for some smooth functions $a^1, \ldots, a^n$.

[4] Many authors use the words "moment map" for what we call the "momentum map." In English, unlike French, one does not use the phrases "linear moment" or "angular moment of a particle," and correspondingly, we prefer to use "momentum map." We shall give some comments on the history of momentum maps in §11.2.

(a) Use the fact just stated to prove that for any bilinear operation $\{,\}$ on $\mathcal{F}(\mathbb{R}^n)$ which is a derivation in each of its arguments, we have

$$\{F,G\} = \sum_{i,j=1}^{n} \{x^i, x^j\} \frac{\partial F}{\partial x^i} \frac{\partial G}{\partial x^j}.$$

(b) Show that the Jacobi identity holds for any operation $\{,\}$ on $\mathcal{F}(\mathbb{R}^n)$ as in (a), if and only if it holds for the coordinate functions.

◇ **1.3-2.** Define, for a fixed function $f : \mathbb{R}^3 \to \mathbb{R}$,

$$\{F,K\}_f = \nabla f \cdot (\nabla F \times \nabla K).$$

(a) Show that this is a Poisson bracket.

(b) Locate the bracket in part (a) in Nambu [1973].

◇ **1.3-3.** Verify directly that (1.3.13) defines a Poisson map.

◇ **1.3-4.** Show that a bracket satisfying the Leibniz identity also satisfies

$$F\{K,L\} - \{FK,L\} = \{F,K\}L - \{F,KL\}.$$

1.4 The Heavy Top

The equations of motion for a rigid body with a fixed point in a gravitational field provide another interesting example of a system that is Hamiltonian relative to a Lie–Poisson bracket. See Figure 1.4.1.

The underlying Lie algebra consists of the algebra of infinitesimal Euclidean motions in $\mathbb{R}^3$. (These do *not* arise as Euclidean motions of the body, since the body has a fixed point.) As we shall see, there is a close parallel with the Poisson structure for compressible fluids.

The basic phase space we start with is again $T^*\,\mathrm{SO}(3)$, coordinatized by Euler angles and their conjugate momenta. In these variables, the equations are in canonical Hamiltonian form; however, the presence of gravity breaks the symmetry, and the system is no longer SO(3) invariant, so it cannot be written entirely in terms of the body angular momentum $\mathbf{\Pi}$. One also needs to keep track of $\mathbf{\Gamma}$, the "direction of gravity" as seen from the body. This is defined by $\mathbf{\Gamma} = \mathbf{A}^{-1}\mathbf{k}$, where $\mathbf{k}$ points upward and $\mathbf{A}$ is the element of SO(3) describing the current configuration of the body. The equations of motion are

$$\begin{aligned}
\dot{\Pi}_1 &= \frac{I_2 - I_3}{I_2 I_3} \Pi_2 \Pi_3 + Mgl(\Gamma^2 \chi^3 - \Gamma^3 \chi^2), \\
\dot{\Pi}_2 &= \frac{I_3 - I_1}{I_3 I_1} \Pi_3 \Pi_1 + Mgl(\Gamma^3 \chi^1 - \Gamma^1 \chi^3), \\
\dot{\Pi}_3 &= \frac{I_1 - I_2}{I_1 I_2} \Pi_1 \Pi_2 + Mgl(\Gamma^1 \chi^2 - \Gamma^2 \chi^1),
\end{aligned} \tag{1.4.1}$$

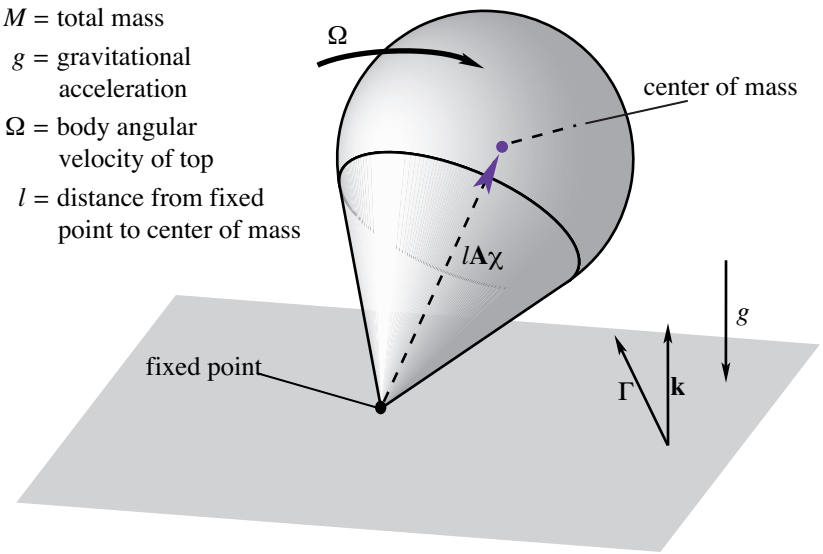

FIGURE 1.4.1. Heavy top

and

$$\dot{\mathbf{\Gamma}} = \mathbf{\Gamma} \times \mathbf{\Omega}, \tag{1.4.2}$$

where M is the body's mass, g is the acceleration of gravity, $\boldsymbol{\chi}$ is the body fixed unit vector on the line segment connecting the fixed point with the body's center of mass, and l is the length of this segment. See Figure 1.4.1.

The Lie algebra of the Euclidean group is $\mathfrak{se}(3) = \mathbb{R}^3 \times \mathbb{R}^3$ with the Lie bracket

$$[(\boldsymbol{\xi}, \mathbf{u}), (\boldsymbol{\eta}, \mathbf{v})] = (\boldsymbol{\xi} \times \boldsymbol{\eta}, \boldsymbol{\xi} \times \mathbf{v} - \boldsymbol{\eta} \times \mathbf{u}). \tag{1.4.3}$$

We identify the dual space with pairs $(\mathbf{\Pi}, \mathbf{\Gamma})$; the corresponding $(-)$ Lie–Poisson bracket, called the ***heavy top bracket***, is

$$\begin{aligned} \{F, G\}(\mathbf{\Pi}, \mathbf{\Gamma}) = &-\mathbf{\Pi} \cdot (\nabla_{\mathbf{\Pi}} F \times \nabla_{\mathbf{\Pi}} G) \\ &- \mathbf{\Gamma} \cdot (\nabla_{\mathbf{\Pi}} F \times \nabla_{\mathbf{\Gamma}} G - \nabla_{\mathbf{\Pi}} G \times \nabla_{\mathbf{\Gamma}} F). \end{aligned} \tag{1.4.4}$$

The above equations for $\mathbf{\Pi}, \mathbf{\Gamma}$ can be checked to be equivalent to

$$\dot{F} = \{F, H\}, \tag{1.4.5}$$

where the ***heavy top Hamiltonian***

$$H(\mathbf{\Pi}, \mathbf{\Gamma}) = \frac{1}{2}\left(\frac{\Pi_1^2}{I_1} + \frac{\Pi_2^2}{I_2} + \frac{\Pi_3^2}{I_3}\right) + Mgl\mathbf{\Gamma} \cdot \boldsymbol{\chi} \tag{1.4.6}$$

is the total energy of the body (Sudarshan and Mukunda [1974]).

The Lie algebra of the Euclidean group has a structure that is a special case of what is called a *semidirect product*. Here it is the product of the group of rotations with the translation group. It turns out that semidirect products occur under rather general circumstances when the symmetry in T^*G is broken. The general theory for semidirect products was developed by Sudarshan and Mukunda [1974], Ratiu [1980, 1981, 1982], Guillemin and Sternberg [1982], Marsden, Weinstein, Ratiu, Schmid, and Spencer [1983], Marsden, Ratiu, and Weinstein [1984a, 1984b], and Holm and Kupershmidt [1983]. The Lagrangian approach to this and related problems is given in Holm, Marsden, and Ratiu [1998a].

Exercises

◇ **1.4-1.** Verify that $\dot{F} = \{F, H\}$ is equivalent to the heavy top equations using the heavy top Hamiltonian and bracket.

◇ **1.4-2.** Work out the Euler–Poincaré equations on $\mathfrak{se}(3)$. Show that with

$$L(\mathbf{\Omega}, \mathbf{\Gamma}) = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2) - Mgl\mathbf{\Gamma} \cdot \boldsymbol{\chi},$$

the Euler–Poincaré equations are *not* the heavy top equations.

1.5 Incompressible Fluids

Arnold [1966a, 1969] showed that the Euler equations for an incompressible fluid could be given a Lagrangian and Hamiltonian description similar to that for the rigid body. His approach[5] has the appealing feature that one sets things up just the way Lagrange and Hamilton would have done: One begins with a configuration space Q and forms a Lagrangian L on the velocity phase space TQ and then H on the momentum phase space T^*Q, just as was outlined in §1.1. Thus, one automatically has variational principles, etc. For ideal fluids, $Q = G$ is the group $\mathrm{Diff}_{\mathrm{vol}}(\Omega)$ of volume-preserving transformations of the fluid container (a region Ω in $\mathbb{R}^2$ or $\mathbb{R}^3$, or a Riemannian manifold in general, possibly with boundary). Group multiplication in G is composition.

Kinematics of a Fluid. The reason we select $G = \mathrm{Diff}_{\mathrm{vol}}(\Omega)$ as the configuration space is similar to that for the rigid body; namely, each φ in G is a mapping of Ω to Ω that takes a reference point $X \in \Omega$ to a

[5] Arnold's approach is consistent with what appears in the thesis of Ehrenfest from around 1904; see Klein [1970]. However, Ehrenfest bases his principles on the more sophisticated curvature principles of Gauss and Hertz.

current point $x = \varphi(X) \in \Omega$; thus, knowing φ tells us where each particle of fluid goes and hence gives us the ***fluid configuration***. We ask that φ be a diffeomorphism to exclude discontinuities, cavitation, and fluid interpenetration, and we ask that φ be volume-preserving to correspond to the assumption of incompressibility.

A ***motion*** of a fluid is a family of time-dependent elements of G, which we write as $x = \varphi(X, t)$. The ***material velocity field*** is defined by

$$\mathbf{V}(X, t) = \frac{\partial \varphi(X, t)}{\partial t},$$

and the ***spatial velocity field*** is defined by $\mathbf{v}(x, t) = \mathbf{V}(X, t)$, where x and X are related by $x = \varphi(X, t)$. If we suppress "t" and write $\dot{\varphi}$ for $\mathbf{V}$, note that

$$\mathbf{v} = \dot{\varphi} \circ \varphi^{-1}, \quad \text{i.e.,} \quad \mathbf{v}_t = \mathbf{V}_t \circ \varphi_t^{-1}, \tag{1.5.1}$$

where $\varphi_t(x) = \varphi(X, t)$. See Figure 1.5.1.

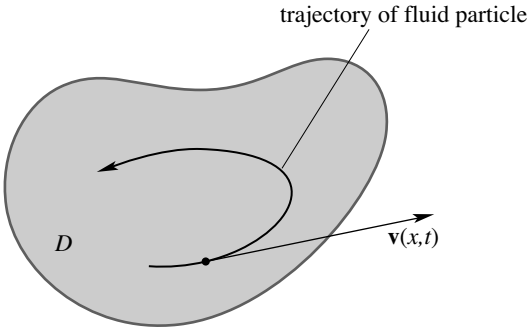

FIGURE 1.5.1. The trajectory and velocity of a fluid particle.

We can regard (1.5.1) as a map from the space of $(\varphi, \dot{\varphi})$ (material or Lagrangian description) to the space of $\mathbf{v}$'s (spatial or Eulerian description). Like the rigid body, the material to spatial map (1.5.1) takes the canonical bracket to a Lie–Poisson bracket; one of our goals is to understand this reduction. Notice that if we replace φ by $\varphi \circ \eta$ for a fixed (time-independent) $\eta \in \mathrm{Diff}_{\mathrm{vol}}(\Omega)$, then $\dot{\varphi} \circ \varphi^{-1}$ is independent of η; this reflects the *right* invariance of the Eulerian description ($\mathbf{v}$ is invariant under composition of φ by η on the right). This is also called the ***particle relabeling symmetry*** of fluid dynamics. The spaces TG and T^*G represent the Lagrangian (material) description, and we pass to the Eulerian (spatial) description by right translations and use the (+) Lie–Poisson bracket. One of the things we want to do later is to better understand the reason for the switch between right and left in going from the rigid body to fluids.

Dynamics of a Fluid. The ***Euler equations*** for an ideal, incompressible, homogeneous fluid moving in the region Ω are

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\nabla p \tag{1.5.2}$$

with the constraint div $\mathbf{v} = 0$ and the boundary condition that $\mathbf{v}$ is tangent to the boundary, $\partial\Omega$.

The ***pressure*** p is determined implicitly by the divergence-free (volume-preserving) constraint div $\mathbf{v} = 0$. (See Chorin and Marsden [1993] for basic information on the derivation of Euler's equations.) The associated Lie algebra $\mathfrak{g}$ is the space of all divergence-free vector fields tangent to the boundary. This Lie algebra is endowed with the *negative* ***Jacobi–Lie bracket*** of vector fields given by

$$[v, w]_L^i = \sum_{j=1}^{n} \left(w^j \frac{\partial v^i}{\partial x^j} - v^j \frac{\partial w^i}{\partial x^j} \right). \tag{1.5.3}$$

(The subscript L on $[\cdot\,,\cdot]$ refers to the fact that it is the *left* Lie algebra bracket on $\mathfrak{g}$. The most common convention for the Jacobi–Lie bracket of vector fields, also the one we adopt, has the opposite sign.) We identify $\mathfrak{g}$ and $\mathfrak{g}^*$ using the pairing

$$\langle \mathbf{v}, \mathbf{w} \rangle = \int_\Omega \mathbf{v} \cdot \mathbf{w}\, d^3x. \tag{1.5.4}$$

Hamiltonian Structure. Introduce the (+) Lie–Poisson bracket, called the ***ideal fluid bracket***, on functions of $\mathbf{v}$ by

$$\{F, G\}(\mathbf{v}) = \int_\Omega \mathbf{v} \cdot \left[\frac{\delta F}{\delta \mathbf{v}}, \frac{\delta G}{\delta \mathbf{v}} \right]_L d^3x, \tag{1.5.5}$$

where $\delta F/\delta \mathbf{v}$ is defined by

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}[F(\mathbf{v} + \varepsilon \delta \mathbf{v}) - F(\mathbf{v})] = \int_\Omega \left(\delta \mathbf{v} \cdot \frac{\delta F}{\delta \mathbf{v}} \right) d^3x. \tag{1.5.6}$$

With the energy function chosen to be the kinetic energy,

$$H(\mathbf{v}) = \frac{1}{2} \int_\Omega \|\mathbf{v}\|^2\, d^3x, \tag{1.5.7}$$

one can verify that the Euler equations (1.5.2) are equivalent to the Poisson bracket equations

$$\dot{F} = \{F, H\} \tag{1.5.8}$$

for all functions F on $\mathfrak{g}$. To see this, it is convenient to use the orthogonal decomposition $\mathbf{w} = \mathbb{P}\mathbf{w} + \nabla p$ of a vector field $\mathbf{w}$ into a divergence-free part $\mathbb{P}\mathbf{w}$ in $\mathfrak{g}$ and a gradient. The Euler equations can be written

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbb{P}(\mathbf{v} \cdot \nabla \mathbf{v}) = 0. \tag{1.5.9}$$

One can express the Hamiltonian structure in terms of the vorticity as a basic dynamic variable and show that the preservation of coadjoint orbits amounts to Kelvin's circulation theorem. Marsden and Weinstein [1983] show that the Hamiltonian structure in terms of Clebsch potentials fits naturally into this Lie–Poisson scheme, and that Kirchhoff's Hamiltonian description of point vortex dynamics, vortex filaments, and vortex patches can be derived in a natural way from the Hamiltonian structure described above.

Lagrangian Structure. The general framework of the Euler–Poincaré and the Lie–Poisson equations gives other insights as well. For example, this general theory shows that the Euler equations are derivable from the "variational principle"

$$\delta \int_a^b \int_\Omega \frac{1}{2} \|\mathbf{v}\|^2 \, d^3x = 0,$$

which is to hold for all variations $\delta \mathbf{v}$ of the form

$$\delta \mathbf{v} = \dot{\mathbf{u}} + [\mathbf{v}, \mathbf{u}]_L$$

(sometimes called ***Lin constraints***), where $\mathbf{u}$ is a vector field (representing the infinitesimal particle displacement) vanishing at the temporal endpoints.[6]

There are important functional-analytic differences between working in material representation (that is, on T^*G) and in Eulerian representation (that is, on $\mathfrak{g}^*$) that are important for proving existence and uniqueness theorems, theorems on the limit of zero viscosity, and the convergence of numerical algorithms (see Ebin and Marsden [1970], Marsden, Ebin, and Fischer [1972], and Chorin, Hughes, Marsden, and McCracken [1978]). Finally, we note that for *two-dimensional flow*, a collection of Casimir functions is given by

$$C(\omega) = \int_\Omega \Phi(\omega(x)) \, d^2x \tag{1.5.10}$$

for $\Phi : \mathbb{R} \to \mathbb{R}$ any (smooth) function, where $\omega \mathbf{k} = \nabla \times \mathbf{v}$ is the ***vorticity***. For three-dimensional flow, (1.5.10) is no longer a Casimir.

---

[6] As mentioned earlier, this form of the variational (strictly speaking, a Lagrange–d'Alembert type) principle is due to Newcomb [1962]; see also Bretherton [1970]. For the case of general Lie algebras, it is due to Marsden and Scheurle [1993b]; see also Cendra and Marsden [1987].

Exercises

◇ **1.5-1.** Show that any divergence-free vector field X on $\mathbb{R}^3$ can be written *globally* as a curl of another vector field and, away from equilibrium points, can *locally* be written as

$$X = \nabla f \times \nabla g,$$

where f and g are real-valued functions on $\mathbb{R}^3$. Assume that this (so-called Clebsch–Monge) representation also holds globally. Particles of fluid follow trajectories satisfying the equation $\dot{x} = X(x)$. Show that these trajectories can be described by a Hamiltonian system with a bracket in the form of Exercise 1.3-2.

1.6 The Maxwell–Vlasov System

Plasma physics provides another beautiful application area for the techniques discussed in the preceding sections. We shall briefly indicate these in this section. The period 1970–1980 saw the development of noncanonical Hamiltonian structures for the Korteweg–de Vries (KdV) equation (due to Gardner, Kruskal, Miura, and others; see Gardner [1971]) and other soliton equations. This quickly became entangled with the attempts to understand integrability of Hamiltonian systems and the development of the algebraic approach; see, for example, Gelfand and Dorfman [1979], Manin [1979] and references therein. More recently, these approaches have come together again; see, for instance, Reyman and Semenov-Tian-Shansky [1990], Moser and Veselov [1991]. KdV type models are usually derived from or are approximations to more fundamental fluid models, and it seems fair to say that the reasons for their complete integrability are not yet completely understood.

Some History. For fluid and plasma systems, some of the key early works on Poisson bracket structures were Dashen and Sharp [1968], Goldin [1971], Iwiński and Turski [1976], Dzyaloshinskii and Volovick [1980], Morrison and Greene [1980], and Morrison [1980]. In Sudarshan and Mukunda [1974], Guillemin and Sternberg [1982], and Ratiu [1980, 1982], a general theory for Lie–Poisson structures for special kinds of Lie algebras, called semidirect products, was begun. This was quickly recognized (see, for example, Marsden [1982], Marsden, Weinstein, Ratiu, Schmid, and Spencer [1983], Holm and Kupershmidt [1983], and Marsden, Ratiu, and Weinstein [1984a, 1984b]) to be relevant to the brackets for compressible flow; see §1.7 below.

Derivation of Poisson Structures. A rational scheme for systematically *deriving* brackets is needed since for one thing, a direct verification of Jacobi's identity can be inefficient and time–consuming. Here we outline a derivation of the Maxwell–Vlasov bracket by Marsden and Weinstein

[1982]. The method is similar to Arnold's, namely by performing a reduction starting with:

(i) canonical brackets in a material representation for the plasma; and

(ii) a potential representation for the electromagnetic field.

One then identifies the symmetry group and carries out reduction by this group in a manner similar to that we described for Lie–Poisson systems.

For plasmas, the physically correct material description is actually slightly more complicated; we refer to Cendra, Holm, Hoyle, and Marsden [1998] for a full account.

Parallel developments can be given for many other brackets, such as the charged fluid bracket by Spencer and Kaufman [1982]. Another method, based primarily on Clebsch potentials, was developed in a series of papers by Holm and Kupershmidt (for example, Holm and Kupershmidt [1983]) and applied to a number of interesting systems, including superfluids and superconductors. They also pointed out that semidirect products are appropriate for the MHD bracket of Morrison and Greene [1980].

The Maxwell–Vlasov System. The Maxwell–Vlasov equations for a collisionless plasma are the fundamental equations in plasma physics.[7] In Euclidean space, the basic dynamical variables are

$$\begin{array}{rcl} f(\mathbf{x},\mathbf{v},t) & : & \text{the plasma particle number density per phase space} \\ & & \text{volume } d^3x\, d^3v; \\ \mathbf{E}(\mathbf{x},t) & : & \text{the electric field;} \\ \mathbf{B}(\mathbf{x},t) & : & \text{the magnetic field.} \end{array}$$

The equations for a collisionless plasma for the case of a single species of particles with mass m and charge e are

$$\begin{aligned} &\frac{\partial f}{\partial t} + \mathbf{v}\cdot\frac{\partial f}{\partial \mathbf{x}} + \frac{e}{m}\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right)\cdot\frac{\partial f}{\partial \mathbf{v}} = 0, \\ &\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = -\operatorname{curl}\mathbf{E}, \\ &\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \operatorname{curl}\mathbf{B} - \frac{1}{c}\mathbf{j}_f, \\ &\operatorname{div}\mathbf{E} = \rho_f, \\ &\operatorname{div}\mathbf{B} = 0. \end{aligned} \tag{1.6.1}$$

The ***current*** defined by f is given by

$$\mathbf{j}_f = e\int \mathbf{v} f(\mathbf{x},\mathbf{v},t)\, d^3v$$

[7] See, for example, Clemmow and Dougherty [1959], van Kampen and Felderhof [1967], Krall and Trivelpiece [1973], Davidson [1972], Ichimaru [1973], and Chen [1974].

and the ***charge density*** by

$$\rho_f = e \int f(\mathbf{x}, \mathbf{v}, t)\, d^3v.$$

Also, $\partial f/\partial \mathbf{x}$ and $\partial f/\partial \mathbf{v}$ denote the gradients of f with respect to $\mathbf{x}$ and $\mathbf{v}$, respectively, and c is the speed of light. The evolution equation for f results from the Lorentz force law and standard transport assumptions. The remaining equations are the standard Maxwell equations with charge density ρ_f and current $\mathbf{j}_f$ produced by the plasma.

Two limiting cases will aid our discussions. First, if the plasma is constrained to be static, that is, f is concentrated at $\mathbf{v} = 0$ and t-independent, we get the ***charge-driven Maxwell equations***:

$$\begin{aligned} &\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = -\operatorname{curl} \mathbf{E}, \\ &\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \operatorname{curl} \mathbf{B}, \\ &\operatorname{div} \mathbf{E} = \rho, \quad \text{and} \quad \operatorname{div} \mathbf{B} = 0. \end{aligned} \tag{1.6.2}$$

Second, if we let $c \to \infty$, electrodynamics becomes electrostatics, and we get the ***Poisson–Vlasov equation***

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} - \frac{e}{m}\frac{\partial \varphi_f}{\partial \mathbf{x}} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0, \tag{1.6.3}$$

where $-\nabla^2 \varphi_f = \rho_f$. In this context, the name "Poisson–Vlasov" seems quite appropriate. The equation is, however, formally the same as the earlier Jeans [1919] equation of stellar dynamics. Henon [1982] has proposed calling it the "collisionless Boltzmann equation."

Maxwell's Equations. For simplicity, we let $m = e = c = 1$. As the basic configuration space we take the space $\mathcal{A}$ of vector potentials $\mathbf{A}$ on $\mathbb{R}^3$ (for the Yang–Mills equations this is generalized to the space of connections on a principal bundle over space). The corresponding phase space $T^*\mathcal{A}$ is identified with the set of pairs $(\mathbf{A}, \mathbf{Y})$, where $\mathbf{Y}$ is also a vector field on $\mathbb{R}^3$. The canonical Poisson bracket is used on $T^*\mathcal{A}$:

$$\{F, G\} = \int \left(\frac{\delta F}{\delta \mathbf{A}} \frac{\delta G}{\delta \mathbf{Y}} - \frac{\delta F}{\delta \mathbf{Y}} \frac{\delta G}{\delta \mathbf{A}} \right) d^3x. \tag{1.6.4}$$

The ***electric field*** is $\mathbf{E} = -\mathbf{Y}$, and the ***magnetic field*** is $\mathbf{B} = \operatorname{curl} \mathbf{A}$. With the Hamiltonian

$$H(\mathbf{A}, \mathbf{Y}) = \frac{1}{2} \int (\|\mathbf{E}\|^2 + \|\mathbf{B}\|^2)\, d^3x, \tag{1.6.5}$$

Hamilton's canonical field equations (1.1.14) are checked to give the equations for $\partial \mathbf{E}/\partial t$ and $\partial \mathbf{A}/\partial t$, which imply the vacuum Maxwell's equations.

Alternatively, one can begin with $T\mathcal{A}$ and the Lagrangian

$$L(\mathbf{A}, \dot{\mathbf{A}}) = \frac{1}{2} \int \left(\|\dot{\mathbf{A}}\|^2 - \|\nabla \times \mathbf{A}\|^2 \right) d^3x \tag{1.6.6}$$

and use the Euler–Lagrange equations and variational principles.

It is of interest to incorporate the equation div $\mathbf{E} = \rho$ and, correspondingly, to use directly the field strengths $\mathbf{E}$ and $\mathbf{B}$, rather than $\mathbf{E}$ and $\mathbf{A}$. To do this, we introduce the ***gauge group*** $\mathcal{G}$, the additive group of real-valued functions $\psi : \mathbb{R}^3 \to \mathbb{R}$. Each $\psi \in \mathcal{G}$ transforms the fields according to the rule

$$(\mathbf{A}, \mathbf{E}) \mapsto (\mathbf{A} + \nabla\psi, \mathbf{E}). \tag{1.6.7}$$

Each such transformation leaves the Hamiltonian H invariant and is a canonical transformation, that is, it leaves Poisson brackets intact. In this situation, as above, there will be a corresponding conserved quantity, or *momentum map*, in the same sense as in §1.3. As mentioned there, some simple general formulas for computing momentum maps will be studied in detail in Chapter 12. For the action (1.6.7) of $\mathcal{G}$ on $T^*\mathcal{A}$, the associated momentum map is

$$\mathbf{J}(\mathbf{A}, \mathbf{Y}) = \text{div } \mathbf{E}, \tag{1.6.8}$$

so we recover the fact that div $\mathbf{E}$ is preserved by Maxwell's equations (this is easy to verify directly using the identity div curl $= 0$). Thus we see that we can incorporate the equation div $\mathbf{E} = \rho$ by restricting our attention to the set $\mathbf{J}^{-1}(\rho)$. The theory of reduction is a general process whereby one reduces the dimension of a phase space by exploiting conserved quantities and symmetry groups. In the present case, the reduced space is $\mathbf{J}^{-1}(\rho)/\mathcal{G}$, which is identified with Max_ρ, the space of $\mathbf{E}$'s and $\mathbf{B}$'s satisfying div $\mathbf{E} = \rho$ and div $\mathbf{B} = 0$.

The space Max_ρ inherits a Poisson structure as follows. If F and K are functions on Max_ρ, we substitute $\mathbf{E} = -\mathbf{Y}$ and $\mathbf{B} = \nabla \times \mathbf{A}$ to express F and K as functionals of $(\mathbf{A}, \mathbf{Y})$. Then we compute the canonical brackets on $T^*\mathcal{A}$ and express the result in terms of $\mathbf{E}$ and $\mathbf{B}$. Carrying this out using the chain rule gives

$$\{F, K\} = \int \left(\frac{\delta F}{\delta \mathbf{E}} \cdot \text{curl} \frac{\delta K}{\delta \mathbf{B}} - \frac{\delta K}{\delta \mathbf{E}} \cdot \text{curl} \frac{\delta F}{\delta \mathbf{B}} \right) d^3x, \tag{1.6.9}$$

where $\delta F/\delta \mathbf{E}$ and $\delta F/\delta \mathbf{B}$ are vector fields, with $\delta F/\delta \mathbf{B}$ *divergence-free*. These are defined in the usual way; for example,

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [F(\mathbf{E} + \varepsilon\delta\mathbf{E}, \mathbf{B}) - F(\mathbf{E}, \mathbf{B})] = \int \frac{\delta F}{\delta \mathbf{E}} \cdot \delta\mathbf{E} \, d^3x. \tag{1.6.10}$$

This bracket makes Max_ρ into a Poisson manifold and the map $(\mathbf{A}, \mathbf{Y}) \mapsto (-\mathbf{Y}, \nabla \times \mathbf{A})$ into a Poisson map. The bracket (1.6.9) was discovered (by

a different procedure) by Pauli [1933] and Born and Infeld [1935]. We refer to (1.6.9) as the ***Pauli–Born–Infeld bracket*** or the ***Maxwell–Poisson bracket*** for Maxwell's equations.

With the energy H given by (1.6.5) regarded as a function of $\mathbf{E}$ and $\mathbf{B}$, Hamilton's equations in bracket form $\dot{F} = \{F, H\}$ on Max_ρ capture the full set of Maxwell's equations (with external charge density ρ).

The Poisson–Vlasov Equation. The papers Iwiński and Turski [1976] and Morrison [1980] showed that the Poisson–Vlasov equations form a Hamiltonian system with

$$H(f) = \frac{1}{2}\int \|\mathbf{v}\|^2 f(\mathbf{x},\mathbf{v},t)\, d^3x\, d^3v + \frac{1}{2}\int \|\nabla\varphi_f\|^2\, d^3x \tag{1.6.11}$$

and the ***Poisson–Vlasov bracket***

$$\{F, G\} = \int f \left\{ \frac{\delta F}{\delta f}, \frac{\delta G}{\delta f} \right\}_{\mathbf{xv}} d^3x\, d^3v, \tag{1.6.12}$$

where $\{\,,\}_{\mathbf{xv}}$ is the canonical bracket on $(\mathbf{x},\mathbf{v})$-space. As was observed in Gibbons [1981] and Marsden and Weinstein [1982], this is the (+) Lie–Poisson bracket associated with the Lie algebra $\mathfrak{g}$ of functions of $(\mathbf{x},\mathbf{v})$ with Lie bracket the canonical Poisson bracket.

According to the general theory, this Lie–Poisson structure is obtained by reduction from canonical brackets on the cotangent bundle of the group underlying $\mathfrak{g}$, just as was the case for the rigid body and incompressible fluids. This time, the group $G = \mathrm{Diff}_{\mathrm{can}}$ is the group of canonical transformations of $(\mathbf{x},\mathbf{v})$-space. The Poisson–Vlasov equations can equally well be written in canonical form on T^*G. This is related to the Lagrangian and Hamiltonian description of a plasma that goes back to Low [1958], Katz [1961], and Lundgren [1963]. Thus, one can start with the particle description with canonical brackets and, through reduction, derive the brackets here. See Cendra, Holm, Hoyle, and Marsden [1998] for exactly how this goes. There are other approaches to the Hamiltonian formulation using analogues of Clebsch potentials; see, for instance, Su [1961], Zakharov [1971], and Gibbons, Holm, and Kupershmidt [1982].

The Poisson–Vlaslov to Compressible Flow Map. Before going on to the Maxwell–Vlasov equations, we point out a remarkable connection between the Poisson–Vlasov bracket (1.6.12) and the bracket for compressible flow.

The Euler equations for compressible flow in a region Ω in $\mathbb{R}^3$ are

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v}\right) = -\nabla p \tag{1.6.13}$$

and

$$\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho\mathbf{v}) = 0, \tag{1.6.14}$$

with the boundary condition

$$\mathbf{v} \text{ tangent to } \partial\Omega.$$

Here the pressure p is determined from an internal energy function per unit mass given by $p = \rho^2 w'(\rho)$, where $w = w(\rho)$ is the constitutive relation. (We ignore entropy for the present discussion—its inclusion is straightforward to deal with.) The ***compressible fluid Hamiltonian*** is

$$H = \frac{1}{2}\int_\Omega \rho\|\mathbf{v}\|^2\, d^3x + \int_\Omega \rho w(\rho)\, d^3x. \tag{1.6.15}$$

The relevant Poisson bracket is most easily expressed if we use the momentum density $\mathbf{M} = \rho\mathbf{v}$ and density ρ as our basic variables. The ***compressible fluid bracket*** is

$$\begin{aligned}\{F, G\} = &\int_\Omega \mathbf{M}\cdot\left[\left(\frac{\delta G}{\delta \mathbf{M}}\cdot\nabla\right)\frac{\delta F}{\delta \mathbf{M}} - \left(\frac{\delta F}{\delta \mathbf{M}}\cdot\nabla\right)\frac{\delta G}{\delta \mathbf{M}}\right] d^3x \\ &+ \int_\Omega \rho\left[\left(\frac{\delta G}{\delta \mathbf{M}}\cdot\nabla\right)\frac{\delta F}{\delta \rho} - \left(\frac{\delta F}{\delta \mathbf{M}}\cdot\nabla\right)\frac{\delta G}{\delta \rho}\right] d^3x.\end{aligned} \tag{1.6.16}$$

Notice the similarities in structure between the Poisson bracket (1.6.16) for compressible flow and (1.4.4). For compressible flow it is the density that prevents a full $\mathrm{Diff}(\Omega)$ invariance; the Hamiltonian is invariant only under those diffeomorphisms that preserve the density.

The space of $(\mathbf{M}, \rho)$'s can be shown to be the dual of a semidirect product Lie algebra and it can also be shown that the preceding bracket is the associated (+) Lie–Poisson bracket (see Marsden, Weinstein, Ratiu, Schmid, and Spencer [1983], Holm and Kupershmidt [1983], and Marsden, Ratiu, and Weinstein [1984a, 1984b]).

The relationship with the Poisson–Vlasov bracket is this: Suppressing the time variable, define the map $f \mapsto (\mathbf{M}, \rho)$ by

$$\mathbf{M}(\mathbf{x}) = \int_\Omega \mathbf{v} f(\mathbf{x}, \mathbf{v}) d^3v \quad \text{and} \quad \rho(\mathbf{x}) = \int_\Omega f(\mathbf{x}, \mathbf{v})\, d^3v. \tag{1.6.17}$$

Remarkably, this plasma to fluid map is a Poisson map taking the Poisson–Vlasov bracket (1.6.12) to the compressible fluid bracket (1.6.16). In fact, this map is a momentum map (Marsden, Weinstein, Ratiu, Schmid, and Spencer [1983]). The Poisson–Vlasov Hamiltonian is *not* invariant under the associated group action, however.

The Maxwell–Vlasov Bracket. A bracket for the Maxwell–Vlasov equations was given by Iwiński and Turski [1976] and Morrison [1980]. Marsden and Weinstein [1982] used systematic procedures involving reduction and momentum maps to derive (and correct) the bracket starting with a canonical bracket.

The procedure starts with the material description[8] of the plasma as the cotangent bundle of the group $\mathrm{Diff}_{\mathrm{can}}$ of canonical transformations of $(\mathbf{x},\mathbf{p})$-space and the space $T^*\mathcal{A}$ for Maxwell's equations. We justify this by noticing that the motion of a charged particle in a fixed (but possibly time-dependent) electromagnetic field via the Lorentz force law defines a (time-dependent) canonical transformation. On $T^*\,\mathrm{Diff}_{\mathrm{can}}\times T^*\mathcal{A}$ we put the sum of the two canonical brackets, and then we reduce. First we reduce by $\mathrm{Diff}_{\mathrm{can}}$, which acts on $T^*\,\mathrm{Diff}_{\mathrm{can}}$ by right translation but does not act on $T^*\mathcal{A}$. Thus we end up with densities $f_{\mathrm{mom}}(\mathbf{x},\mathbf{p},t)$ on position-momentum space and with the space $T^*\mathcal{A}$ used for the Maxwell equations. On this space we get the (+) Lie–Poisson bracket, plus the canonical bracket on $T^*\mathcal{A}$. Recalling that $\mathbf{p}$ is related to $\mathbf{v}$ and $\mathbf{A}$ by $\mathbf{p}=\mathbf{v}+\mathbf{A}$, we let the gauge group $\mathcal{G}$ of electromagnetism act on this space by

$$\begin{aligned}(f_{\mathrm{mom}}(\mathbf{x},\mathbf{p},t),\mathbf{A}(\mathbf{x},t),\mathbf{Y}(\mathbf{x},t))\mapsto\\ (f_{\mathrm{mom}}(\mathbf{x},\mathbf{p}+\nabla\varphi(\mathbf{x}),t),\mathbf{A}(\mathbf{x},t)+\nabla\varphi(x),\mathbf{Y}(\mathbf{x},t)).\end{aligned}\tag{1.6.18}$$

The momentum map associated with this action is computed to be

$$\mathbf{J}(f_{\mathrm{mom}},\mathbf{A},\mathbf{Y})=\operatorname{div}\mathbf{E}-\int f_{\mathrm{mom}}(\mathbf{x},\mathbf{p})\,d^3p.\tag{1.6.19}$$

This corresponds to $\operatorname{div}\mathbf{E}-\rho_f$ if we write $f(\mathbf{x},\mathbf{v},t)=f_{\mathrm{mom}}(\mathbf{x},\mathbf{p}-\mathbf{A},t)$. This reduced space $\mathbf{J}^{-1}(0)/\mathcal{G}$ can be identified with the space $\mathcal{MV}$ of triples $(f,\mathbf{E},\mathbf{B})$ satisfying $\operatorname{div}\mathbf{E}=\rho_f$ and $\operatorname{div}\mathbf{B}=0$. The bracket on $\mathcal{MV}$ is computed by the same procedure as for Maxwell's equations. These computations yield the following ***Maxwell–Vlasov bracket***:

$$\begin{aligned}\{F,K\}(f,\mathbf{E},\mathbf{B})=&\int f\left\{\frac{\delta F}{\delta f},\frac{\delta K}{\delta f}\right\}_{xv}d^3x\,d^3v\\ &+\int\left(\frac{\delta F}{\delta\mathbf{E}}\cdot\operatorname{curl}\frac{\delta K}{\delta\mathbf{B}}-\frac{\delta K}{\delta\mathbf{E}}\cdot\operatorname{curl}\frac{\delta F}{\delta\mathbf{B}}\right)d^3x\\ &+\int\left(\frac{\delta F}{\delta\mathbf{E}}\cdot\frac{\delta f}{\delta\mathbf{v}}\frac{\delta K}{\delta f}-\frac{\delta K}{\delta\mathbf{E}}\cdot\frac{\delta f}{\delta\mathbf{v}}\frac{\delta F}{\delta f}\right)d^3x\,d^3v\\ &+\int f\mathbf{B}\cdot\left(\frac{\partial}{\partial\mathbf{v}}\frac{\delta F}{\delta f}\times\frac{\partial}{\partial\mathbf{v}}\frac{\delta K}{\delta f}\right)d^3x\,d^3v.\end{aligned}\tag{1.6.20}$$

With the ***Maxwell–Vlasov Hamiltonian***

$$\begin{aligned}H(f,\mathbf{E},\mathbf{B})=&\frac{1}{2}\int\|\mathbf{v}\|^2f(\mathbf{x},\mathbf{v},t)\,d^3x\,d^3v\\ &+\frac{1}{2}\int(\|\mathbf{E}(x,t)\|^2+\|\mathbf{B}(\mathbf{x},t)\|^2)\,d^3x,\end{aligned}\tag{1.6.21}$$

[8] As shown in Cendra, Holm, Hoyle, and Marsden [1998], the correct physical description of the material representation of a plasma is a bit more complicated than simply $\mathrm{Diff}_{\mathrm{can}}$; however the end result is the same.

the Maxwell–Vlasov equations take the Hamiltonian form

$$\dot{F} = \{F, H\} \tag{1.6.22}$$

on the Poisson manifold $\mathcal{MV}$.

Exercises

◇ **1.6-1.** Verify that one obtains the Maxwell equations from the Maxwell–Poisson bracket.

◇ **1.6-2.** Verify that the action (1.6.7) has the momentum map $\mathbf{J}(\mathbf{A}, \mathbf{Y}) = \operatorname{div} \mathbf{E}$ in the sense given in §1.3.

1.7 Nonlinear Stability

There are various meanings that can be given to the word "stability." Intuitively, stability means that small disturbances do not grow large as time passes. Being more precise about this notion is not just capricious mathematical nitpicking; indeed, different interpretations of the word stability can lead to *different* stability criteria. Examples like the double spherical pendulum and stratified shear flows, which are sometimes used to model oceanographic phenomena show that one can get *different* criteria if one uses linearized or nonlinear analyses (see Marsden and Scheurle [1993a] and Abarbanel, Holm, Marsden, and Ratiu [1986]).

Some History. The history of stability theory in mechanics is very complex, but certainly has its roots in the work of Riemann [1860, 1861], Routh [1877], Thomson and Tait [1879], Poincaré [1885, 1892], and Liapunov [1892, 1897].

Since these early references, the literature has become too vast to even survey roughly. We do mention, however, that a guide to the large Soviet literature may be found in Mikhailov and Parton [1990].

The basis of the nonlinear stability method discussed below was originally given by Arnold [1965b, 1966b] and applied to two-dimensional ideal fluid flow, substantially augmenting the pioneering work of Rayleigh [1880]. Related methods were also found in the plasma physics literature, notably by Newcomb [1958], Fowler [1963], and Rosenbluth [1964]. However, these works did not provide a general setting or key convexity estimates needed to deal with the nonlinear nature of the problem. In retrospect, we may view other stability results, such as the stability of solitons in the Korteweg–de Vries (KdV) equation (Benjamin [1972] and Bona [1975]) as being instances of the same method used by Arnold. A crucial part of the method exploits the fact that the basic equations of nondissipative fluid and plasma dynamics are Hamiltonian in character. We shall explain below how the Hamilto-

nian structures discussed in the previous sections are used in the stability analysis.

Dynamics and Stability. Stability is a dynamical concept. To explain it, we shall use some fundamental notions from the theory of dynamical systems (see, for example, Hirsch and Smale [1974] and Guckenheimer and Holmes [1983]). The laws of dynamics are usually presented as equations of motion, which we write in the abstract form of a ***dynamical system***:

$$\dot{u} = X(u). \tag{1.7.1}$$

Here, u is a variable describing the state of the system under study, X is a system-specific function of u, and $\dot{u} = du/dt$, where t is time. The set of all allowed u's forms the state, or phase space P. We usually view X as a vector field on P. For a classical mechanical system, u is often a $2n$-tuple $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ of positions and momenta, and for fluids, u is a velocity field in physical space.

As time evolves, the state of the system changes; the state follows a curve $u(t)$ in P. The trajectory $u(t)$ is assumed to be uniquely determined if its initial condition $u_0 = u(0)$ is specified. An ***equilibrium state*** is a state u_e such that $X(u_e) = 0$. The unique trajectory starting at u_e is u_e itself; that is, u_e does not move in time.

The language of dynamics has been an extraordinarily useful tool in the physical and biological sciences, especially during the last few decades. The study of systems that develop spontaneous oscillations through a mechanism called the Poincaré–Andronov–Hopf bifurcation is an example of such a tool (see Marsden and McCracken [1976], Carr [1981], and Chow and Hale [1982], for example). More recently, the concept of "chaotic dynamics" has sparked a resurgence of interest in dynamical systems. This occurs when dynamical systems possess trajectories that are so complex that they behave as if they were, in some sense, random. Some believe that the theory of turbulence will use such notions in its future development. We are not concerned with chaos directly, although it plays a role in some of what follows. In particular, we remark that in the definition of stability below, stability does not preclude chaos. In other words, the trajectories near a stable point can still be temporally very complex; stability just prevents them from moving very far from equilibrium.

To define stability, we choose a measure of nearness in P using a "metric" d. For two points u_1 and u_2 in P, d determines a positive number denoted by $d(u_1, u_2)$, the ***distance*** from u_1 to u_2. In the course of a stability analysis, it is necessary to specify, or construct, a metric appropriate for the problem at hand. In this setting, one says that an equilibrium state u_e is ***stable*** when trajectories that start near u_e remain near u_e for all $t \geq 0$. In precise terms, given any number $\epsilon > 0$, there is $\delta > 0$ such that if $d(u_0, u_e) < \delta$, then $d(u(t), u_e) < \epsilon$ for all $t > 0$. Figure 1.7.1 shows examples of stable and unstable equilibria for dynamical systems whose state space is the plane.

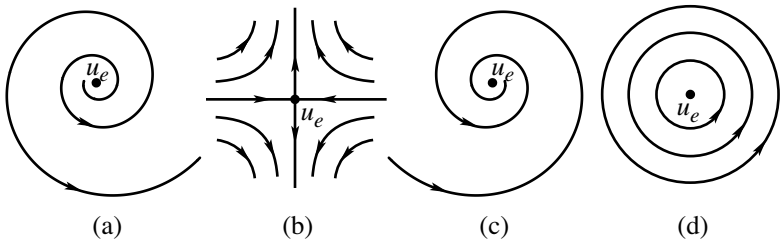

FIGURE 1.7.1. The equilibrium point (a) is unstable because the trajectory $u(t)$ does not remain near u_e. Similarly, (b) is unstable, since most trajectories (eventually) move away from u_e. The equilibria in (c) and (d) are stable because all trajectories near u_e stay near u_e.

Fluids can be stable relative to one distance measure and, simultaneously, unstable relative to another. This seeming pathology actually reflects important physical processes; see Wan and Pulvirente [1984].

Rigid-Body Stability. A physical example illustrating the definition of stability is the motion of a free rigid body. This system can be simulated by tossing a book, held shut with a rubber band, into the air. It rotates stably when spun about its longest and shortest axes, but unstably when spun about the middle axis (Figure 1.7.2). One possible choice of a distance measure defining stability in this example is a metric in body angular momentum space. We shall return to this example in detail in Chapter 15 when we study rigid-body stability.

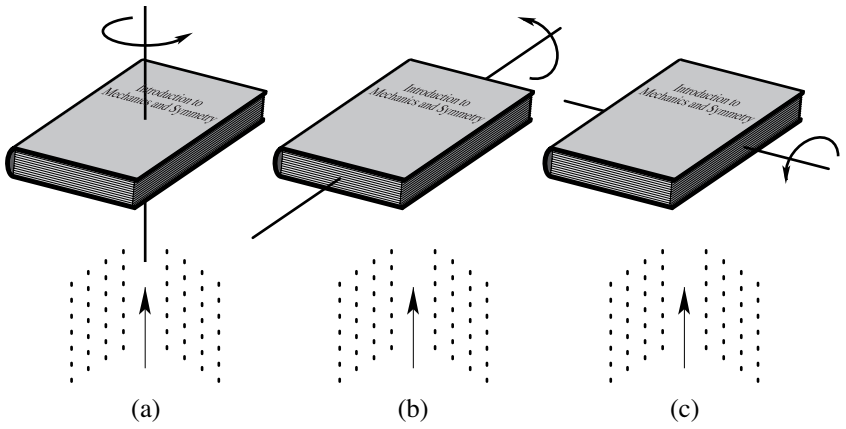

FIGURE 1.7.2. If you toss a book into the air, you can make it spin stably about its shortest axis (a), and its longest axis (b), but it is unstable when it rotates about its middle axis (c).

Linearized and Spectral Stability. There are two other ways of treating stability. First of all, one can linearize equation (1.7.1); if δu denotes a variation in u and $X'(u_e)$ denotes the linearization of X at u_e (the matrix of partial derivatives in the case of finitely many degrees of freedom), the linearized equations describe the time evolution of "infinitesimal" disturbances of u_e:

$$\frac{d}{dt}(\delta u) = X'(u_e) \cdot \delta u. \tag{1.7.2}$$

Equation (1.7.1), on the other hand, describes the nonlinear evolution of *finite* disturbances $\Delta u = u - u_e$. We say that u_e is ***linearly stable*** if (1.7.2) is stable at $\delta u = 0$, in the sense defined above. Intuitively, this means that there are no infinitesimal disturbances that are growing in time. If $(\delta u)_0$ is an eigenfunction of $X'(u_e)$, that is, if

$$X'(u_e) \cdot (\delta u)_0 = \lambda(\delta u)_0 \tag{1.7.3}$$

for a complex number λ, then the corresponding solution of (1.7.2) with initial condition $(\delta u)_0$ is

$$\delta u = e^{t\lambda}(\delta u)_0. \tag{1.7.4}$$

The right side of this equation is growing when λ has positive real part. This leads us to the third notion of stability: We say that (1.7.1) or (1.7.2) is ***spectrally stable*** if the eigenvalues (more precisely, points in the spectrum) all have nonpositive real parts. In finite dimensions and, under appropriate technical conditions in infinite dimensions, one has the following implications:

$$\text{(stability)} \Rightarrow \text{(spectral stability)}$$

and

$$\text{(linear stability)} \Rightarrow \text{(spectral stability)}.$$

If the eigenvalues all lie strictly in the left half-plane, then a classical result of Liapunov guarantees stability. (See, for instance, Hirsch and Smale [1974] for the finite-dimensional case and Marsden and McCracken [1976] or Abraham, Marsden, and Ratiu [1988] for the infinite-dimensional case.) However, in many systems of interest, the dissipation is very small and are modeled as being conservative. For such systems the eigenvalues must be symmetrically distributed under reflection in the real and imaginary axes (We prove this later in the text). This implies that the only possibility for spectral stability occurs when the eigenvalues lie exactly on the imaginary axis. Thus, *this version* of *the Liapunov theorem is of no help in the Hamiltonian case*.

Spectral stability need not imply stability; instabilities can be generated (even in Hamiltonian systems) through, for example, *resonance*. Thus, to

obtain general stability results, one must use other techniques to augment or replace the linearized theory. We give such a technique below.

Here is a planar example of a system that is spectrally stable at the origin but that is unstable there. In polar coordinates (r, θ), consider the evolution of $u = (r, \theta)$ given by

$$\dot{r} = r^3(1 - r^2) \quad \text{and} \quad \dot{\theta} = 1. \tag{1.7.5}$$

In (x, y) coordinates this system takes the form

$$\begin{aligned} \dot{x} &= x(x^2 + y^2)(1 - x^2 - y^2) - y, \\ \dot{y} &= y(x^2 + y^2)(1 - x^2 - y^2) + x. \end{aligned}$$

The eigenvalues of the linearized system at the origin are readily verified to be $\pm\sqrt{-1}$, so the origin is spectrally stable; however, the phase portrait, shown in Figure 1.7.3, shows that the origin is unstable. (We include the factor $1 - r^2$ to give the system an attractive periodic orbit—this is merely to enrich the example and show how a stable periodic orbit can attract the orbits expelled by an unstable equilibrium.) This is not, however, a conservative system; next, we give two examples of Hamiltonian systems with similar features.

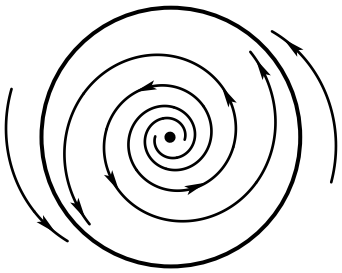

FIGURE 1.7.3. The phase portrait for $\dot{r} = r^3(1 - r^2)$, $\dot{\theta} = 1$.

Resonance Example. The linear system in $\mathbb{R}^2$ whose Hamiltonian is given by

$$H(q, p) = \frac{1}{2}p^2 + \frac{1}{2}q^2 + pq$$

has zero as a double eigenvalue, so it is spectrally stable. On the other hand,

$$q(t) = (q_0 + p_0)t + q_0 \quad \text{and} \quad p(t) = -(q_0 + p_0)t + p_0$$

is the solution of this system with initial condition (q_0, p_0), which clearly leaves any neighborhood of the origin no matter how close to it (q_0, p_0) is. Thus, *spectral stability need not imply even linear stability*. An even simpler

example of the same phenomenon is given by the free particle Hamiltonian $H(q,p) = p^2/2$.

Another higher-dimensional example with resonance in $\mathbb{R}^8$ is given by the linear system whose Hamiltonian is

$$H = q_2p_1 - q_1p_2 + q_4p_3 - q_3p_4 + q_2q_3.$$

The general solution with initial condition $(q_1^0, \dots, p_4^0)$ is given by

$$\begin{aligned} q_1(t) &= q_1^0 \cos t + q_2^0 \sin t, \\ q_2(t) &= -q_1^0 \sin t + q_2^0 \cos t, \\ q_3(t) &= q_3^0 \cos t + q_4^0 \sin t, \\ q_4(t) &= -q_3^0 \sin t + q_4^0 \cos t, \end{aligned}$$

and

$$\begin{aligned} p_1(t) &= -\frac{q_3^0}{2} t \sin t + \frac{q_4^0}{2}(t\cos t - \sin t) + p_1^0 \cos t + p_2^0 \sin t, \\ p_2(t) &= -\frac{q_3^0}{2}(t\cos t + \sin t) - \frac{q_4^0}{2} t \sin t - p_1^0 \sin t + p_2^0 \cos t, \\ p_3(t) &= \frac{q_1^0}{2} t \sin t - \frac{q_2^0}{2}(t\cos t + \sin t) + p_3^0 \cos t + p_4^0 \sin t, \\ p_4(t) &= \frac{q_1^0}{2}(t\cos t - \sin t) + \frac{q_2^0}{2} t \sin t - p_3^0 \sin t + p_4^0 \cos t. \end{aligned}$$

One sees that $p_i(t)$ leaves any neighborhood of the origin, no matter how close to the origin the initial conditions $(q_1^0, \dots, p_4^0)$ are; that is, the system is linearly unstable. On the other hand, all eigenvalues of this linear system are $\pm i$, each a quadruple eigenvalue. Thus, this linear system is spectrally stable.

Cherry's Example (Cherry [1959, 1968]). This example is a *Hamiltonian system that is spectrally stable* **and** *linearly stable but is nonlinearly unstable*. Consider the Hamiltonian on $\mathbb{R}^4$ given by

$$H = \frac{1}{2}(q_1^2 + p_1^2) - (q_2^2 + p_2^2) + \frac{1}{2} p_2(p_1^2 - q_1^2) - q_1q_2p_1. \tag{1.7.6}$$

This system has an equilibrium at the origin, which is linearly stable, since the linearized system consists of two uncoupled oscillators in the $(\delta q_2, \delta p_2)$ and $(\delta q_1, \delta p_1)$ variables, respectively, with frequencies in the ratio $2:1$ (the eigenvalues are $\pm i$ and $\pm 2i$, so the frequencies are in resonance). A family of solutions (parametrized by a constant τ) of Hamilton's equations for (1.7.6) is given by

$$\begin{aligned} q_1 &= -\sqrt{2}\frac{\cos(t-\tau)}{t-\tau}, \qquad & q_2 &= \frac{\cos 2(t-\tau)}{t-\tau}, \\ p_1 &= \sqrt{2}\frac{\sin(t-\tau)}{t-\tau}, \qquad & p_2 &= \frac{\sin 2(t-\tau)}{t-\tau}. \end{aligned} \tag{1.7.7}$$

The solutions (1.7.7) clearly blow up in finite time; however, they start at time $t = 0$ at a distance $\sqrt{3}/\tau$ from the origin, so by choosing τ large, we can find solutions starting arbitrarily close to the origin, yet going to infinity in a finite time, so *the origin is nonlinearly unstable*.

Despite the above situation relating the linear and nonlinear theories, there has been much effort devoted to the development of spectral stability methods. When *instabilities* are present, spectral estimates give important information on growth rates. As far as stability goes, spectral stability gives necessary, but not sufficient, conditions for stability. In other words, for the nonlinear problems *spectral instability can predict instability, but not stability*. This is a basic result of Liapunov; see Abraham, Marsden, and Ratiu [1988], for example. Our immediate purpose is the opposite: *to describe sufficient conditions for stability*.

Casimir Functions. Besides the energy, there are other conserved quantities associated with group symmetries such as linear and angular momentum. Some of these are associated with the group that underlies the passages from material to spatial or body coordinates. These are called ***Casimir functions***; such a quantity, denoted by C, is characterized by the fact that it Poisson commutes with every function, that is,

$$\{C, F\} = 0 \tag{1.7.8}$$

for all functions F on phase space P. We shall study such functions and their relation with momentum maps in Chapters 10 and 11. For example, if Φ is any function of one variable, the quantity

$$C(\mathbf{\Pi}) = \Phi(\|\mathbf{\Pi}\|^2) \tag{1.7.9}$$

is a Casimir function for the rigid-body bracket, as is seen by using the chain rule. Likewise,

$$C(\omega) = \int_\Omega \Phi(\omega)\, dx\, dy \tag{1.7.10}$$

is a Casimir function for the two-dimensional ideal fluid bracket. (This calculation ignores boundary terms that arise in an integration by parts—see Lewis, Marsden, Montgomery, and Ratiu [1986] for a treatment of these boundary terms.)

Casimir functions are conserved by the dynamics associated with any Hamiltonian H, since $\dot{C} = \{C, H\} = 0$. Conservation of (1.7.9) corresponds to conservation of total angular momentum for the rigid body, while conservation of (1.7.10) represents Kelvin's circulation theorem for the Euler equations. It provides infinitely many independent constants of the motion that mutually Poisson commute; that is, $\{C_1, C_2\} = 0$, but this does *not* imply that these equations are integrable.

Lagrange–Dirichlet Criterion. For Hamiltonian systems in canonical form, an equilibrium point (q_e, p_e) is a point at which the partial derivatives of H vanish, that is, it is a critical point of H. If the $2n \times 2n$ *matrix* $\delta^2 H$ *of second partial derivatives evaluated at* (q_e, p_e) *is positive or negative definite* (*that is, all the eigenvalues of* $\delta^2 H(q_e, p_e)$ *have the same sign*), *then* (q_e, p_e) *is stable.* This follows from conservation of energy and the fact from calculus that the level sets of H near (q_e, p_e) are approximately ellipsoids. As mentioned earlier, this condition implies, but is not implied by, spectral stability. The KAM (Kolmogorov, Arnold, Moser) theorem, which gives stability of periodic solutions for *two*-degree-of-freedom systems, and the Lagrange–Dirichlet theorem are the most basic *general* stability theorems for equilibria of Hamiltonian systems.

For example, let us apply the Lagrange–Dirichlet theorem to a classical mechanical system whose Hamiltonian has the form kinetic plus potential energy. If (q_e, p_e) is an equilibrium, it follows that p_e is zero. Moreover, the matrix $\delta^2 H$ of second-order partial derivatives of H evaluated at (q_e, p_e) block diagonalizes, with one of the blocks being the matrix of the quadratic form of the kinetic energy, which is always positive definite. Therefore, if $\delta^2 H$ is definite, it must be positive definite, and this in turn happens if and only if $\delta^2 V$ is positive definite at q_e, where V is the potential energy of the system. We conclude that *for a mechanical system* whose Lagrangian is kinetic minus potential energy, $(q_e, 0)$ *is a stable equilibrium, provided that the matrix* $\delta^2 V(q_e)$ *of second-order partial derivatives of the potential* V *at* q_e *is positive definite* (*or, more generally,* q_e *is a strict local minimum for* V). *If* $\delta^2 V$ *at* q_e *has a negative definite direction, then* q_e *is an unstable equilibrium.*

The second statement is seen in the following way. The linearized Hamiltonian system at $(q_e, 0)$ is again a Hamiltonian system whose Hamiltonian is of the form kinetic plus potential energy, the potential energy being given by the quadratic form $\delta^2 V(q_e)$. From a standard theorem in linear algebra, which states that two quadratic forms, one of which is positive definite, can be simultaneously diagonalized, we conclude that the linearized Hamiltonian system decouples into a family of Hamiltonian systems of the form

$$\frac{d}{dt}(\delta p_k) = -c_k \delta q^k, \qquad \frac{d}{dt}(\delta q^k) = \frac{1}{m_k}\delta p_k,$$

where $1/m_k > 0$ are the eigenvalues of the positive definite quadratic form given by the kinetic energy in the variables δp_j, and c_k are the eigenvalues of $\delta^2 V(q_e)$. Thus the eigenvalues of the linearized system are given by $\pm\sqrt{-c_k/m_k}$. Therefore, if some c_k is negative, the linearized system has at least one positive eigenvalue, and thus $(q_e, 0)$ is spectrally and hence linearly and nonlinearly unstable. For generalizations of this, see Oh [1987], Grillakis, Shatah, and Strauss [1987], Chern [1997] and references therein.

The Energy–Casimir Method. This is a generalization of the classical Lagrange–Dirichlet method. Given an equilibrium u_e for $\dot{u} = X_H(u)$ on a Poisson manifold P, it proceeds in the following steps.

To test an equilibrium (satisfying $X_H(z_e) = 0$) for stability:

Step 1. *Find a conserved function C (C will typically be a Casimir function plus other conserved quantities) such that the first variation vanishes*:

$$\delta(H + C)(z_e) = 0.$$

Step 2. *Calculate the second variation*

$$\delta^2(H + C)(z_e).$$

Step 3. *If $\delta^2(H + C)(z_e)$ is definite (either positive or negative), then z_e is called* ***formally stable.***

With regard to Step 3, we point out that an equilibrium solution need not be a critical point of H alone; in general, $\delta H(z_e) \neq 0$. An example where this occurs is a rigid body spinning about one of its principal axes of inertia. In this case, a critical point of H alone would have zero angular velocity; but a critical point of $H + C$ is a (nontrivial) stationary rotation about one of the principal axes.

The argument used to establish the Lagrange–Dirichlet test formally works in infinite dimensions too. Unfortunately, for systems with infinitely many degrees of freedom (like fluids and plasmas), there is a serious technical snag. The calculus argument used before runs into problems; one might think that these are just technical and that we just need to be more careful with the calculus arguments. In fact, there is widespread belief in this "energy criterion" (see, for instance, the discussion and references in Marsden and Hughes [1983, Chapter 6], and Potier-Ferry [1982]). However, Ball and Marsden [1984] have shown using an example from elasticity theory that the difficulty is genuine: They produce a critical point of H at which $\delta^2 H$ is positive definite, yet this point is *not* a local minimum of H. On the other hand, Potier-Ferry [1982] shows that asymptotic stability is restored if suitable dissipation is added. Another way to overcome this difficulty is to modify Step 3 using a convexity argument of Arnold [1966b].

Modified Step 3. Assume that P is a *linear* space.

(a) *Let $\Delta u = u - u_e$ denote a finite variation in phase space.*

(b) *Find quadratic functions Q_1 and Q_2 such that*

$$Q_1(\Delta u) \leq H(u_e + \Delta u) - H(u_e) - \delta H(u_e) \cdot \Delta u$$

and

$$Q_2(\Delta u) \leq C(u_e + \Delta u) - C(u_e) - \delta C(u_e) \cdot \Delta u,$$

(c) *Require* $Q_1(\Delta u) + Q_2(\Delta u) > 0$ *for all* $\Delta u \neq 0$.

(d) *Introduce the norm* $\|\Delta u\|$ *by*

$$\|\Delta u\|^2 = Q_1(\Delta u) + Q_2(\Delta u),$$

so $\|\Delta u\|$ *is a measure of the distance from* u *to* u_e*; that is, we choose* $d(u, u_e) = \|\Delta u\|$.

(e) *Require*

$$|H(u_e + \Delta u) - H(u_e)| \leq C_1 \|\Delta u\|^\alpha$$

and

$$|C(u_e + \Delta u) - C(u_e)| \leq C_2 \|\Delta u\|^\alpha$$

for constants $\alpha, C_1, C_2 > 0$ *and* $\|\Delta u\|$ *sufficiently small.*

These conditions guarantee stability of u_e and provide the distance measure relative to which stability is defined. The key part of the proof is simply the observation that if we add the two inequalities in (b), we get

$$\|\Delta u\|^2 \leq H(u_e + \Delta u) + C(u_e + \Delta u) - H(u_e) - C(u_e)$$

using the fact that $\delta H(u_e) \cdot \Delta u$ and $\delta C(u_e) \cdot \Delta u$ add up to zero by Step 1. But H and C are constant in time, so

$$\|(\Delta u)_{\text{time}=t}\|^2 \leq \left[H(u_e + \Delta u) + C(u_e + \Delta u) - H(u_e) - C(u_e)\right]\big|_{\text{time}=0}.$$

Now employ the inequalities in (e) to get

$$\|(\Delta u)_{\text{time}=t}\|^2 \leq (C_1 + C_2)\|(\Delta u)_{\text{time}=0}\|^\alpha.$$

This estimate bounds the temporal growth of finite perturbations in terms of initial perturbations, which is what is needed for stability. For a survey of this method, additional references, and numerous examples, see Holm, Marsden, Ratiu, and Weinstein [1985].

There are some situations (such as the stability of elastic rods) in which the above techniques do not apply. The chief reason is that there may be a lack of sufficiently many Casimir functions to achieve even the first step. For this reason a modified (but more sophisticated) method has been developed called the "energy–momentum method." The key to the method is to avoid the use of Casimir functions by applying the method *before* any reduction has taken place. This method was developed in a series of papers of Simo, Posbergh, and Marsden [1990, 1991] and Simo, Lewis, and Marsden [1991]. A discussion and additional references are found later in this section.

Gyroscopic Systems. The distinctions between "stability by energy methods," that is, ***energetics*** and "spectral stability" become especially interesting when one adds dissipation. In fact, building on the classical work of Kelvin and Chetaev, one can prove that if $\delta^2 H$ is indefinite, yet the spectrum is on the imaginary axis, then adding dissipation necessarily makes the system *linearly unstable*. That is, at least one pair of eigenvalues of the linearized equations move into the right half-plane. This is a phenomenon called ***dissipation-induced instability.*** This result, along with related developments, is proved in Bloch, Krishnaprasad, Marsden, and Ratiu [1991, 1994, 1996]. For example, consider the linear ***gyroscopic system***

$$M\ddot{\mathbf{q}} + S\dot{\mathbf{q}} + V\mathbf{q} = 0, \tag{1.7.11}$$

where $\mathbf{q} \in \mathbb{R}^n$, M is a positive definite symmetric $n \times n$ matrix, S is skew, and V is symmetric. This system is Hamiltonian (Exercise 1.7-2). If V has negative eigenvalues, then (1.7.11) is *formally unstable*. However, due to S, the system can be spectrally stable. However, if R is positive definite symmetric and $\epsilon > 0$ is small, the system with friction

$$M\ddot{\mathbf{q}} + S\dot{\mathbf{q}} + \epsilon R\dot{\mathbf{q}} + V\mathbf{q} = 0 \tag{1.7.12}$$

is linearly unstable. A specific example is given in Exercise 1.7-4.

Outline of the Energy–Momentum Method. The energy momentum method is an extension of the Arnold (or energy–Casimir) method for the study of stability of relative equilibria, which was developed for Lie–Poisson systems on duals of Lie algebras, especially those of fluid dynamical type. In addition, the method extends and refines the fundamental stability techniques going back to Routh, Liapunov, and, in more recent times, to the work of Smale.

The motivation for these extensions is threefold.

First of all, the energy–momentum method can deal with Lie–Poisson systems for which there are not sufficient Casimir functions available, such as 3-D ideal flow and certain problems in elasticity. In fact, Abarbanel and Holm [1987] use what can be recognized retrospectively as the energy–momentum method to show that 3-D equilibria for ideal flow are generally formally unstable due to vortex stretching. Other fluid and plasma situations, such as those considered by Chern and Marsden [1990] for ABC flows and certain multiple-hump situations in plasma dynamics (see Holm, Marsden, Ratiu, and Weinstein [1985] and Morrison [1987], for example), provided additional motivation in the Lie–Poisson setting.

A second motivation is to extend the method to systems that need not be Lie–Poisson and still make use of the powerful idea of using reduced spaces, as in the original Arnold method. Examples such as rigid bodies with vibrating antennas (Sreenath, Oh, Krishnaprasad, and Marsden [1988], Oh,

Sreenath, Krishnaprasad, and Marsden [1989], Krishnaprasad and Marsden [1987]) and coupled rigid bodies (Patrick [1989]) motivated the need for such an extension of the theory.

Finally, it gives sharper stability conclusions in material representation and links with geometric phases.

The Idea of the Energy–Momentum Method. The setting of the energy–momentum method is that of a mechanical system with symmetry with a configuration space Q and phase space T^*Q and a symmetry group G acting, with a standard momentum map $\mathbf{J} : T^*Q \to \mathfrak{g}^*$, where $\mathfrak{g}^*$ is the Lie algebra of G. Of course, one gets the Lie–Poisson case when $Q = G$.

The rough idea for the energy momentum method is first to formulate the problem directly on the unreduced space. Here, relative equilibria associated with a Lie algebra element ξ are critical points of the augmented Hamiltonian $H_\xi := H - \langle \mathbf{J}, \xi \rangle$. The idea is now to compute the second variation of H_ξ at a relative equilibrium z_e with momentum value μ_e subject to the constraint $J = \mu_e$ and on a space transverse to the action of G_{μ_e}, the subgroup of G that leaves μ_e fixed. Although the augmented Hamiltonian plays the role of $H+C$ in the Arnold method, notice that Casimir functions are not required to carry out the calculations.

The surprising thing is that the second variation of H_ξ at the relative equilibrium can be arranged to be block diagonal, using splittings that are based on the mechanical connection, while *at the same time*, the symplectic structure also has a simple block structure, so that the linearized equations are put into a useful canonical form. Even in the Lie–Poisson setting, this leads to situations in which one gets much simpler second variations. This block diagonal structure is what gives the method its computational power.

The general theory for carrying out this procedure was developed in Simo, Posbergh, and Marsden [1990, 1991] and Simo, Lewis, and Marsden [1991]. An exposition of the method may be found, along with additional references, in Marsden [1992]. It is of interest to extend this to the singular case, which is the subject of ongoing work; see Ortega and Ratiu [1997, 1998] and references therein.

The energy–momentum method may also be usefully formulated in the Lagrangian setting, which is very convenient for the calculations in many examples. The general theory for this was developed in Lewis [1992] and Wang and Krishnaprasad [1992]. This Lagrangian setting is closely related to the general theory of Lagrangian reduction. In this context one reduces variational principles rather than symplectic and Poisson structures, and for the case of reducing the tangent bundle of a Lie group, this leads to the Euler–Poincaré equations rather than the Lie–Poisson equations.

Effectiveness in Examples. The energy–momentum method has proven its effectiveness in a number of examples. For instance, Lewis and Simo [1990] were able to deal with the stability problem for pseudo-rigid bodies, which was thought up to that time to be analytically intractable.

The energy–momentum method can sometimes be used in contexts where the reduced space is singular or at nongeneric points in the dual of the Lie algebra. This is done at singular points in Lewis, Ratiu, Simo, and Marsden [1992], who analyze the heavy top in great detail and, in the Lie–Poisson setting for compact groups at nongeneric points in the dual of the Lie algebra, in Patrick [1992, 1995]. One of the key things is to keep track of group drifts, because the isotropy group G_μ can change for nearby points, and these are important for the reconstruction process and for understanding the Hannay–Berry phase in the context of reduction (see Marsden, Montgomery, and Ratiu [1990] and references therein). For noncompact groups and an application to the dynamics of rigid bodies in fluids (underwater vehicles), see Leonard and Marsden [1997]. Additional work in this area is still needed in the context of singular reduction.

The Benjamin–Bona theorem on stability of solitons for the KdV equation can be viewed as an instance of the energy momentum method, see also Maddocks and Sachs [1993], and for example, Oh [1987] and Grillakis, Shatah, and Strauss [1987], although there are many subtleties in the PDE context.

Hamiltonian Bifurcations. The energy–momentum method has also been used in the context of Hamiltonian bifurcation problems. We shall give some simple examples of this in §1.8. One such context is that of free boundary problems building on the work of Lewis, Marsden, Montgomery, and Ratiu [1986], which gives a Hamiltonian structure for dynamic free boundary problems (surface waves, liquid drops, etc.), generalizing Hamiltonian structures found by Zakharov. Along with the Arnold method itself, this is used for a study of the bifurcations of such problems in Lewis, Marsden, and Ratiu [1987], Lewis [1989, 1992], Kruse, Marsden, and Scheurle [1993], and other references cited therein.

Converse to the Energy–Momentum Method. Because of the block structure mentioned, it has also been possible to prove, in a sense, a converse of the energy–momentum method. That is, if the second variation is indefinite, then the system is unstable. One cannot, of course, hope to do this literally as stated, since there are many systems (e.g., gyroscopic system mentioned earlier—an explicit example is given in Exercise 1.7-4) that are formally unstable, and yet their linearizations have eigenvalues lying on the imaginary axis. Most of these are presumably unstable due to "Arnold diffusion," but of course this is a very delicate situation to prove analytically. Instead, the technique is to show that with the addition of dissipation, the system is destabilized. This idea of *dissipation-induced instability* goes back to Thomson and Tait in the last century. In the context of the energy–momentum method, Bloch, Krishnaprasad, Marsden, and Ratiu [1994, 1996] show that with the addition of appropriate dissipation, the indefiniteness of the second variation is sufficient to induce linear instability in the problem.

There are related eigenvalue movement formulas (going back to Krein) that are used to study non-Hamiltonian perturbations of Hamiltonian normal forms in Kirk, Marsden, and Silber [1996]. There are interesting analogues of this for reversible systems in O'Reilly, Malhotra, and Namamchchivaya [1996].

Extension to Nonholonomic Systems. It is possible to partially extend the energy–momentum method to the case of nonholonomic systems. Building on the work on nonholonomic systems in Arnold [1988], Bates and Sniatycki [1993] and Bloch, Krishnaprasad, Marsden, and Murray [1996], on the example of the Routh problem in Zenkov [1995], and on the large Russian literature in this area, Zenkov, Bloch, and Marsden [1998] show that there is a generalization to this setting. The method is effective in the sense that it applies to a wide variety of interesting examples, such as the rolling disk, a three-wheeled vehicle known as the the roller racer and the rattleback.

Exercises

◇ **1.7-1.** Work out Cherry's example of the Hamiltonian system in $\mathbb{R}^4$ whose energy function is given by (1.7.6). Show explicitly that the origin is a linearly and spectrally stable equilibrium but that it is nonlinearly unstable by proving that (1.7.7) is a solution for every $\tau > 0$ that can be chosen to start arbitrarily close to the origin and that goes to infinity for $t \to \tau$.

◇ **1.7-2.** Show that (1.7.11) is Hamiltonian with $\mathbf{p} = M\dot{\mathbf{q}}$,

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\mathbf{p} \cdot M^{-1}\mathbf{p} + \frac{1}{2}\mathbf{q} \cdot V\mathbf{q},$$

and

$$\{F, K\} = \frac{\partial F}{\partial q^i}\frac{\partial K}{\partial p_i} - \frac{\partial K}{\partial q^i}\frac{\partial F}{\partial p_i} - S^{ij}\frac{\partial F}{\partial p_i}\frac{\partial K}{\partial p_j}.$$

◇ **1.7-3.** Show that (up to an overall factor) the characteristic polynomial for the linear system (1.7.11) is

$$p(\lambda) = \det[\lambda^2 M + \lambda S + V]$$

and that this actually is a polynomial of degree n in λ^2.

◇ **1.7-4.** Consider the two-degree-of-freedom system

$$\ddot{x} - g\dot{y} + \gamma\dot{x} + \alpha x = 0,$$
$$\ddot{y} + g\dot{x} + \delta\dot{y} + \beta y = 0.$$

(a) Write it in the form (1.7.12).

(b) For $\gamma = \delta = 0$ show:

(i) it is spectrally stable if $\alpha > 0$, $\beta > 0$;

(ii) for $\alpha\beta < 0$, it is spectrally unstable;

(iii) for $\alpha < 0$, $\beta < 0$, it is formally unstable (that is, the energy function, which is a quadratic form, is indefinite); and

A. if $D := (g^2 + \alpha + \beta)^2 - 4\alpha\beta < 0$, then there are two roots in the right half-plane and two in the left; the system is spectrally unstable;

B. if $D = 0$ and $g^2 + \alpha + \beta \geq 0$, the system is spectrally stable, but if $g^2 + \alpha + \beta < 0$ then it is spectrally unstable; and

C. if $D > 0$ and $g^2 + \alpha + \beta \geq 0$, the system is spectrally stable, but if $g^2 + \alpha + \beta < 0$, then it is spectrally unstable.

(c) For a polynomial $p(\lambda) = \lambda^4 + \rho_1\lambda^3 + \rho_2\lambda^2 + \rho_3\lambda + \rho_4$, the *Routh–Hurwitz criterion* (see Gantmacher [1959, Volume 2]) says that the number of right half-plane zeros of p is the number of sign changes of the sequence

$$\left\{1, \rho_1, \frac{\rho_1\rho_2 - \rho_3}{\rho_1}, \frac{\rho_3\rho_1\rho_2 - \rho_3^2 - \rho_4\rho_1^2}{\rho_1\rho_2 - \rho_3}, \rho_4\right\}.$$

Apply this to the case in which $\alpha < 0$, $\beta < 0$, $g^2 + \alpha + \beta > 0$, $\gamma > 0$, and $\delta > 0$ to show that the system is spectrally unstable.

1.8 Bifurcation

When the energy–momentum or energy–Casimir method indicates that an instability might be possible, techniques of bifurcation theory can be brought to bear to determine the emerging dynamical complexities such as the development of multiple equilibria and periodic orbits.

Ball in a Rotating Hoop. For example, consider a particle moving with no friction in a rotating hoop (Figure 1.8.1).

In §2.8 we derive the equations and study the phase portraits for this system. One finds that as ω increases past $\sqrt{g/R}$, the stable equilibrium at $\theta = 0$ becomes unstable through a *Hamiltonian pitchfork bifurcation* and two new solutions are created. These solutions are symmetric in the vertical axis, a reflection of the original $\mathbb{Z}_2$ symmetry of the mechanical system in Figure 1.8.1. Breaking this symmetry by, for example, putting the rotation axis slightly off center is an interesting topic that we shall discuss in §2.8.

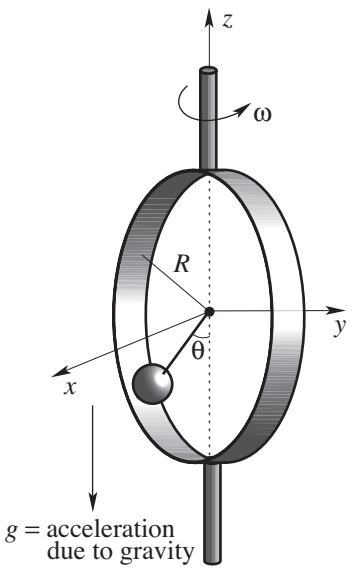

FIGURE 1.8.1. A particle moving in a hoop rotating with angular velocity ω.

Rotating Liquid Drop. The system consists of the two-dimensional Euler equations for an ideal fluid with a free boundary. An equilibrium solution consists of a rigidly rotating circular drop. The energy–Casimir method shows stability, provided that

$$\Omega < 2\sqrt{\frac{3\tau}{R^3}}. \tag{1.8.1}$$

In this formula, Ω is the angular velocity of the circular drop, R is its radius, and τ is the surface tension, a constant. As Ω increases and (1.8.1) is violated, the stability of the circular solution is lost and is picked up by elliptical-like solutions with $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry. The bifurcation is actually subcritical relative to the *angular velocity* Ω (that is, the new solutions occur *below* the critical value of Ω) and is supercritical (the new solutions occur *above* criticality) relative to the *angular momentum*. This is proved in Lewis, Marsden, and Ratiu [1987] and Lewis [1989], where other references may also be found (see Figure 1.8.2).

For the ball in the hoop, the eigenvalue evolution for the linearized equations is shown in Figure 1.8.3(a). For the rotating liquid drop, the movement of eigenvalues is the same: They are constrained to *stay* on the imaginary axis because of the symmetry of the problem. Without this symmetry, eigenvalues typically split, as in Figure 1.8.3(b). These are examples of a general theory of the movement of such eigenvalues given in Golubitsky and Stewart [1987], Dellnitz, Melbourne, and Marsden [1992], Knobloch, Mahalov, and Marsden [1994], and Kirk, Marsden, and Silber [1996].

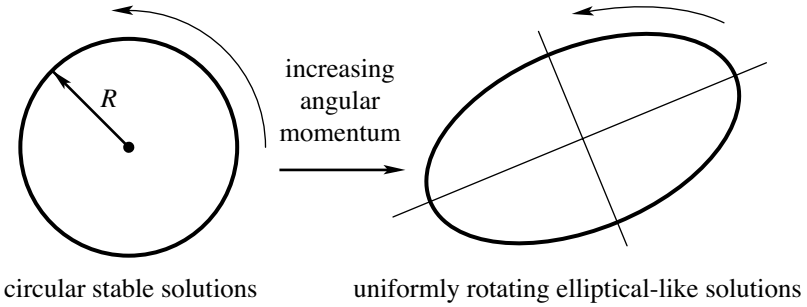

FIGURE 1.8.2. A circular liquid drop losing its stability and its symmetry.

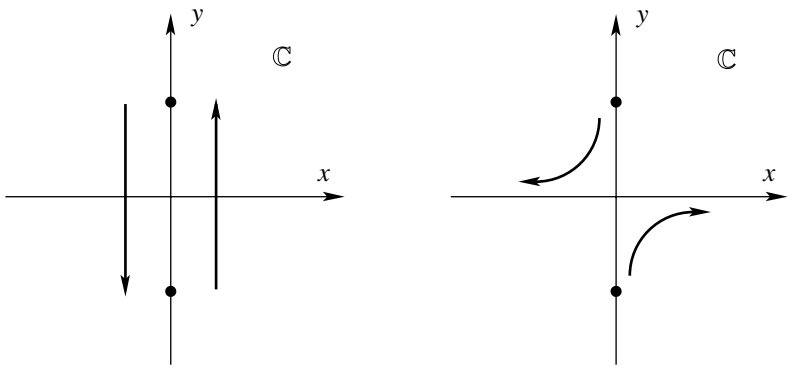

FIGURE 1.8.3. The movement of eigenvalues in bifurcation of equilibria.

More Examples. Another example is the heavy top: a rigid body with one point fixed, moving in a gravitational field. When the top makes the transition from a fast top to a slow top, the angular velocity ω decreases past the critical value

$$\omega_c = \frac{2\sqrt{MglI_1}}{I_3}, \tag{1.8.2}$$

stability is lost, and a ***resonance bifurcation*** occurs. Here, when the bifurcation occurs, the eigenvalues of the equations linearized at the equilibrium behave as in Figure 1.8.4.

For an extensive study of bifurcations and stability in the dynamics of a heavy top, see Lewis, Ratiu, Simo, and Marsden [1992]. Behavior of this sort is sometimes called a ***Hamiltonian Krein–Hopf bifurcation***, or a ***gyroscopic instability*** (see van der Meer [1985, 1990]). Here more complex dynamic behavior ensues, including periodic and chaotic motions (see

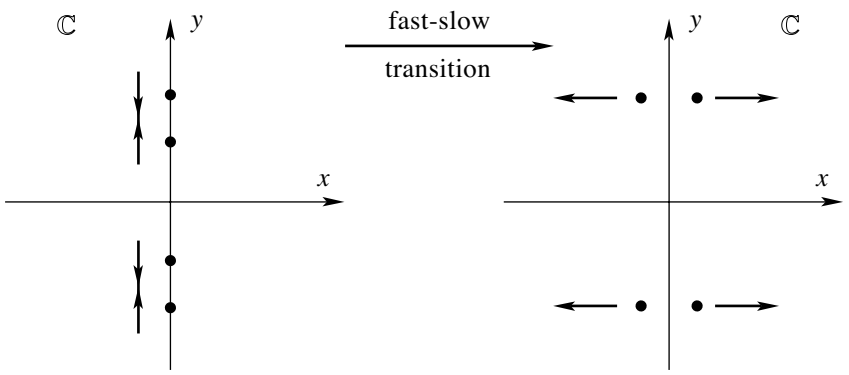

FIGURE 1.8.4. Eigenvalue movement in the Hamiltonian Hopf bifurcation.

Holmes and Marsden [1983]). In some systems with symmetry, the eigenvalues can ***pass*** as well as ***split***, as has been shown by Dellnitz, Melbourne, and Marsden [1992] and references therein.

More sophisticated examples, such as the dynamics of two coupled three-dimensional rigid bodies, requires a systematic development of the basic theory of Golubitsky and Schaeffer [1985] and Golubitsky, Stewart, and Schaeffer [1988]. This theory is begun in, for example, Duistermaat [1983], Lewis, Marsden, and Ratiu [1987], Lewis [1989], Patrick [1989], Meyer and Hall [1992], Broer, Chow, Kim, and Vegter [1993], and Golubitsky, Marsden, Stewart, and Dellnitz [1995]. For bifurcations in the double spherical pendulum (which includes a Hamiltonian–Krein–Hopf bifurcation), see Dellnitz, Marsden, Melbourne, and Scheurle [1992] and Marsden and Scheurle [1993a].

Exercises

◇ **1.8-1.** Study the bifurcations (changes in the phase portrait) for the equation

$$\ddot{x} + \mu x + x^2 = 0$$

as μ passes through zero. Use the second derivative test on the potential energy.

◇ **1.8-2.** Repeat Exercise 1.8-1 for

$$\ddot{x} + \mu x + x^3 = 0$$

as μ passes through zero.

1.9 The Poincaré–Melnikov Method

The Forced Pendulum. To begin with a simple example, consider the equation of a forced pendulum:

$$\ddot{\phi} + \sin\phi = \epsilon\cos\omega t. \tag{1.9.1}$$

Here ω is a constant angular forcing frequency and ϵ is a small parameter. Systems of this or a similar nature arise in many interesting situations. For example, a double planar pendulum and other "executive toys" exhibit chaotic motion that is analogous to the behavior of this equation; see Burov [1986] and Shinbrot, Grebogi, Wisdom, and Yorke [1992].

For $\epsilon = 0$ (1.9.1) has the phase portrait of a simple pendulum (the same as shown later in Figure 2.8.2a). For ϵ small but nonzero, (1.9.1) possesses no analytic integrals of the motion. In fact, it possesses transversal intersecting stable and unstable manifolds (separatrices); that is, the Poincaré map $P_{t_0} : \mathbb{R}^2 \to \mathbb{R}^2$ defined as the map that advance solutions by one period $T = 2\pi/\omega$ starting at time t_0 possess transversal homoclinic points. This type of dynamic behavior has several consequences, besides precluding the existence of analytic integrals, that lead one to use the term "chaotic." For example, (1.9.1) has infinitely many periodic solutions of arbitrarily high period. Also, using the shadowing lemma, one sees that given any bi–infinite sequence of zeros and ones[9], there exists a corresponding solution of (1.9.1) that successively crosses the plane $\phi = 0$ (the pendulum's vertically downward configuration) with $\phi > 0$ corresponding to a zero and $\phi < 0$ corresponding to a one. The origin of this chaos on an intuitive level lies in the motion of the pendulum near its unperturbed homoclinic orbit, the orbit that does one revolution in infinite time. Near the top of its motion (where $\phi = \pm\pi$) small nudges from the forcing term can cause the pendulum to fall to the left or right in a temporally complex way.

The dynamical systems theory needed to justify the preceding statements is available in Smale [1967], Moser [1973], Guckenheimer and Holmes [1983], and Wiggins [1988, 1990]. Some key people responsible for the development of the basic theory are Poincaré, Birkhoff, Kolmogorov, Melnikov, Arnold, Smale, and Moser. The idea of transversal intersecting separatrices comes from Poincaré's famous paper on the three-body problem (Poincaré [1890]). His goal, not quite achieved for reasons we shall comment on later, was to prove the nonintegrability of the restricted three-body problem and that various series expansions used up to that point diverged (he began the theory of asymptotic expansions and dynamical systems in the course of

[9] For example, build such a sequence out of digits from the binary expansion of π and e using the former for the left infinite sequence and the latter for the right infinite sequence.

this work). See Diacu and Holmes [1996] for additional information about Poincaré's work.

Although Poincaré had all the essential tools needed to prove that equations like (1.9.1) are not integrable (in the sense of having no analytic integrals), his interests lay with harder problems, and he did not develop the easier basic theory very much. Important contributions were made by Melnikov [1963] and Arnold [1964] that lead to a simple procedure for proving that (1.9.1) is not integrable. The Poincaré–Melnikov method was revived by Chirikov [1979], Holmes [1980b], and Chow, Hale, and Mallet-Paret [1980]. We shall give the method for Hamiltonian systems. We refer to Guckenheimer and Holmes [1983] and to Wiggins [1988, 1990] for generalizations and further references.

The Poincaré–Melnikov Method. This method proceeds as follows:

1. Write the dynamical equation to be studied in the form

$$\dot{x} = X_0(x) + \epsilon X_1(x,t), \tag{1.9.2}$$

where $x \in \mathbb{R}^2$, X_0 is a Hamiltonian vector field with energy H_0, X_1 is periodic with period T and is Hamiltonian with energy a T–periodic function H_1. Assume that X_0 has a homoclinic orbit $\overline{x}(t)$, so $\overline{x}(t) \to x_0$, a hyperbolic saddle point, as $t \to \pm\infty$.

2. Compute the ***Poincaré–Melnikov function*** defined by

$$M(t_0) = \int_{-\infty}^{\infty} \{H_0, H_1\}(\overline{x}(t-t_0), t)\, dt, \tag{1.9.3}$$

where $\{\,,\}$ denotes the Poisson bracket.

If $M(t_0)$ has simple zeros as a function of t_0, then (1.9.2) has, for sufficiently small ϵ, homoclinic chaos in the sense of transversal intersecting separatrices (in the sense of Poincaré maps as mentioned above).

We shall prove this result in §2.11. To apply it to equation (1.9.1) one proceeds as follows. Let $x = (\phi, \dot{\phi})$, so we get

$$\frac{d}{dt}\begin{bmatrix} \phi \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} \dot{\phi} \\ -\sin\phi \end{bmatrix} + \epsilon \begin{bmatrix} 0 \\ \cos\omega t \end{bmatrix}.$$

The homoclinic orbits for $\epsilon = 0$ are given by (see Exercise 1.9-1)

$$\overline{x}(t) = \begin{bmatrix} \phi(t) \\ \dot{\phi}(t) \end{bmatrix} = \begin{bmatrix} \pm 2\tan^{-1}(\sinh t) \\ \pm 2 \text{ sech } t \end{bmatrix},$$

and one has

$$H_0(\phi,\dot\phi)=\tfrac{1}{2}\dot\phi^2-\cos\phi \quad \text{and} \quad H_1(\phi,\dot\phi,t)=\phi\cos\omega t. \tag{1.9.4}$$

Hence (1.9.3) gives

$$\begin{aligned} M(t_0) &= \int_{-\infty}^{\infty}\left(\frac{\partial H_0}{\partial\phi}\frac{\partial H_1}{\partial\dot\phi}-\frac{\partial H_0}{\partial\dot\phi}\frac{\partial H_1}{\partial\phi}\right)(\overline{x}(t-t_0),t)\,dt \\ &= -\int_{-\infty}^{\infty}\dot\phi(t-t_0)\cos\omega t\,dt \\ &= \mp\int_{-\infty}^{\infty}[2\ \mathrm{sech}(t-t_0)\cos\omega t]\,dt. \end{aligned}$$

Changing variables and using the fact that sech is even and sin is odd, we get

$$M(t_0)=\mp 2\left(\int_{-\infty}^{\infty}\mathrm{sech}\ t\cos\omega t\,dt\right)\cos(\omega t_0).$$

The integral is evaluated by residues (see Exercise 1.9-2):

$$M(t_0)=\mp 2\pi\ \mathrm{sech}\left(\frac{\pi\omega}{2}\right)\cos(\omega t_0), \tag{1.9.5}$$

which clearly has simple zeros. Thus, this equation has chaos for ϵ small enough.

Exercises

◇ **1.9-1.** Verify directly that the homoclinic orbits for the simple pendulum equation $\ddot\phi+\sin\phi=0$ are given by $\phi(t)=\pm 2\tan^{-1}(\sinh t)$.

◇ **1.9-2.** Evaluate the integral $\int_{-\infty}^{\infty}\mathrm{sech}\ t\cos\omega t\,dt$ to prove (1.9.5) as follows. Write $\mathrm{sech}\ t=2/(e^t+e^{-t})$ and note that there is a simple pole of

$$f(z)=\frac{e^{i\omega z}+e^{-i\omega z}}{e^z+e^{-z}}$$

in the complex plane at $z=\pi i/2$. Evaluate the residue there and apply Cauchy's theorem.[10]

[10] Consult a book on complex variables such as Marsden and Hoffman, *Basic Complex Analysis*, Third Edition, Freeman, 1998.

1.10 Resonances, Geometric Phases, and Control

The work of Smale [1970] shows that topology plays an important role in mechanics. Smale's work employs Morse theory applied to conserved quantities such as the energy–momentum map. In this section we point out other ways in which geometry and topology enter mechanical problems.

The One-to-One Resonance. When one considers resonant systems, one often encounters Hamiltonians of the form

$$H = \frac{1}{2}(q_1^2 + p_1^2) + \frac{\lambda}{2}(q_2^2 + p_2^2) + \text{ higher-order terms.} \tag{1.10.1}$$

The quadratic terms describe two oscillators that have the same frequency when $\lambda = 1$, which is why one speaks of a one-to-one resonance. To analyze the dynamics of H, it is important to utilize a good geometric picture for the critical case

$$H_0 = \frac{1}{2}(q_1^2 + p_1^2 + q_2^2 + p_2^2). \tag{1.10.2}$$

The energy level $H_0 =$ constant is the three-sphere $S^3 \subset \mathbb{R}^4$. If we think of H_0 as a function on complex two-space $\mathbb{C}^2$ by letting

$$z_1 = q_1 + ip_1 \quad \text{and} \quad z_2 = q_2 + ip_2,$$

then $H_0 = (|z_1|^2 + |z_2|^2)/2$, so H_0 is left-invariant by the action of SU(2), the group of complex 2×2 unitary matrices of determinant one. The corresponding conserved quantities are

$$\begin{aligned} W_1 &= 2(q_1q_2 + p_1p_2), \\ W_2 &= 2(q_2p_1 - q_1p_2), \\ W_3 &= q_1^2 + p_1^2 - q_2^2 - p_2^2, \end{aligned} \tag{1.10.3}$$

which comprise the components of a (momentum) map

$$\mathbf{J} : \mathbb{R}^4 \to \mathbb{R}^3. \tag{1.10.4}$$

From the readily verified relation $4H_0^2 = W_1^2 + W_2^2 + W_3^2$, one finds that $\mathbf{J}$ restricted to S^3 gives a map

$$j : S^3 \to S^2. \tag{1.10.5}$$

The fibers j^{-1}(point) are circles, and the trajectories for the dynamics of H_0 move along these circles. The map j is the ***Hopf fibration***, which describes S^3 as a topologically nontrivial circle bundle over S^2. The role of the Hopf fibration in mechanics was known to Reeb [1949].

One also finds that the study of systems like (1.10.1) that are close to H_0 can, to a good approximation, be reduced to dynamics on S^2. These dynamics are in fact Lie–Poisson and S^2 sits as a coadjoint orbit in $\mathfrak{so}(3)^*$, so the evolution is of rigid-body type, just with a different Hamiltonian. For a computer study of the Hopf fibration in the one-to-one resonance, see Kocak, Bisshopp, Banchoff, and Laidlaw [1986].

The Hopf Fibration in Rigid-Body Mechanics. When doing reduction for the rigid body, one studies the reduced space

$$\mathbf{J}^{-1}(\mu)/G_\mu = \mathbf{J}^{-1}(\mu)/S^1,$$

which in this case is the sphere S^2. As we shall see in Chapter 15, $\mathbf{J}^{-1}(\mu)$ is topologically the same as the rotation group SO(3), which in turn is the same as $S^3/\mathbb{Z}_2$. Thus, the reduction map is a map of SO(3) to S^2. Such a map is given explicitly by taking an orthogonal matrix A and mapping it to the vector on the sphere given by $A\mathbf{k}$, where $\mathbf{k}$ is the unit vector along the z-axis. This map, which does the projection, is in fact a restriction of a momentum map and is, when composed with the map of $S^3 \cong \mathrm{SU}(2)$ to SO(3), just the Hopf fibration again. Thus, not only does the Hopf fibration occur in the one-to-one resonance, *it occurs in the rigid body in a natural way as the reduction map from material to body representation*!

Geometric Phases. The history of this concept is complex. We refer to Berry [1990] for a discussion of the history, going back to Bortolotti in 1926, Vladimirskii and Rytov in 1938 in the study of polarized light, Kato in 1950, and Longuet–Higgins and others in 1958 in atomic physics. Some additional historical comments regarding phases in rigid-body mechanics are given below.

We pick up the story with the classical example of the Foucault pendulum. The Foucault pendulum gives an interesting phase shift (a shift in the angle of the plane of the pendulum's swing) when the overall system undergoes a cyclic evolution (the pendulum is carried in a circular motion due to the Earth's rotation). This phase shift is geometric in character: If one parallel transports an orthonormal frame along the same line of latitude, it returns with a phase shift equaling that of the Foucault pendulum. This phase shift $\Delta\theta = 2\pi\cos\alpha$ (where α is the co-latitude) has the geometric meaning shown in Figure 1.10.1.

In geometry, when an orthonormal frame returns to its original position after traversing a closed path but is rotated, the rotation is referred to as ***holonomy*** (or ***anholonomy***). This is a unifying mathematical concept that underlies many geometric phases in systems such as fiber optics, MRI (magnetic resonance imaging), amoeba propulsion, molecular dynamics, and micromotors. These applications represent one reason the subject is of such current interest.

In the quantum case a seminal paper on geometric phases is Kato [1950]. It was Berry [1984, 1985], Simon [1983], Hannay [1985], and Berry and

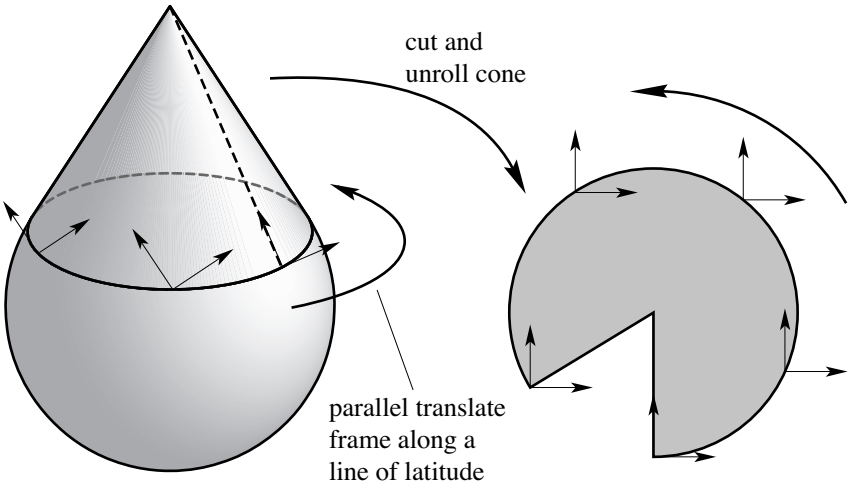

FIGURE 1.10.1. The geometric interpretation of the Foucault pendulum phase shift.

Hannay [1988] who realized that holonomy is the crucial geometric unifying thread. On the other hand, Golin, Knauf, and Marmi [1989], Montgomery [1988], and Marsden, Montgomery, and Ratiu [1989, 1990] demonstrated that averaging connections and reduction of mechanical systems with symmetry also plays an important role, both classically and quantum-mechanically. Aharonov and Anandan [1987] have shown that the geometric phase for a closed loop in projectivized complex Hilbert space occurring in quantum mechanics equals the exponential of the symplectic area of a two–dimensional manifold whose boundary is the given loop. The symplectic form in question is naturally induced on the projective space from the canonical symplectic form of complex Hilbert space (minus the imaginary part of the inner product) via reduction. Marsden, Montgomery, and Ratiu [1990] show that this formula is the holonomy of the closed loop relative to a principal S^1-connection on the unit ball of complex Hilbert space and is a particular case of the holonomy formula in principal bundles with abelian structure group.

Geometric Phases and Locomotion. Geometric phases naturally occur in families of integrable systems depending on parameters. Consider an integrable system with action-angle variables

$$(I_1, I_2, \dots, I_n, \theta_1, \theta_2, \dots, \theta_n);$$

assume that the Hamiltonian $H(I_1, I_2, \dots, I_n; m)$ depends on a parameter $m \in M$. This just means that we have a Hamiltonian independent of the angular variables θ and we can identify the configuration space with an n-torus $\mathbb{T}^n$. Let c be a loop based at a point m_0 in M. We want to compare the

angular variables in the torus over m_0, while the system is slowly changed as the parameters traverse the circuit c. Since the dynamics in the fiber vary as we move along c, even if the actions vary by a negligible amount, there will be a shift in the angle variables due to the frequencies $\omega^i = \partial H/\partial I^i$ of the integrable system; correspondingly, one defines

$$\textbf{\textit{dynamic phase}} = \int_0^1 \omega^i \left(I, c(t)\right) dt.$$

Here we assume that the loop is contained in a neighborhood whose standard action coordinates are defined. In completing the circuit c, we return to the same torus, so a comparison between the angles makes sense. The actual shift in the angular variables during the circuit is the ***dynamic phase*** plus a correction term called the ***geometric phase***. One of the key results is that this geometric phase is the holonomy of an appropriately constructed connection (called the ***Hannay–Berry connection***) on the torus bundle over M that is constructed from the action–angle variables. The corresponding angular shift, computed by Hannay [1985], is called ***Hannay's angles***, so the actual phase shift is given by

$$\Delta\theta = \text{dynamic phases} + \text{Hannay's angles}.$$

The geometric construction of the Hannay–Berry connection for classical systems is given in terms of momentum maps and averaging in Golin, Knauf, and Marmi [1989] and Montgomery [1988]. Weinstein [1990] makes precise the geometric structures that make possible a definition of the Hannay angles for a cycle in the space of Lagrangian submanifolds, even without the presence of an integrable system. Berry's phase is then seen as a "primitive" for the Hannay angles. A summary of this work is given in Woodhouse [1992].

Another class of examples where geometric phases naturally arise is the dynamics of coupled rigid bodies. The three-dimensional single rigid body is discussed below. For several coupled rigid bodies, the dynamics can be quite complex. For instance, even for three coupled bodies in the plane, the dynamics are known to be chaotic, despite the presence of stable relative equilibria; see Oh, Sreenath, Krishnaprasad, and Marsden [1989]. Geometric phase phenomena for this type of example are quite interesting as they are in some of the work of Wilczek and Shapere on locomotion in microorganisms. (See, for example, Shapere and Wilczek [1987, 1989] and Wilczek and Shapere [1989].) In this problem, control of the system's *internal or shape variables* can lead to phase changes in the *external or group variables*. These choices of variables are related to the variables in the reduced and the unreduced phase spaces. In this setting one can formulate interesting questions of optimal control such as "When a falling cat turns itself over in mid-flight (all the time with zero angular momentum!), does it do so with optimal efficiency in terms of, say, energy expended?" There are interesting answers to these questions that are related to the dynamics of

Yang–Mills particles moving in the associated gauge field of the problem. See Montgomery [1984, 1990] and references therein.

We give two simple examples of geometric phases for linked rigid bodies. Additional details can be found in Marsden, Montgomery, and Ratiu [1990]. First, consider three uniform coupled bars (or coupled planar rigid bodies) linked together with pivot (or pin) joints, so the bars are free to rotate relative to each other. Assume that the bars are moving freely in the plane with no external forces and that the angular momentum is zero. However, assume that the joint angles can be controlled with, say, motors in the joints. Figure 1.10.2 shows how the joints can be manipulated, each one going through an angle of 2π and yet the overall assemblage rotates through an angle π.

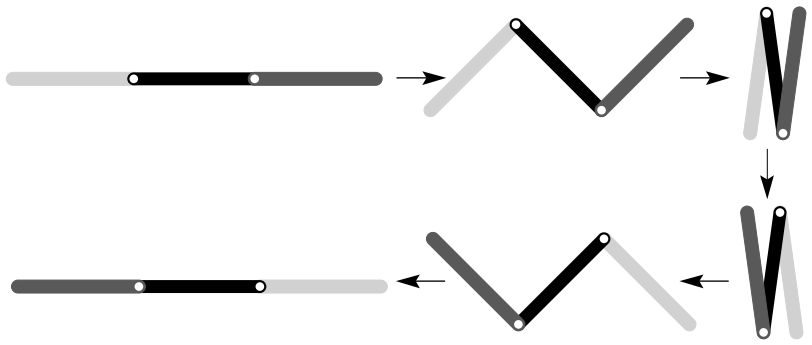

FIGURE 1.10.2. Manipulating the joint angles can lead to an overall rotation of the system.

Here we assume that the moments of inertia of the two outside bars (about an axis through their centers of mass and perpendicular to the page) are each one-half that of the middle bar. The statement is verified by examining the equation for zero angular momentum (see, for example Sreenath, Oh, Krishnaprasad, and Marsden [1988] and Oh, Sreenath, Krishnaprasad, and Marsden [1989]). General formulas for the reconstruction phase applicable to examples of this type are given in Krishnaprasad [1989].

A second example is the dynamics of linkages. This type of example is considered in Krishnaprasad [1989], Yang and Krishnaprasad [1990], including comments on the relation with the three-manifold theory of Thurston. Here one considers a linkage of rods, say four rods linked by pivot joints as in Figure 1.10.3.

The system is free to rotate without external forces or torques, but there are assumed to be torques at the joints. When one turns the small "crank" the whole assemblage turns, even though the angular momentum, as in the previous example, stays zero.

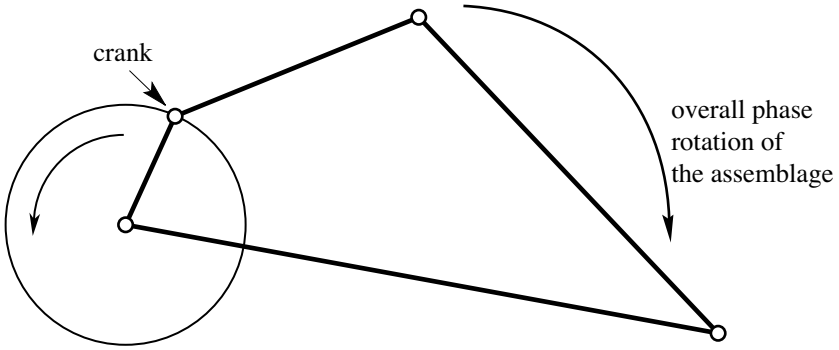

FIGURE 1.10.3. Turning the crank can lead to an overall phase shift.

For an overview of how geometric phases are used in robotic locomotion problems, see Marsden and Ostrowski [1998]. (This paper is available at `http://www.cds.caltech.edu/~marsden`.)

Phases in Rigid-Body Dynamics. As we shall see in Chapter 15, the motion of a rigid body is a geodesic with respect to a left-invariant Riemannian metric (the inertia tensor) on the rotation group SO(3). The corresponding phase space is $P = T^*\,\mathrm{SO}(3)$ and the momentum map $\mathbf{J} : P \to \mathbb{R}^3$ for the *left* SO(3) action is *right* translation to the identity. We identify $\mathfrak{so}(3)^*$ with $\mathfrak{so}(3)$ via the standard inner product and identify $\mathbb{R}^3$ with $\mathfrak{so}(3)$ via the map $v \mapsto \hat{v}$, where $\hat{v}(w) = v \times w$, $\times$ being the standard cross product. Points in $\mathfrak{so}(3)^*$ are regarded as the left reduction of $T^*\,\mathrm{SO}(3)$ by $G = \mathrm{SO}(3)$ and are the angular momenta as seen from a *body-fixed* frame.

The reduced spaces $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$ are identified with spheres in $\mathbb{R}^3$ of Euclidean radius $\|\mu\|$, with their symplectic form $\omega_\mu = -dS/\|\mu\|$, where dS is the standard area form on a sphere of radius $\|\mu\|$ and where G_μ consists of rotations about the μ-axis. The trajectories of the reduced dynamics are obtained by intersecting a family of homothetic ellipsoids (the energy ellipsoids) with the angular momentum spheres. In particular, all but at most four of the reduced trajectories are periodic. These four exceptional trajectories are the well-known homoclinic trajectories; we shall determine them explicitly in §15.8.

Suppose a reduced trajectory $\mathbf{\Pi}(t)$ is given on P_μ, with period T. *After time T, by how much has the rigid body rotated in space*? The spatial angular momentum is $\pi = \mu = g\mathbf{\Pi}$, which is the conserved value of $\mathbf{J}$. Here $g \in \mathrm{SO}(3)$ is the attitude of the rigid body and $\mathbf{\Pi}$ is the body angular momentum. If $\mathbf{\Pi}(0) = \mathbf{\Pi}(T)$, then

$$\mu = g(0)\mathbf{\Pi}(0) = g(T)\mathbf{\Pi}(T),$$

and so $g(T)^{-1}\mu = g(0)^{-1}\mu$, that is, $g(T)g(0)^{-1}\mu = \mu$, so $g(T)g(0)^{-1}$ is a rotation about the axis μ. We want to give the angle of this rotation.

To determine this angle, let $c(t)$ be the corresponding trajectory in $\mathbf{J}^{-1}(\mu) \subset P$. Identify $T^*\,\mathrm{SO}(3)$ with $\mathrm{SO}(3)\times\mathbb{R}^3$ by left trivialization, so $c(t)$ gets identified with $(g(t), \mathbf{\Pi}(t))$. Since the reduced trajectory $\mathbf{\Pi}(t)$ closes after time T, we recover the fact that $c(T) = gc(0)$ for some $g \in G_\mu$. Here, $g = g(T)g(0)^{-1}$ in the preceding notation. Thus, we can write

$$g = \exp[(\Delta\theta)\zeta], \tag{1.10.6}$$

where $\zeta = \mu/\|\mu\|$ identifies $\mathfrak{g}_\mu$ with $\mathbb{R}$ by $a\zeta \mapsto a$, for $a \in \mathbb{R}$. Let D be one of the two spherical caps on S^2 enclosed by the reduced trajectory, let Λ be the corresponding oriented solid angle, that is, $|\Lambda| = (\text{area } D)/\|\mu\|^2$, and let H_μ be the energy of the reduced trajectory. See Figure 1.10.4. All norms are taken relative to the Euclidean metric of $\mathbb{R}^3$. Montgomery [1991a] and Marsden, Montgomery, and Ratiu [1990] show that modulo 2π, we have the ***rigid-body phase formula***

$$\Delta\theta = \frac{1}{\|\mu\|}\left\{\int_D \omega_\mu + 2H_\mu T\right\} = -\Lambda + \frac{2H_\mu T}{\|\mu\|}. \tag{1.10.7}$$

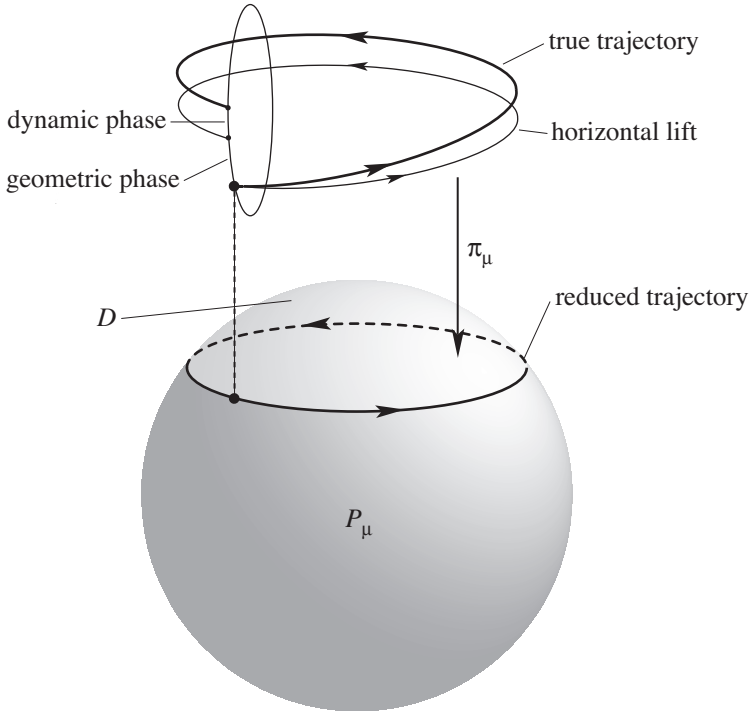

FIGURE 1.10.4. The geometry of the rigid-body phase formula.

More History. The history of the rigid-body phase formula is quite interesting and seems to have proceeded independently of the other developments above.[11] The formula has its roots in work of MacCullagh dating back to 1840 and Thomson and Tait [1867, §§123, 126]. (See Zhuravlev [1996] and O'Reilly [1997] for a discussion and extensions.) A special case of formula (1.10.7) is given in Ishlinskii [1952]; see also Ishlinskii [1963].[12] The formula referred to covers a special case in which only the geometric phase is present. For example, in certain precessional motions in which, up to a certain order in averaging, one can ignore the dynamic phase, and only the geometric phase survives. Even though Ishlinskii found only special cases of the result, he recognized that it is related to the geometric concept of parallel transport. A formula like the one above was found by Goodman and Robinson [1958] in the context of drift in gyroscopes; their proof is based on the Gauss–Bonnet theorem. Another interesting approach to formulas of this sort, also based on averaging and solid angles, is given in Goldreich and Toomre [1969], who applied it to the interesting geophysical problem of polar wander (see also Poincaré [1910]!).

The special case of the above formula for a *symmetric* free rigid body was given by Hannay [1985] and Anandan [1988, formula (20)]. The proof of the general formula based on the theory of connections and the formula for holonomy in terms of curvature was given by Montgomery [1991a] and Marsden, Montgomery, and Ratiu [1990]. The approach using the Gauss–Bonnet theorem and its relation to the Poinsot construction along with additional results is taken up by Levi [1993]. For applications to general resonance problems (such as the three-wave interaction) and nonlinear optics, see Alber, Luther, Marsden and Robbins [1998].

An analogue of the rigid-body phase formula for the heavy top and the Lagrange top (symmetric heavy top) was given in Marsden, Montgomery, and Ratiu [1990]. Links with vortex filament configurations were given in Fukumoto and Miyajima [1996] and Fukumoto [1997].

Satellites with Rotors and Underwater Vehicles. Another example that naturally gives rise to geometric phases is the rigid body with one or more internal rotors. Figure 1.10.5 illustrates the system considered. To specify the position of this system we need an element of the group of rigid motions of $\mathbb{R}^3$ to place the center of mass and the attitude of the carrier, and an angle (element of S^1) to position each rotor. Thus the configuration space is $Q = \mathrm{SE}(3) \times S^1 \times S^1 \times S^1$. The equations of motion of this system are an extension of Euler's equations of motion for a freely spinning rotor. Just as holding a spinning bicycle wheel while sitting on a swivel chair can affect the carrier's motion, so the spinning rotors can affect the dynamics

[11] We thank V. Arnold for valuable help with these comments.

[12] On page 195 of Ishlinskii [1976], a later book on mechanics, it is stated that "the formula was found by the author in 1943 and was published in Ishlinskii [1952]."

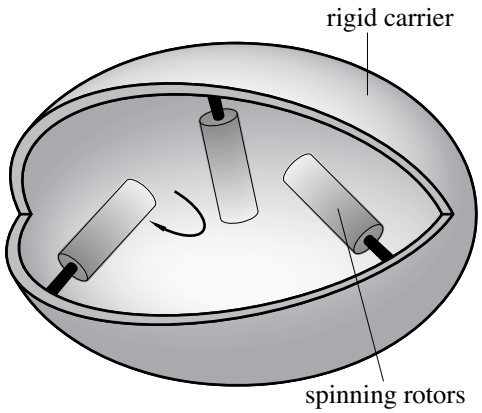

FIGURE 1.10.5. The rigid body with internal rotors.

of the rigid carrier.

In this example, one can analyze equilibria and their stability in much the same way as one can with the rigid body. However, what one often wants to do is to forcibly spin, or control, the rotors so that one can achieve attitude control of the structure in the same spirit that a falling cat has *control* of its attitude by manipulating its body parts while falling. For example, one can attempt to prescribe a relation between the rotor dynamics and the rigid-body dynamics by means of a *feedback law*. This has the property that the total system angular momentum is still preserved and that the resulting dynamic equations can be expressed entirely in terms of the free rigid-body variable. (A falling cat has zero angular momentum even though it is able to turn over!) In some cases the resulting equations are again Hamiltonian on the invariant momentum sphere. Using this fact, one can compute the geometric phase for the problem generalizing the free rigid-body phase formula. (See Bloch, Krishnaprasad, Marsden, and Sánchez de Alvarez [1992] and Bloch, Leonard, and Marsden [1997, 1998] for details.) This type of analysis is useful in designing and understanding attitude control devices.

Another example that combines some features of the satellite and the heavy top is the *underwater vehicle*. This is in the realm of the dynamics of rigid bodies in fluids, a subject going back to Kirchhoff in the late 1800s. We refer to Leonard and Marsden [1997] and Holmes, Jenkins, and Leonard [1998] for modern accounts and many references.

Miscellaneous Links. There are many continuum-mechanical examples to which the techniques of geometric mechanics apply. Some of those are free boundary problems (Lewis, Marsden, Montgomery, and Ratiu [1986], Montgomery, Marsden, and Ratiu [1984], Mazer and Ratiu [1989]), spacecraft with flexible attachments (Krishnaprasad and Marsden [1987]), elas-

ticity (Holm and Kupershmidt [1983], Kupershmidt and Ratiu [1983], Marsden, Ratiu, and Weinstein [1984a, 1984b], Simo, Marsden, and Krishnaprasad [1988]), and reduced MHD (Morrison and Hazeltine [1984] and Marsden and Morrison [1984]). We also wish to look at these theories from both the spatial (Eulerian) and body (convective) points of view as reductions of the canonical material picture. These two reductions are, in an appropriate sense, dual to each other.

The geometric-analytic approach to mechanics finds use in a number of other diverse areas as well. We mention just a few samples.

- Integrable systems (Moser [1980], Perelomov [1990], Adams, Harnad, and Previato [1988], Fomenko and Trofimov [1989], Fomenko [1988a, 1988b], Reyman and Semenov-Tian-Shansky [1990], and Moser and Veselov [1991]).

- Applications of integrable systems to numerical analysis (like the QR algorithm and sorting algorithms); see Deift and Li [1989] and Bloch, Brockett, and Ratiu [1990, 1992].

- Numerical integration (Sanz-Serna and Calvo [1994], Marsden, Patrick, and Shadwick [1996], Wendlandt and Marsden [1997], Marsden, Patrick, and Shkoller [1998]).

- Hamiltonian chaos (Arnold [1964], Ziglin [1980a, 1980b, 1981], Holmes and Marsden [1981, 1982a, 1982b, 1983], Wiggins [1988]).

- Averaging (Cushman and Rod [1982], Iwai [1982, 1985], Ercolani, Forest, McLaughlin, and Montgomery [1987]).

- Hamiltonian bifurcations (van der Meer [1985], Golubitsky and Schaeffer [1985], Golubitsky and Stewart [1987], Golubitsky, Stewart, and Schaeffer [1988], Lewis, Marsden, and Ratiu [1987], Lewis, Ratiu, Simo, and Marsden [1992], Montaldi, Roberts, and Stewart [1988], Golubitsky, Marsden, Stewart, and Dellnitz [1995]).

- Algebraic geometry (Atiyah [1982, 1983], Kirwan [1984, 1985 1998]).

- Celestial mechanics (Deprit [1983], Meyer and Hall [1992]).

- Vortex dynamics (Ziglin [1980b], Koiller, Soares, and Melo Neto [1985], Wan and Pulvirente [1984], Wan [1986, 1988a, 1988b, 1988c], Kirwan [1988], Szeri and Holmes [1988], Newton [1994], Pekarsky and Marsden [1998]).

- Solitons (Flaschka, Newell, and Ratiu [1983a, 1983b], Newell [1985], Kovačič and Wiggins [1992], Alber and Marsden [1992]).

- Multisymplectic geometry, PDEs, and nonlinear waves (Gotay, Isenberg, and Marsden [1997], Bridges [1994, 1997], Marsden and Shkoller [1999], and Marsden, Patrick, and Shkoller [1998]).

- Relativity and Yang–Mills theory (Fischer and Marsden [1972, 1979], Arms [1981], Arms, Marsden, and Moncrief [1981, 1982]).

- Fluid variational principles using Clebsch variables and "Lin constraints" (Seliger and Whitham [1968], Cendra and Marsden [1987], Cendra, Ibort, and Marsden [1987], Holm, Marsden, and Ratiu [1998a]).

- Control, stabilization, satellite and underwater vehicle dynamics (Krishnaprasad [1985], van der Schaft and Crouch [1987], Aeyels and Szafranski [1988], Bloch, Krishnaprasad, Marsden, and Sánchez de Alvarez [1992], Wang, Krishnaprasad, and Maddocks [1991], Leonard [1997], Leonard and Marsden [1997]), Bloch, Leonard, and Marsden [1998], and Holmes, Jenkins, and Leonard [1998]).

- Nonholonomic systems (Naimark and Fufaev [1972], Koiller [1992], Bates and Sniatycki [1993], Bloch, Krishnaprasad, Marsden, and Murray [1996], Koon and Marsden [1997a, 1997b, 1998], Zenkov, Bloch, and Marsden [1998]).

Reduction theory for mechanical systems with symmetry is a natural historical continuation of the works of Liouville (for integrals in involution) and of Jacobi (for angular momentum) for reducing the phase space dimension in the presence of first integrals. It is intimately connected with work on momentum maps, and its forerunners appear already in Jacobi [1866], Lie [1890], Cartan [1922], and Whittaker [1927]. It was developed later in Kirillov [1962], Arnold [1966a], Kostant [1970], Souriau [1970], Smale [1970], Nekhoroshev [1977], Meyer [1973], and Marsden and Weinstein [1974]. See also Guillemin and Sternberg [1984] and Marsden and Ratiu [1986] for the Poisson case and Sjamaar and Lerman [1991] for basic work on the singular symplectic case.