



A Crash Course in Geometric Mechanics

Tudos Ratiu

► To cite this version:

Tudos Ratiu. A Crash Course in Geometric Mechanics. 3rd cycle. Monastir (Tunisie), 2005, pp.134.
cel-00391890

HAL Id: cel-00391890

<https://cel.hal.science/cel-00391890>

Submitted on 5 Jun 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

V

A Crash Course in Geometric Mechanics

T.S. Ratiu, R. Tudoran, L. Sbano, E. Sousa Dias & G. Terra

Notes of the courses given by Tudor Ratiu

1 Introduction

These lecture notes are the direct result of presentations held in two consecutive years (2000 and 2001) at the Peyresq Conference Center, in the Alpes de Haute Provence, North of Nice. These summer schools were organized in conjunction with the Research Training Network MASIE (Mechanics and Symmetry in Europe) of the Fifth Framework Program of the European Commission and were intended for graduate students and postdocs who needed a crash course in geometric mechanics. They were also tailored to link with the other lectures and provide the necessary background for them. There are already many books on this subject and its links to symplectic and Poisson geometry (see, e.g. [AbMa78], [Arnold79], [GuSt84], [JoSa98], [LiMa87], [MaRa94], [McDSal95]) and the literature on this subject is overwhelming. So the goal of these two one-week intensive lectures was to find a quick way through this subject and give the young researchers enough tools to be able to sift and sort through the books and papers necessary for their own work. This is why these lectures present occasionally detailed proofs and sometimes only quick surveys of more extensive subjects that are, however, explained with care. The examples, on the other hand, are all carried out with detailed computations in order to show how one applies the theory in concrete cases. There are, essentially, four main examples that reappear throughout the lectures: particle dynamics, the free and heavy tops, the motion of a charged particle in a magnetic field, and ideal incompressible fluid flow as well as related systems such as the Korteweg-de Vries and the Camassa-Holm equations. Each example illustrates several different constructions prevalent in geometric mechanics and is at the root of many developments and generalizations.

There is nothing original in these lectures and they are entirely based on three main sources: [MaRa94], the yet unfinished book [MaRa03], and some unpublished notes [MaRa95] on the geometric theory of fluid dynamics. When carrying out reduction, several strong regularity hypotheses will be made. The singular case is considerably more involved and we refer to [OR04] for an in-depth analysis of this case. All the missing proofs of results quoted here can be found in these works as well as the books referred to before.

The reader is assumed to be familiar with calculus on manifolds and the elementary theory of Lie groups and Lie algebras, as found in e.g. [AbMa78], [AMR88], [DFN95], [Jost], [Lang], [MaRa94], [Sp79], [Serre], or [W83]. [Bou71, Bou89] is always helpful when a quick recall of the statement of a theorem is needed.

1.1 Lagrangian and Hamiltonian Formalism

Let us start with Newton's equations for N particles $\mathbf{q} := (\mathbf{q}_1, \dots, \mathbf{q}_N) \in \mathbb{R}^{3N}$ with masses $m_1, \dots, m_N \in \mathbb{R}$. If $\mathbf{F} = (F_1, \dots, F_N)$ are the forces acting on these particles then *Newton's equations* are

$$m\mathbf{a} = \mathbf{F}, \quad (1.1)$$

where $\mathbf{a} = \ddot{\mathbf{q}}$ is the acceleration of the system. Assuming that the forces are induced by a potential $V : \mathbb{R}^{3N} \rightarrow \mathbb{R}$, that is,

$$\mathbf{F}(\mathbf{q}) = -\nabla V(\mathbf{q}), \quad (1.2)$$

equations (1.1) become

$$m_i \ddot{\mathbf{q}}_i = -\frac{\partial V}{\partial \mathbf{q}_i}, \quad i = 1, \dots, N, \quad (1.3)$$

where $\partial V / \partial \mathbf{q}_i$ denotes the gradient relative to the variable $\mathbf{q}_i$.

A straightforward verification shows that if one defines the *Lagrangian* function $L : \mathbb{R}^{6N} = \{(\mathbf{q}, \dot{\mathbf{q}}) \mid \mathbf{q}, \dot{\mathbf{q}} \in \mathbb{R}^{3N}\} \rightarrow \mathbb{R}$ by

$$L(\mathbf{q}, \dot{\mathbf{q}}) := \frac{1}{2} \sum_{i=1}^N m_i \|\dot{\mathbf{q}}_i\|^2 - V(\mathbf{q}). \quad (1.4)$$

and assumes that $\dot{\mathbf{q}} = d\mathbf{q}/dt$, then Newton's equations (1.2) are equivalent to *Lagrange's equations*

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\mathbf{q}}_i} \right) - \frac{\partial L}{\partial \mathbf{q}_i} = 0, \quad i = 1, \dots, N, \quad (1.5)$$

where $\partial L/\partial \dot{\mathbf{q}}_i, \partial L/\partial \mathbf{q}_i \in \mathbb{R}^3$ denote the gradients in $\mathbb{R}^3$ of L relative to $\dot{\mathbf{q}}_i, \mathbf{q}_i \in \mathbb{R}^3$.

On other hand, the Lagrange equations (1.5) are equivalent with the variational principle of Hamilton: *the solutions of (1.5) are critical points of the action functional defined on the space of smooth paths with fixed endpoints.* More precisely, let $\Lambda([a, b], \mathbb{R}^{3N})$ be the space of all possible smooth trajectories $\mathbf{q} : [a, b] \rightarrow \mathbb{R}^{3N}$ with fixed endpoints $\mathbf{q}_a = \gamma(a), \mathbf{q}_b = \gamma(b)$. The **action functional** is defined by:

$$\mathcal{A}[\mathbf{q}(\cdot)] := \int_a^b L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) dt, \quad (1.6)$$

where $\dot{\mathbf{q}} = d\mathbf{q}(t)/dt$. In $\Lambda([a, b], \mathbb{R}^{3N})$ consider a deformation $\mathbf{q}(t, s), s \in (-\epsilon, \epsilon), \epsilon > 0$, with fixed endpoints $\mathbf{q}_a, \mathbf{q}_b$, of a curve $\mathbf{q}_0(t)$, that is, $\mathbf{q}(t, 0) = \mathbf{q}_0(t)$ for all $t \in [a, b]$ and $\mathbf{q}(a, s) = \mathbf{q}_0(a) = \mathbf{q}_a, \mathbf{q}(b, s) = \mathbf{q}_0(b) = \mathbf{q}_b$ for all $s \in (-\epsilon, \epsilon)$. Define a **variation** of the curve $\mathbf{q}_0(\cdot)$ in $\Lambda([a, b], \mathbb{R}^{3N})$ by

$$\delta \mathbf{q}(\cdot) := \left. \frac{d}{ds} \right|_{s=0} \mathbf{q}(\cdot, s) \in T_{\mathbf{q}_0(\cdot)} \Lambda([a, b], \mathbb{R}^{3N}),$$

and the **first variation** of $\mathcal{A}$ at $\mathbf{q}_0(t)$ to be the following derivative:

$$\mathbf{D}\mathcal{A}[\mathbf{q}_0(\cdot)](\delta \mathbf{q}(\cdot)) := \left. \frac{d}{ds} \right|_{s=0} \mathcal{A}[\mathbf{q}(\cdot, s)]. \quad (1.7)$$

Note that $\delta \mathbf{q}(a) = \delta \mathbf{q}(b) = \mathbf{0}$. With these notations, the **variational principle of Hamilton** states that the curve $\mathbf{q}_0(t)$ satisfies the Lagrange equations (1.5) if and only if $\mathbf{q}_0(\cdot)$ is a critical point of the action functional, that is, $\mathbf{D}\mathcal{A}[\mathbf{q}_0(\cdot)] = \mathbf{0}$. Indeed, using the equality of mixed partials, integrating by parts, and taking into account that $\delta \mathbf{q}(a) = \delta \mathbf{q}(b) = \mathbf{0}$, we get

$$\begin{aligned} \mathbf{D}\mathcal{A}[\mathbf{q}_0(\cdot)](\delta \mathbf{q}(\cdot)) &= \left. \frac{d}{ds} \right|_{s=0} \mathcal{A}[\mathbf{q}(\cdot, s)] = \left. \frac{d}{ds} \right|_{s=0} \int_a^b L(\mathbf{q}(t, s), \dot{\mathbf{q}}(t, s)) dt \\ &= \int_a^b \left[\frac{\partial L}{\partial \mathbf{q}_i} \delta \mathbf{q}(t, s) + \frac{\partial L}{\partial \dot{\mathbf{q}}_i} \delta \dot{\mathbf{q}}_i \right] dt \\ &= - \int_a^b \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\mathbf{q}}_i} \right) - \frac{\partial L}{\partial \mathbf{q}_i} \right] \delta \mathbf{q}_i dt = 0 \end{aligned}$$

for all smooth $\delta \mathbf{q}_i(t)$ satisfying $\delta \mathbf{q}_i(a) = \delta \mathbf{q}_i(b) = \mathbf{0}$, which proves the claim.

Next, introduce the **conjugate momenta**

$$\mathbf{p}^i := \frac{\partial L}{\partial \dot{\mathbf{q}}_i} = m_i \dot{\mathbf{q}}_i \in \mathbb{R}^3, \quad i = 1, \dots, N. \quad (1.8)$$

Define the change of variables $(\mathbf{q}, \dot{\mathbf{q}}) \mapsto (\mathbf{q}, \mathbf{p})$, called the **Legendre transform**, and the **Hamiltonian**

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}) &:= \mathbf{p} \cdot \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}) - L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p})) \\ &= \frac{1}{2} \sum_{i=1}^N m_i \|\dot{\mathbf{q}}_i\|^2 + V(\mathbf{q}) \\ &= \frac{1}{2} \sum_{i=1}^N \frac{1}{m_i} \|\mathbf{p}^i\|^2 + V(\mathbf{q}) \end{aligned} \quad (1.9)$$

which is the total energy of the system, expressed in the variables $(\mathbf{q}, \mathbf{p})$. Then one has

$$\frac{\partial H}{\partial \mathbf{p}^i} = \frac{1}{m_i} \mathbf{p}^i = \dot{\mathbf{q}}_i = \frac{d\mathbf{q}_i}{dt}$$

and

$$\frac{\partial H}{\partial \mathbf{q}_i} = \frac{\partial V}{\partial \mathbf{q}_i} = -\frac{\partial L}{\partial \mathbf{q}_i}.$$

Therefore, by the Lagrange equations (1.5) we have

$$\dot{\mathbf{p}}^i = \frac{d\mathbf{p}^i}{dt} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\mathbf{q}}_i} \right) = \frac{\partial L}{\partial \mathbf{q}_i} = -\frac{\partial H}{\partial \mathbf{q}_i}.$$

This shows that Lagrange's equations (1.5) are equivalent to **Hamilton's equations**

$$\dot{\mathbf{q}}_i = \frac{\partial H}{\partial \mathbf{p}^i} \quad \dot{\mathbf{p}}^i = -\frac{\partial H}{\partial \mathbf{q}_i}, \quad (1.10)$$

where, as before, $\partial H / \partial \mathbf{q}_i, \partial H / \partial \mathbf{p}^i \in \mathbb{R}^3$ are the gradients of H relative to $\mathbf{q}_i, \mathbf{p}^i \in \mathbb{R}^3$.

Note that, whereas Lagrange's equations are of second order and concern curves in the configuration space (the space of $\mathbf{q}$'s), Hamilton's equations are of first order and describe the dynamics of curves belonging to **phase space**, a space of dimension twice the dimension of the configuration space whose points are pairs formed by configurations $\mathbf{q}$ and conjugate momenta $\mathbf{p}$.

An easy verification shows that Hamilton's equations (1.10) can be equivalently written as

$$\dot{F} = \{F, H\} \quad \text{for all } F \in \mathcal{F}(P), \quad (1.11)$$

where $\mathcal{F}(P)$ denotes the smooth functions on the phase space $P := \mathbb{R}^{6N}$, and

the **Poisson bracket** is defined by

$$\{G, K\} := \sum_{i=1}^N \left(\frac{\partial G}{\partial \mathbf{q}_i} \cdot \frac{\partial K}{\partial \mathbf{p}^i} - \frac{\partial G}{\partial \mathbf{p}^i} \cdot \frac{\partial K}{\partial \mathbf{q}_i} \right) \quad \text{for all } G, K \in \mathcal{F}(P). \quad (1.12)$$

Indeed, from (1.12) and (1.10) we have, for any $F \in \mathcal{F}(P)$,

$$\begin{aligned} \sum_{i=1}^N \left(\frac{\partial F}{\partial \mathbf{q}_i} \cdot \dot{\mathbf{q}}_i + \frac{\partial F}{\partial \mathbf{p}^i} \cdot \dot{\mathbf{p}}^i \right) &= \frac{dF}{dt} \\ &= \{F, H\} = \sum_{i=1}^N \left(\frac{\partial F}{\partial \mathbf{q}_i} \cdot \frac{\partial H}{\partial \mathbf{p}^i} - \frac{\partial F}{\partial \mathbf{p}^i} \cdot \frac{\partial H}{\partial \mathbf{q}_i} \right) \end{aligned}$$

which is equivalent to (1.10) since $F \in \mathcal{F}(P)$ is arbitrary.

Summarizing, for classical mechanical systems in Euclidean space describing particle motion, whose total energy is given by kinetic plus potential energy, we have shown that Newton's equations are equivalent to:

- Lagrange's equations
- Hamilton's variational principle
- Hamilton's equations of motion
- Hamilton's equations in Poisson bracket formulation.

In the course of these lectures we shall focus on each one of these four pictures and shall explain the geometric structure underlying them when the configuration space is a general manifold. It turns out that, in general, they are not equivalent and, moreover, some of these formulations have very useful generalizations, particularly appropriate for systems with symmetry, the case we shall consider next by means of an example.

1.2 The Heavy Top

In these lectures we shall discuss the equivalences just described in the context of systems with symmetry when one can eliminate variables. To see what is involved in this case, let us consider in detail an example, namely the motion of a heavy top moving about a fixed point; the exposition below is mostly based on [MaRa94, MaRaWe84a, MaRaWe84b, HMR98]. In this case, all computations can be easily carried out explicitly. We shall describe this system both in the Lagrangian and Hamiltonian picture and shall find two additional equivalent formulations that take into account the symmetries of this system. This example will serve then as model for the reduction theory presented in these lectures.

The Lie algebra $\mathfrak{so}(3)$ and its dual. To be efficient in the computations that follow we briefly recall the main formulas regarding the special orthogonal group $SO(3) := \{A \mid A \text{ a } 3 \times 3 \text{ orthogonal matrix, } \det(A) = 1\}$, its Lie algebra $\mathfrak{so}(3)$ formed by 3×3 skew symmetric matrices, and its dual $\mathfrak{so}(3)^*$. All these formulas will be proved in §5.1 and §6.3. The Lie algebra $(\mathfrak{so}(3), [\cdot, \cdot])$, where $[\cdot, \cdot]$ is the commutator bracket of matrices, is isomorphic to the Lie algebra $(\mathbb{R}^3, \times)$, where $\times$ denotes the vector product in $\mathbb{R}^3$, by the isomorphism

$$\mathbf{u} := (u^1, u^2, u^3) \in \mathbb{R}^3 \mapsto \hat{\mathbf{u}} := \begin{bmatrix} 0 & -u^3 & u^2 \\ u^3 & 0 & -u^1 \\ -u^2 & u^1 & 0 \end{bmatrix} \in \mathfrak{so}(3). \quad (1.13)$$

Equivalently, this isomorphism is given by

$$\hat{\mathbf{u}}\mathbf{v} = \mathbf{u} \times \mathbf{v} \quad \text{for all } \mathbf{u}, \mathbf{v} \in \mathbb{R}^3. \quad (1.14)$$

The following properties for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ are easily checked:

$$(\mathbf{u} \times \mathbf{v})^\wedge = [\hat{\mathbf{u}}, \hat{\mathbf{v}}] \quad (1.15)$$

$$[\hat{\mathbf{u}}, \hat{\mathbf{v}}]\mathbf{w} = (\mathbf{u} \times \mathbf{v}) \times \mathbf{w} \quad (1.16)$$

$$\mathbf{u} \cdot \mathbf{v} = -\frac{1}{2} \text{trace}(\hat{\mathbf{u}}\hat{\mathbf{v}}). \quad (1.17)$$

If $A \in SO(3)$ and $\hat{\mathbf{u}} \in \mathfrak{so}(3)$ denote, as usual, by $\text{Ad}_A \hat{\mathbf{u}} := A\hat{\mathbf{u}}A^{-1}$ the adjoint action of $SO(3)$ on its Lie algebra $\mathfrak{so}(3)$. Then

$$(A\mathbf{u})^\wedge = \text{Ad}_A \hat{\mathbf{u}} := A\hat{\mathbf{u}}A^T \quad (1.18)$$

since $A^{-1} = A^T$, the transpose of A . Also

$$A(\mathbf{u} \times \mathbf{v}) = A\mathbf{u} \times A\mathbf{v} \quad (1.19)$$

for any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ and $A \in SO(3)$. It should be noted that this relation is *not* valid if A is just an orthogonal matrix; if A is not in the component of the identity matrix, then one gets a minus sign on the right hand side.

The dual $\mathfrak{so}(3)^*$ is identified with $\mathbb{R}^3$ by the isomorphism $\mathbf{\Pi} \in \mathbb{R}^3 \mapsto \tilde{\mathbf{\Pi}} \in \mathfrak{so}(3)^*$ given by $\tilde{\mathbf{\Pi}}(\hat{\mathbf{u}}) := \mathbf{\Pi} \cdot \mathbf{u}$ for any $\mathbf{u} \in \mathbb{R}^3$. Then the coadjoint action of $SO(3)$ on $\mathfrak{so}(3)^*$ is given by (see §5.1 for the explicit computation)

$$\text{Ad}_{A^{-1}}^* \tilde{\mathbf{\Pi}} = (A\mathbf{\Pi})^\sim. \quad (1.20)$$

The coadjoint action of $\mathfrak{so}(3)$ on $\mathfrak{so}(3)^*$ is given by (see §6.3 for the detailed computation)

$$\text{ad}_{\hat{\mathbf{u}}}^* \tilde{\mathbf{\Pi}} = (\mathbf{\Pi} \times \mathbf{u})^\sim. \quad (1.21)$$

Euler angles. The Lie group $SO(3)$ is diffeomorphic to the real three dimensional projective space $\mathbb{RP}(3)$. The Euler angles that we shall review below provide a very convenient chart for $SO(3)$.

Let $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ be an orthonormal basis of $\mathbb{R}^3$ thought of as the *reference configuration*. Points in the reference configuration, called *material* or *Lagrangian points*, are denoted by $\mathbf{X}$ and their components, called *material* or *Lagrangian coordinates* by (X^1, X^2, X^3) . Another copy of $\mathbb{R}^3$ is thought of as the *spatial* or *Eulerian configuration*; its points, called *spatial* or *Eulerian points* are denoted by $\mathbf{x}$ whose components (x^1, x^2, x^3) relative to an orthonormal basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are called *spatial* or *Eulerian coordinates*. A *configuration* is a map from the reference to the spatial configuration that will be assumed to be an orientation preserving diffeomorphism. If the configuration is defined only on a subset of $\mathbb{R}^3$ with certain good properties such as being a submanifold, as will be the case for the heavy top, then it is assumed that the configuration is a diffeomorphism onto its image. A *motion* $\mathbf{x}(\mathbf{X}, t)$ is a time dependent family of configurations. In what follows we shall only consider motions that are given by rotations, that is, we shall assume that $\mathbf{x}(\mathbf{X}, t) = A(t)\mathbf{X}$ with $A(t)$ an orthogonal matrix. Since the motion is assumed to be smooth and equal to the identity at $t = 0$, it follows that $A(t) \in SO(3)$.

Define the time dependent orthonormal basis ξ_1, ξ_2, ξ_3 by $\xi_i := A(t)\mathbf{E}_i$, for $i = 1, 2, 3$. This basis is anchored in the body and moves together with it. The *body* or *convected coordinates* are the coordinates of a point relative to the basis ξ_1, ξ_2, ξ_3 . Note that the components of a vector $\mathbf{V}$ relative to the basis $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ are the same as the components of the vector $A(t)\mathbf{V}$ relative to the basis ξ_1, ξ_2, ξ_3 . In particular, the body coordinates of $\mathbf{x}(\mathbf{X}, t) = A(t)\mathbf{X}$ are X^1, X^2, X^3 .

The Euler angles encode the passage from the spatial basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ to the body basis ξ_1, ξ_2, ξ_3 by means of three consecutive counterclockwise rotations performed in a specific order: first rotate around the axis $\mathbf{e}_3$ by the angle φ and denote the resulting position of $\mathbf{e}_1$ by $\mathbf{ON}$ (line of nodes), then rotate about $\mathbf{ON}$ by the angle θ and denote the resulting position of $\mathbf{e}_3$ by ξ_3 , and finally rotate about ξ_3 by the angle ψ . Note that, by construction, $0 \leq \varphi, \psi < 2\pi$ and $0 \leq \theta < \pi$ and that the method just described provides a bijective map between (φ, ψ, θ) variables and the group $SO(3)$. However, this bijective map is not a chart since its differential vanishes at $\varphi = \psi = \theta = 0$. So for $0 < \varphi, \psi < 2\pi, 0 < \theta < \pi$ the *Euler angles* (φ, ψ, θ) form a chart. If one carries out explicitly the rotation just described the resulting linear map performing the motion $\mathbf{x}(\mathbf{X}, t) = A(t)\mathbf{X}$ has the matrix relative to the bases ξ_1, ξ_2, ξ_3

and $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ equal to

$$A = \begin{bmatrix} \cos \psi \cos \varphi - \cos \theta \sin \varphi \sin \psi & \cos \psi \sin \varphi + \cos \theta \cos \varphi \sin \psi & \sin \theta \sin \psi \\ -\sin \psi \cos \varphi - \cos \theta \sin \varphi \cos \psi & -\sin \psi \sin \varphi + \cos \theta \cos \varphi \cos \psi & \sin \theta \cos \psi \\ \sin \theta \sin \varphi & -\sin \theta \cos \varphi & \cos \theta \end{bmatrix}; \quad (1.22)$$

this computation is carried out in practically any mechanics book such as [Arnold79] or [MaRa94].

The total energy of the heavy top. A heavy top is by definition a rigid body moving about a fixed point in $\mathbb{R}^3$. Let $\mathcal{B}$ be an open bounded set whose closure is a reference configuration. Points on the reference configuration are denoted, as before, by $\mathbf{X} = (X^1, X^2, X^3)$, with X^1, X^2, X^3 the material coordinates relative to a fixed orthonormal frame $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$. The map $\eta : \mathcal{B} \rightarrow \mathbb{R}^3$, with enough smoothness properties so that all computations below make sense, which is, in addition, orientation preserving and invertible on its image, is a configuration of the top. The spatial points $\mathbf{x} := \eta(\mathbf{X}) \in \eta(\mathcal{B})$ have coordinates x^1, x^2, x^3 relative to an orthonormal basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. Since the body is rigid and has a fixed point, its motion $\eta_t : \mathcal{B} \rightarrow \mathbb{R}^3$ is necessarily of the form

$$\eta_t(\mathbf{X}) := \mathbf{x}(\mathbf{X}, t) = A(t)\mathbf{X}$$

with $A(t) \in SO(3)$; this is a 1932 theorem of Mazur and Ulam which states that any isometry of $\mathbb{R}^3$ that leaves the origin fixed is necessarily a rotation. If ξ_1, ξ_2, ξ_3 is the orthonormal basis of $\mathbb{R}^3$ defined by $\xi_i := A(t)\mathbf{E}_i$, for $i = 1, 2, 3$, then the body coordinates of a vector are its components relative to this basis anchored in the body and moving together with it.

The **material** or **Lagrangian velocity** is defined by

$$\mathbf{V}(\mathbf{X}, t) := \frac{\partial \mathbf{x}(\mathbf{X}, t)}{\partial t} = \dot{A}(t)\mathbf{X}. \quad (1.23)$$

The **spatial** or **Eulerian velocity** is defined by

$$\mathbf{v}(\mathbf{x}, t) := \mathbf{V}(\mathbf{X}, t) = \dot{A}(t)\mathbf{X} = \dot{A}(t)A(t)^{-1}\mathbf{x}. \quad (1.24)$$

The **body** or **convective velocity** is defined by

$$\begin{aligned} \mathcal{V}(\mathbf{X}, t) &:= -\frac{\partial \mathbf{X}(\mathbf{x}, t)}{\partial t} = A(t)^{-1}\dot{A}(t)A(t)^{-1}\mathbf{x} \\ &= A(t)^{-1}\mathbf{V}(\mathbf{X}, t) = A(t)^{-1}\mathbf{v}(\mathbf{x}, t). \end{aligned} \quad (1.25)$$

Denote by ρ_0 the density of the top in the reference configuration. Then the

kinetic energy at time t in material, spatial, and convective representation is given by

$$K(t) = \frac{1}{2} \int_{\mathcal{B}} \rho_0(\mathbf{X}) \|\mathbf{V}(\mathbf{X}, t)\|^2 d^3\mathbf{X} \quad (1.26)$$

$$= \frac{1}{2} \int_{A(t)\mathcal{B}} \rho_0(A(t)^{-1}\mathbf{x}) \|\mathbf{v}(\mathbf{x}, t)\|^2 d^3\mathbf{x} \quad (1.27)$$

$$= \frac{1}{2} \int_{\mathcal{B}} \rho_0(\mathbf{X}) \|\mathcal{V}(\mathbf{X}, t)\|^2 d^3\mathbf{X} \quad (1.28)$$

If we denote

$$\hat{\omega}_S(t) := \dot{A}(t)A(t)^{-1} \quad (1.29)$$

$$\hat{\omega}_B(t) := A(t)^{-1}\dot{A}(t) \quad (1.30)$$

and take into account (1.24), (1.25), and (1.14), we conclude that

$$\mathbf{v}(\mathbf{x}, t) = \omega_S(t) \times \mathbf{x}$$

$$\mathcal{V}(\mathbf{X}, t) = \omega_B(t) \times \mathbf{X}$$

which shows that ω_S and ω_B are the *spatial* and *body angular velocities* respectively. Using the Euler angles representation (1.22), the expressions for ω_S and ω_B are

$$\omega_S = \begin{bmatrix} \dot{\theta} \cos \varphi + \dot{\psi} \sin \varphi \sin \theta \\ \dot{\theta} \sin \varphi - \dot{\psi} \cos \varphi \sin \theta \\ \dot{\varphi} + \dot{\psi} \cos \theta \end{bmatrix} \quad (1.31)$$

$$\omega_B = \begin{bmatrix} \dot{\theta} \cos \psi + \dot{\varphi} \sin \psi \sin \theta \\ -\dot{\theta} \sin \psi + \dot{\varphi} \cos \psi \sin \theta \\ \dot{\varphi} \cos \theta + \dot{\psi} \end{bmatrix}. \quad (1.32)$$

Thus, by (1.28), the kinetic energy in convective representation has the expression

$$K(t) = \frac{1}{2} \int_{\mathcal{B}} \rho_0(\mathbf{X}) \|\omega_B(t) \times \mathbf{X}\|^2 d^3\mathbf{X} =: \frac{1}{2} \langle\langle \omega_B(t), \omega_B(t) \rangle\rangle. \quad (1.33)$$

This is the quadratic form associated to the bilinear symmetric form on $\mathbb{R}^3$

$$\langle\langle \mathbf{a}, \mathbf{b} \rangle\rangle := \int_{\mathcal{B}} \rho_0(\mathbf{X}) (\mathbf{a} \times \mathbf{X}) \cdot (\mathbf{b} \times \mathbf{X}) d^3\mathbf{X} = \mathbb{I} \mathbf{a} \cdot \mathbf{b}, \quad (1.34)$$

where $\mathbb{I} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is the symmetric isomorphism (relative to the dot product) whose components are given by $\mathbb{I}_{ij} := \mathbb{I} \mathbf{E}_j \cdot \mathbf{E}_i = \langle\langle \mathbf{E}_j, \mathbf{E}_i \rangle\rangle$, that is,

$$\mathbb{I}_{ij} = - \int_{\mathcal{B}} \rho_0(\mathbf{X}) X^i X^j d^3\mathbf{X} \quad \text{if } i \neq j$$

and

$$\mathbb{I}_{ii} = \int_{\mathcal{B}} \rho_0(\mathbf{X}) (\|\mathbf{X}\|^2 - (X^i)^2) d^3\mathbf{X}.$$

These are the expressions of the moment of inertia tensor in classical mechanics, that is, $\mathbb{I}$ is the **moment of inertia tensor**. Since $\mathbb{I}$ is symmetric, it can be diagonalized. The basis in which it is diagonal is called in classical mechanics the **principal axis body frame** and the diagonal elements I_1, I_2, I_3 of $\mathbb{I}$ in this basis are called the **principal moments of inertia** of the top. From now on, we choose the basis $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ to be a principal axis body frame.

Identify in what follows the linear functional $\langle\langle \omega_B, \cdot \rangle\rangle$ on $\mathbb{R}^3$ with the vector $\mathbf{\Pi} := \mathbb{I}\omega_B \in \mathbb{R}^3$. In Euler angles this equals

$$\mathbf{\Pi} = \begin{bmatrix} I_1(\dot{\varphi} \sin \psi \sin \theta + \dot{\theta} \cos \psi) \\ I_2(\dot{\varphi} \cos \psi \sin \theta - \dot{\theta} \sin \psi) \\ I_3(\dot{\varphi} \cos \theta + \dot{\psi}) \end{bmatrix}. \quad (1.35)$$

Using (1.33) and (1.34), and noting that $\omega_B = \mathbb{I}^{-1}\mathbf{\Pi}$, the expression of the kinetic energy on the dual of $\mathfrak{so}(3)^*$ identified with $\mathbb{R}^3$, is

$$K(\mathbf{\Pi}) = \frac{1}{2}\mathbf{\Pi} \cdot \mathbb{I}^{-1}\mathbf{\Pi} = \frac{1}{2} \left(\frac{\Pi_1^2}{I_1} + \frac{\Pi_2^2}{I_2} + \frac{\Pi_3^2}{I_3} \right). \quad (1.36)$$

The kinetic energy on $\mathbb{R}^3$ given by (1.33) can be expressed as a function on $\mathfrak{so}(3)$ using (1.17), namely

$$\begin{aligned} K(\omega_B) &= \frac{1}{2}\omega_B \cdot \mathbb{I}\omega_B = -\frac{1}{4} \text{trace}(\hat{\omega}_B(\mathbb{I}\omega_B)^\wedge) \\ &= -\frac{1}{4} \text{trace}(\hat{\omega}_B(\hat{\omega}_B J + J\hat{\omega}_B)), \end{aligned} \quad (1.37)$$

where J is a diagonal matrix whose entries are given by the relations $I_1 = J_2 + J_3$, $I_2 = J_3 + J_1$, and $I_3 = J_1 + J_2$, that is, $J_1 = (-I_1 + I_2 + I_3)/2$, $J_2 = (I_1 - I_2 + I_3)/2$, and $J_3 = (I_1 + I_2 - I_3)/2$. The last equality in (1.37) follows from the identity $(\mathbb{I}\omega_B)^\wedge = \hat{\omega}_B J + J\hat{\omega}_B$, proved by a direct verification. Formulas (1.37) and (1.30) immediately yield the expression of the kinetic energy on the tangent bundle $TSO(3)$:

$$K(A, \dot{A}) = -\frac{1}{4} \text{trace}((JA^{-1}\dot{A} + A^{-1}\dot{A}J)A^{-1}\dot{A}). \quad (1.38)$$

Since left translation of $SO(3)$ on itself lifts to the left action $B \cdot (A, \dot{A}) := (BA, B\dot{A})$ on $TSO(3)$, the expression (1.38) of $K(A, \dot{A})$ immediately implies that K is invariant relative to this action. Thus, *the kinetic energy of the heavy top is left invariant*.

Left translating the inner product $\langle\langle \cdot, \cdot \rangle\rangle$ from the tangent space to the identity

to the tangent space at an arbitrary point of $SO(3)$, defines a left invariant Riemannian metric on $SO(3)$ whose kinetic energy is (1.38). Relative to this metric, the Legendre transformation gives the canonically conjugate variables

$$p_\varphi := \frac{\partial K}{\partial \dot{\varphi}}, \quad p_\psi := \frac{\partial K}{\partial \dot{\psi}}, \quad p_\theta := \frac{\partial K}{\partial \dot{\theta}}.$$

We shall summarize at the end of this discussion various formulas in terms of the Euler angles, including this one. Expressing now the kinetic energy in the variables $(\varphi, \psi, \theta, p_\varphi, p_\psi, p_\theta)$ will give thus a left invariant function on $T^*SO(3)$.

Next we turn to the expression of the potential energy. It is given by the height of the center of mass over the horizontal plane perpendicular to the direction of gravity. Let ℓ denote the length of the segment between the fixed point and the center of mass and let χ be the unit vector supported by this line segment. Let $M = \int_B \rho_0(\mathbf{X}) d^3\mathbf{X}$ be the total mass of the top, g the value of the gravitational acceleration, and $\mathbf{k}$ the spatial unit vector pointing in opposite direction to gravity. Then the potential energy at time t equals

$$V(t) = Mg\mathbf{k} \cdot A\ell\chi = Mg\ell\mathbf{k} \cdot \boldsymbol{\lambda} = Mg\ell\boldsymbol{\Gamma} \cdot \chi \quad (1.39)$$

where $\boldsymbol{\lambda} := A\chi$ and $\boldsymbol{\Gamma} := A^{-1}\mathbf{k}$. The three expressions represent the potential energy in material, spatial, and body representation, respectively. It is clear that the potential energy is invariant only with respect to rotations about the axis of gravity, which shows that the total energy

$$H(A, \dot{A}) = -\frac{1}{4} \text{trace}((JA^{-1}\dot{A} + A^{-1}\dot{A}J)A^{-1}\dot{A}) + Mg\ell\mathbf{k} \cdot A\chi \quad (1.40)$$

is also invariant only under this circle subgroup of $SO(3)$. In Euler angles and their conjugate momenta, (1.40) becomes

$$\begin{aligned} H = \frac{1}{2} & \left[\frac{[(p_\varphi - p_\psi \cos \theta) \sin \psi + p_\theta \sin \theta \cos \psi]^2}{I_1 \sin^2 \theta} \right. \\ & + \left. \frac{[(p_\varphi - p_\psi \cos \theta) \cos \psi - p_\theta \sin \theta \sin \psi]^2}{I_2 \sin^2 \theta} + \frac{p_\psi^2}{I_3} \right] \\ & + Mg\ell \cos \theta, \end{aligned} \quad (1.41)$$

where, without loss of generality, we assumed that χ in body coordinates is equal to $(0, 0, 1)$.

Since $\dot{p}_\varphi = -\partial H / \partial \varphi = 0$, it follows that $p_\varphi = \boldsymbol{\Pi} \cdot \boldsymbol{\Gamma}$ is conserved.

On $\mathbb{R}^3 \times \mathbb{R}^3$ the expression of the total energy is hence

$$H(\boldsymbol{\Pi}, \boldsymbol{\Gamma}) = \frac{1}{2} \boldsymbol{\Pi} \cdot \mathbb{I}^{-1} \boldsymbol{\Pi} + Mg\ell \boldsymbol{\Gamma} \cdot \chi. \quad (1.42)$$

In addition to the conservation of $\mathbf{\Pi} \cdot \mathbf{\Gamma}$, we also have $\|\mathbf{\Gamma}\| = 1$. The significance of these two conserved quantities in body representation will become clear only after understanding the Poisson geometry underlying the motion given by (1.42).

For completeness we summarize in Table 1.1 the relationship between the variables introduced till now.

$\begin{aligned}\Pi_1 &= [(p_\varphi - p_\psi \cos \theta) \sin \psi + p_\theta \sin \theta \cos \psi] / \sin \theta \\ &= I_1(\dot{\varphi} \sin \theta \sin \psi + \dot{\theta} \cos \psi) \\ \Pi_2 &= [(p_\varphi - p_\psi \cos \theta) \cos \psi - p_\theta \sin \theta \sin \psi] / \sin \theta \\ &= I_2(\dot{\varphi} \sin \theta \cos \psi - \dot{\theta} \sin \psi) \\ \Pi_3 &= p_\psi = I_3(\dot{\varphi} \cos \theta + \dot{\psi}) \\ \Gamma_1 &= \sin \theta \sin \psi \\ \Gamma_2 &= \sin \theta \cos \psi \\ \Gamma_3 &= \cos \theta \\ p_\varphi &= \mathbf{\Pi} \cdot \mathbf{\Gamma} \\ &= I_1(\dot{\varphi} \sin \theta \sin \psi + \dot{\theta} \cos \psi) \sin \theta \sin \psi \\ &\quad + I_2(\dot{\varphi} \sin \theta \cos \psi - \dot{\theta} \sin \psi) \sin \theta \cos \psi \\ &\quad + I_3(\dot{\varphi} \cos \theta + \dot{\psi}) \cos \theta \\ p_\psi &= \Pi_3 = I_3(\dot{\varphi} \cos \theta + \dot{\psi}) \\ p_\theta &= (\Gamma_2 \Pi_1 - \Gamma_1 \Pi_2) / \sqrt{1 - \Gamma_3^2} \\ &= I_1(\dot{\varphi} \sin \theta \sin \psi + \dot{\theta} \cos \psi) \cos \psi \\ &\quad - I_2(\dot{\varphi} \sin \theta \cos \psi - \dot{\theta} \sin \psi) \sin \psi \\ \dot{\varphi} &= \frac{1}{I_1} \frac{\Pi_1 \Gamma_1}{1 - \Gamma_3^2} + \frac{1}{I_2} \frac{\Pi_2 \Gamma_2}{1 - \Gamma_3^2} \\ \dot{\psi} &= \frac{\Pi_3}{I_3} - \frac{\Pi_1 \Gamma_1 \Gamma_3}{I_1(1 - \Gamma_3^2)} - \frac{\Pi_2 \Gamma_2 \Gamma_3}{I_2(1 - \Gamma_3^2)} \\ \dot{\theta} &= \frac{\Pi_1 \Gamma_2}{I_1 \sqrt{1 - \Gamma_3^2}} - \frac{\Pi_2 \Gamma_1}{I_2 \sqrt{1 - \Gamma_3^2}}\end{aligned}$

Table 1.1. *Summary of the variables for the heavy top*

The equations of motion of the heavy top. In a chart on $T^*SO(3)$ given by the Euler angles and their conjugate momenta, the equations of motion are

$$\begin{aligned}\dot{\varphi} &= \frac{\partial H}{\partial p_\varphi}, & \dot{\psi} &= \frac{\partial H}{\partial p_\psi}, & \dot{\theta} &= \frac{\partial H}{\partial p_\theta} \\ \dot{p}_\varphi &= -\frac{\partial H}{\partial \varphi}, & \dot{p}_\psi &= -\frac{\partial H}{\partial \psi}, & \dot{p}_\theta &= -\frac{\partial H}{\partial \theta}\end{aligned}$$

with H given by (1.41).

Consider now the map

$$\mathbf{J} : (\varphi, \psi, \theta, p_\varphi, p_\psi, p_\theta) \mapsto (\mathbf{\Pi}, \mathbf{\Gamma}) \quad (1.43)$$

given by the formulas above. This is *not* a change of variables because $\|\mathbf{\Gamma}\| = 1$. A lengthy direct computation, using the formulas above, shows that these equations imply the **Euler-Poisson equations**

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbf{\Omega} + Mg\ell\mathbf{\Gamma} \times \boldsymbol{\chi}, \quad \dot{\mathbf{\Gamma}} = \mathbf{\Gamma} \times \mathbf{\Omega} \quad (1.44)$$

where $\mathbf{\Omega} := \omega_B = \mathbb{I}^{-1}\mathbf{\Pi}$.

These equations can be obtained in two ways.

(i) The canonical Poisson bracket of two functions $f, h : T^*SO(3) \rightarrow \mathbb{R}$ in a chart given by the Euler angles and their conjugate momenta is

$$\{f, h\} = \frac{\partial f}{\partial \varphi} \frac{\partial h}{\partial p_\varphi} - \frac{\partial f}{\partial p_\varphi} \frac{\partial h}{\partial \varphi} + \frac{\partial f}{\partial \psi} \frac{\partial h}{\partial p_\psi} - \frac{\partial f}{\partial p_\psi} \frac{\partial h}{\partial \psi} + \frac{\partial f}{\partial \theta} \frac{\partial h}{\partial p_\theta} - \frac{\partial f}{\partial p_\theta} \frac{\partial h}{\partial \theta}.$$

A direct long computation shows that if $F, H : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$, then

$$\{F \circ \mathbf{J}, H \circ \mathbf{J}\} = \{F, H\}_- \circ \mathbf{J},$$

where $\mathbf{J}$ is given by (1.43) and

$$\begin{aligned} \{F, H\}_-(\mathbf{\Pi}, \mathbf{\Gamma}) &= -\mathbf{\Pi} \cdot (\nabla_{\mathbf{\Pi}} F \times \nabla_{\mathbf{\Pi}} H) \\ &\quad - \mathbf{\Gamma} \cdot (\nabla_{\mathbf{\Pi}} F \times \nabla_{\mathbf{\Gamma}} H + \nabla_{\mathbf{\Gamma}} F \times \nabla_{\mathbf{\Pi}} H); \end{aligned} \quad (1.45)$$

$\nabla_{\mathbf{\Pi}} F$ and $\nabla_{\mathbf{\Gamma}} F$ denote the partial gradients relative to the variables $\mathbf{\Pi}$ and $\mathbf{\Gamma}$ respectively. An additional long computation shows that this defines a Poisson bracket, that is, it is bilinear, skew symmetric, and satisfies both the Jacobi and the Leibniz identities. Finally, if H is given by (1.42), it is easy to see that the equation $\dot{F} = \{F, H\}$ for any $F : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ is equivalent to the Euler-Poisson equations (1.44).

Note that the bracket of any function with an arbitrary function of $\|\mathbf{\Gamma}\|^2$ and $\mathbf{\Pi} \cdot \mathbf{\Gamma}$ is zero. The functions $\|\mathbf{\Gamma}\|^2$ and $\mathbf{\Pi} \cdot \mathbf{\Gamma}$ are the **Casimir functions** of the bracket (1.45).

(ii) Equations (1.44) can also be obtained from a variational principle. Given is the Lagrangian

$$L(\mathbf{\Omega}, \mathbf{\Gamma}) := \frac{1}{2} \mathbb{I} \mathbf{\Omega} \cdot \mathbf{\Omega} - Mg\ell \mathbf{\Gamma} \cdot \boldsymbol{\chi} \quad (1.46)$$

and the second Euler-Poisson equation $\dot{\Gamma} = \Gamma \times \Omega$ whose solution is $\Gamma(t) = A(t)^{-1}\mathbf{k}$, where $\Omega(t) = A(t)^{-1}\dot{A}(t)$. Consider the variational principle for L

$$\delta \int_a^b L(\Omega, \Gamma) dt = 0$$

but only subject to the restricted variations of the form

$$\delta\Omega := \dot{\Sigma} + \Omega \times \Sigma \quad \delta\Gamma := \Gamma \times \Sigma \quad (1.47)$$

where $\Sigma(t)$ is an arbitrary curve vanishing at the endpoints a and b , i.e

$$\Sigma(a) = \Sigma(b) = 0. \quad (1.48)$$

Using integration by parts together with the vanishing conditions at the endpoints, $\nabla_{\Omega} L(\Omega, \Gamma) = \mathbb{I}\Omega = \Pi$, and $\nabla_{\Gamma} L(\Omega, \Gamma) = -Mg\ell\chi$, we get

$$\begin{aligned} 0 &= \delta \int_a^b L(\Omega, \Gamma) dt = \int_a^b \nabla_{\Omega} L(\Omega, \Gamma) \cdot \delta\Omega dt + \int_a^b \nabla_{\Gamma} L(\Omega, \Gamma) \cdot \delta\Gamma dt \\ &= \int_a^b \Pi \cdot \delta\Omega dt - Mg\ell \int_a^b \chi \cdot \delta\Gamma dt \\ &= \int_a^b \Pi \cdot (\dot{\Sigma} + \Omega \times \Sigma) dt - Mg\ell \int_a^b \chi \cdot (\Gamma \times \Sigma) dt \\ &= - \int_a^b \dot{\Pi} \cdot \Sigma dt + \int_a^b \Pi \cdot (\Omega \times \Sigma) dt - Mg\ell \int_a^b \Sigma \cdot (\chi \times \Gamma) dt \\ &= \int_a^b \left(-\dot{\Pi} + \Pi \times \Omega + Mg\ell \Gamma \times \chi \right) \cdot \Sigma dt. \end{aligned}$$

The arbitrariness of Σ yields the first of the Euler-Poisson equations (1.44).

Remark. If $g\ell = 0$, then the heavy top becomes a free rigid body. In this case there is only one equation, namely the Euler equation $\dot{\Pi} = \Pi \times \Omega$, the Poisson bracket of two smooth functions $F, H : \mathbb{R}^3 \rightarrow \mathbb{R}$ is given by $\{F, H\}(\Pi) = -\Pi \cdot (\nabla F \times \nabla H)$, and the variational principle is $\delta \int_a^b L(\Omega) dt = 0$ for variations $\delta\Omega = \dot{\Sigma} + \Omega \times \Sigma$ where $\Sigma(a) = \Sigma(b) = 0$. The Poisson bracket of any function with an arbitrary smooth function of $\|\Pi\|^2$ vanishes and thus the motion of the rigid body takes place on Π -spheres of constant radius. For the free rigid body, both the Lagrangian and the total energy coincide with the kinetic energy $L(\Omega) = H(\Pi) = \Pi \cdot \Omega/2$, where $\Pi = \mathbb{I}\Omega$, and thus the solutions of the free rigid body motion are geodesics on $SO(3)$ relative to the left invariant metric whose value at the identity is (1.34). The solutions of the Euler equation $\dot{\Pi} = \Pi \times \Omega$ are therefore obtained by intersecting concentric spheres $\{\Pi \mid \|\Pi\| = R\}$ with the family of ellipsoids $\{\Pi \mid \Pi \cdot \mathbb{I}^{-1}\Pi = C\}$ for any constants $R, C \geq 0$. In this way one immediately sees that there are

six equilibria, four of them stable and two of them unstable. The stable ones correspond to rotations about the short and long axes of the moment of inertia and the unstable one corresponds to rotations about the middle axis.

One of the goals of these lectures is to explain the geometry behind all these phenomena that have appeared here as computational accidents. As we shall see, none of them are arbitrary occurrences and they all have a symplectic geometrical underpinning.

2 Hamiltonian Formalism

In this lecture we recall the fundamental concepts on symplectic manifolds and canonical transformations, as found, for example in, [Arnold79], [AbMa78], [MaRa94], [LiMa87], or [McDSal95].

2.1 Symplectic Manifolds

Definition 2.1 A *symplectic manifold* is a pair (P, Ω) where P is a Banach manifold and Ω is a closed (weakly) nondegenerate two-form on P . If Ω is strongly nondegenerate, (P, Ω) is called a *strong symplectic manifold*.

Recall that Ω is *weakly*, respectively *strongly, nondegenerate* if the smooth vector bundle map covering the identity $\flat : TP \rightarrow T^*P$ given by $v \mapsto v^\flat := \Omega(v, \cdot)$ is injective, respectively bijective, on each fiber. If P is finite dimensional, there is no distinction between these concepts and nondegeneracy implies that P is even dimensional.

If the manifold is a Banach space V and the two-form is constant on V , then (V, Ω) is called a *symplectic Banach space*. For example, $\mathbb{R}^{2n}$ endowed with the canonical symplectic structure

$$\Omega((\mathbf{u}, \mathbf{v}), (\mathbf{u}', \mathbf{v}')) = \mathbf{u} \cdot \mathbf{v}' - \mathbf{u}' \cdot \mathbf{v},$$

for $\mathbf{u}, \mathbf{v}, \mathbf{u}', \mathbf{v}' \in \mathbb{R}^n$, is a symplectic finite dimensional vector space. Another standard example is a complex Hilbert space with the symplectic form given by the negative of the imaginary part of the Hermitian inner product. For example, $\mathbb{C}^n$ has Hermitian inner product given by $\mathbf{z} \cdot \mathbf{w} := \sum_{j=1}^n z_j \overline{w_j}$, where $\mathbf{z} = (z_1, \dots, z_n), \mathbf{w} = (w_1, \dots, w_n) \in \mathbb{C}^n$. The symplectic form is thus given by $\Omega(\mathbf{z}, \mathbf{w}) := -\text{Im}(\mathbf{z} \cdot \mathbf{w})$ and it is identical to the one given before on $\mathbb{R}^{2n}$ by identifying $\mathbf{z} = \mathbf{u} + i\mathbf{v} \in \mathbb{C}^n$ with $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^{2n}$ and $\mathbf{w} = \mathbf{u}' + i\mathbf{v}' \in \mathbb{C}^n$ with $(\mathbf{u}', \mathbf{v}') \in \mathbb{R}^{2n}$.

The local structure of strong symplectic manifolds is given by the following basic result.

Theorem 2.2 (Darboux) *If (P, Ω) is a strong symplectic manifold then each point of P admits a chart in which Ω is constant. If P is finite dimensional then around each point there are coordinates $(q^1, \dots, q^n, p_1, \dots, p_n)$, where $2n = \dim P$, such that $\Omega = \mathbf{d}q^i \wedge \mathbf{d}p_i$ (with the usual summation convention).*

The Darboux theorem for $2n$ -dimensional symplectic manifolds hence states that, locally, the two-form Ω is given by the matrix

$$\mathbb{J} = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{bmatrix}$$

where $\mathbf{I}$ denotes the $n \times n$ identity matrix, that is, locally the symplectic form is the one given above on $\mathbb{R}^{2n}$.

Remark 2.3 (i) The Darboux theorem is false for weak symplectic manifolds. For a counterexample, see Exercise 5.1-3 in [MaRa94].

(ii) There is a relative version of the Darboux theorem, namely, if S is a submanifold of P and Ω_0, Ω_1 are two strong symplectic forms on P that coincide when evaluated at points of S , then there is an open neighborhood V of S and a diffeomorphism $\varphi : V \rightarrow \varphi(V)$ such that $\varphi|_S = \text{id}$ and $\varphi^*\Omega_1 = \Omega_0$.

(iii) There is a generalization of the Darboux theorem for G -equivariant forms if G is a compact Lie group. It is still true that every point admits a G -invariant neighborhood on which the symplectic form is constant. However, the constant symplectic forms are no longer equivalent under G -equivariant changes of coordinates and the number of inequivalent symplectic forms depends on the representation type of the compact Lie group G , i.e., if it is real, complex, or quaternionic; see [DeMe93] and [MD93] for details.

Denote by $\mathfrak{X}(P)$ the set of vector fields on P and by $\Omega^k(P)$ the set of k -forms on P . The map $\flat$ induces a similar map between $\mathfrak{X}(P)$ and $\Omega^1(P)$, namely

$$\mathfrak{X}(P) \ni X \mapsto \mathbf{i}_X \Omega := \Omega(X, \cdot) =: X^\flat \in \Omega^1(P).$$

If Ω is strongly nondegenerate, this map is an isomorphism, so for any smooth map $H : P \rightarrow \mathbb{R}$ there exists a vector field X_H , called the **Hamiltonian vector field** defined by H , such that

$$\mathbf{i}_{X_H} \Omega = \mathbf{d}H,$$

where $\mathbf{d} : \Omega^k(P) \rightarrow \Omega^{k+1}(P)$ denotes the exterior differential. If P is only weakly nondegenerate, the uniqueness but not the existence of the Hamiltonian vector field is guaranteed. So for weak symplectic manifolds, not every

function defines a Hamiltonian vector field. See [ChMa74] for an approach to infinite dimensional Hamiltonian systems.

If P is $2n$ -dimensional, the local expression of X_H in Darboux coordinates is

$$X_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p_i}$$

and **Hamilton's equations** $\dot{z} = X_H(z)$ are given by

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i} \quad i = 1, \dots, n.$$

The two-form Ω induces the volume **Liouville form**

$$\Lambda := \frac{(-1)^{n(n-1)/2}}{n!} \underbrace{(\Omega \wedge \dots \wedge \Omega)}_{n \text{ times}}, \quad (2.1)$$

where $2n = \dim P$, which has the local expression

$$\Lambda = \mathbf{d}q^1 \wedge \dots \wedge \mathbf{d}q^n \wedge \mathbf{d}p_1 \wedge \dots \wedge \mathbf{d}p_n.$$

In particular, any symplectic manifold is oriented.

2.2 Symplectic Transformations

Given two symplectic manifolds (P_1, Ω_1) , (P_2, Ω_2) , a C^∞ mapping

$$\phi : (P_1, \Omega_1) \rightarrow (P_2, \Omega_2)$$

is said to be a **symplectic transformation** if $\phi^*\Omega_2 = \Omega_1$, that is,

$$\Omega_2(\phi(z))(T_z\phi(v_1), T_z\phi(v_2)) = \Omega_1(z)(v_1, v_2)$$

for all $z \in P_1$ and $v_1, v_2 \in T_zP_1$, where $T_z\phi : T_zP_1 \rightarrow T_zP_2$ is the tangent map of ϕ at $z \in P_1$. Any symplectic map has injective derivative due to the weak nondegeneracy of the symplectic forms.

Proposition 2.4 *Let (P, Ω) be a symplectic manifold. The flow ϕ_t of $X \in \mathfrak{X}(P)$ consists of symplectic transformations if and only if it is **locally Hamiltonian**, that is, $\mathcal{L}_X\Omega = 0$, where $\mathcal{L}_X$ denotes the Lie derivative operator defined by X .*

Remark 2.5 There is a general formula relating the Lie derivative and the

dynamics of a vector field. If $\mathcal{T}$ is an arbitrary (time independent) tensor field on P and F_t the flow of the (time independent) vector field X , then

$$\frac{d}{dt}F_t^*\mathcal{T} = F_t^*\mathcal{L}_X\mathcal{T} \quad (2.2)$$

where $\mathcal{L}_X\mathcal{T}$ denotes the Lie derivative of $\mathcal{T}$ along X . We shall use this formula in the proof below.

Proof Note that

$$\phi_t^*(\Omega) = \Omega \iff \frac{d}{dt}\phi_t^*(\Omega) = 0,$$

which by (2.2) is equivalent to

$$\phi_t^*\mathcal{L}_X\Omega = 0 \iff \mathcal{L}_X\Omega = 0$$

since ϕ_t^* is an isomorphism. ■

The name “locally Hamiltonian vector field” is justified by the following fact. Since $\mathcal{L}_X\Omega = \mathbf{i}_X\mathbf{d}\Omega + \mathbf{d}\mathbf{i}_X\Omega$ and $\mathbf{d}\Omega = 0$, the condition $\mathcal{L}_X\Omega = 0$ is equivalent to $\mathbf{d}\mathbf{i}_X\Omega = 0$, that is, $\mathbf{i}_X\Omega$ is closed. The Poincaré lemma implies the existence of a *local* function H such that $\mathbf{i}_X\Omega = \mathbf{d}H$.

Proposition 2.6 (Energy conservation) *Let ϕ_t be the flow of the Hamiltonian vector field X_H . Then $H \circ \phi_t = H$.*

Proof Since

$$\frac{d}{dt}\phi_t(z) = X_H(\phi_t(z))$$

and

$$\Omega(z)(X_H(z), u) = \langle \mathbf{d}H(z), u \rangle \quad \text{for all } z \in P \text{ and } u \in T_zP,$$

where $\langle \cdot, \cdot \rangle$ is the pairing between one-forms and vectors, we have by the chain rule

$$\begin{aligned} \frac{d}{dt}H(\phi_t(z)) &= \left\langle \mathbf{d}H(\phi_t(z)), \frac{d\phi_t(z)}{dt} \right\rangle = \langle \mathbf{d}H(\phi_t(z)), X_H(\phi_t(z)) \rangle \\ &= \Omega(\phi_t(z))(X_H(\phi_t(z)), X_H(\phi_t(z))) = 0, \end{aligned}$$

which means that $H \circ \phi_t$ is constant relative to t . Since $\phi_0(z) = z$ for any $z \in P$, the result follows. ■

Proposition 2.7 *A diffeomorphism $\phi : P_1 \rightarrow P_2$ between strong symplectic manifolds is symplectic if and only if $\phi^* X_H = X_{H \circ \phi}$ for any $H : U \rightarrow \mathbb{R}$, where U is an arbitrary open subset of P_2 .*

Proof For a strong symplectic manifold (P, Ω) , the tangent space $T_z P$ equals the collection of all vectors of the form $X_K(z)$, for all smooth functions $K : U \rightarrow \mathbb{R}$ and all open neighborhoods U of z .

With this observation in mind and using the fact that ϕ is a diffeomorphism and the symplectic form Ω_2 strong, we have for all V open in P_1 , $z \in V$, $v \in T_z P_1$, and all $H : \phi(V) \rightarrow \mathbb{R}$ the following equivalence

$$\begin{aligned} \phi^* \Omega_2 = \Omega_1 &\iff \Omega_2(\phi(z)) (T_z \phi(X_{H \circ \phi}(z)), T_z \phi(v)) \\ &= \Omega_1(z) (X_{H \circ \phi}(z), v) \end{aligned} \quad (2.3)$$

However,

$$\begin{aligned} \Omega_1(z) (X_{H \circ \phi}(z), v) &= \langle \mathbf{d}(H \circ \phi)(z), v \rangle \\ &= \langle \mathbf{d}H(\phi(z)), T_z \phi(v) \rangle \\ &= \Omega_2(\phi(z)) (X_H(\phi(z)), T_z \phi(v)). \end{aligned} \quad (2.4)$$

By (2.3) and (2.4) the relation $\phi^* \Omega_2 = \Omega_1$ is thus equivalent to

$$\Omega_2(\phi(z)) (T_z \phi(X_{H \circ \phi}(z)), T_z \phi(v)) = \Omega_2(\phi(z)) (X_H(\phi(z)), T_z \phi(v))$$

for any smooth function $H : U \rightarrow \mathbb{R}$, $U \subset P_2$ open, $\phi(z) \in U$, and $v \in T_z P_1$. By nondegeneracy of Ω_2 this is equivalent to

$$T_z \phi \circ X_{H \circ \phi} = X_H \circ \phi \iff \phi^* X_H = X_{H \circ \phi}$$

which proves the statement. ■

2.3 Poisson Brackets

Given a strong symplectic manifold (P, Ω) , define the **Poisson bracket** of two smooth functions $F, G : P \rightarrow \mathbb{R}$ by

$$\{F, G\} := \Omega(X_F, X_G). \quad (2.5)$$

In this section all symplectic manifolds are assumed to be strong. As a consequence of Proposition 2.7 and of the Poisson bracket definition we have the following.

Proposition 2.8 A diffeomorphism $\phi : P_1 \rightarrow P_2$ is symplectic if and only if

$$\{F, G\}_2 \circ \phi = \{F \circ \phi, G \circ \phi\}_1$$

for all smooth functions $F, G : U \rightarrow \mathbb{R}$, where U is any open subset of P_2 .

Combined with Proposition 2.4 this yields the following statement.

Corollary 2.9 The flow ϕ_t is that of a locally Hamiltonian vector field if and only if

$$\phi_t^* \{F, G\} = \{\phi_t^* F, \phi_t^* G\}$$

for all smooth functions $F, G : U \rightarrow \mathbb{R}$, where U is an arbitrary open subset of P_2 .

Proposition 2.10 The vector space of smooth functions on the strong symplectic manifold (P, Ω) is a Lie algebra under the Poisson bracket.

Proof The definition of $\{F, G\}$ immediately implies that it is bilinear, skew-symmetric, and it satisfies the Leibniz identity in each factor. To prove the Jacobi identity, note that

$$\{F, G\} = \langle \mathbf{i}_{X_F} \Omega, X_G \rangle = \langle \mathbf{d}F, X_G \rangle = X_G[F]$$

and so

$$\{\{F, G\}, H\} = X_H[\{F, G\}].$$

Thus the Jacobi identity is equivalent to

$$X_H[\{F, G\}] = \{X_H[F], G\} + \{F, X_H[G]\} \quad (2.6)$$

which is obtained by differentiating in t at $t = 0$ the identity

$$\phi_t^* \{F, G\} = \{\phi_t^* F, \phi_t^* G\},$$

where ϕ_t is the flow of X_H (see Corollary 2.9). Indeed,

$$\begin{aligned} X_H[\{F, G\}] &= \left. \frac{d}{dt} \right|_{t=0} \phi_t^* \{F, G\} = \left. \frac{d}{dt} \right|_{t=0} \{\phi_t^* F, \phi_t^* G\} \\ &= \left. \frac{d}{dt} \right|_{t=0} \Omega(X_{\phi_t^* F}, X_{\phi_t^* G}) \\ &= \Omega(X_{X_H[F]}, X_G) + \Omega(X_F, X_{X_H[G]}) \\ &= \{X_H[F], G\} + \{F, X_H[G]\} \end{aligned}$$

which proves the derivation identity (2.6). ■

Corollary 2.11 *The set of Hamiltonian vector fields on P is a Lie subalgebra of the set of the vector fields on P because*

$$[X_F, X_G] = X_{-\{F, G\}}.$$

Proof As derivations,

$$\begin{aligned} [X_F, X_G][H] &= X_F[X_G[H]] - X_G[X_F[H]] = X_F[\{H, G\}] - X_G[\{H, F\}] \\ &= \{\{H, G\}, F\} - \{\{H, F\}, G\} = -\{H, \{F, G\}\} \\ &= -X_{\{F, G\}}[H], \end{aligned}$$

where we have applied the Jacobi identity in the fourth equality. $\blacksquare$

The next corollary gives Hamilton's equations in Poisson bracket form.

Corollary 2.12 *If ϕ_t is the flow of X_H and $F : U \rightarrow \mathbb{R}$ is an arbitrary smooth function defined on the open subset U of P then*

$$\frac{d}{dt}(F \circ \phi_t) = \{F \circ \phi_t, H\} = \{F, H\} \circ \phi_t. \quad (2.7)$$

Proof We have

$$\begin{aligned} \frac{d}{dt}(F \circ \phi_t)(z) &= \left\langle \mathbf{d}F(\phi_t(z)), \frac{d\phi_t(z)}{dt} \right\rangle = \langle \mathbf{d}F(\phi_t(z)), X_H(\phi_t(z)) \rangle \\ &= \{F, H\}(\phi_t(z)). \end{aligned}$$

As ϕ_t is symplectic, this is just $\{F \circ \phi_t, H \circ \phi_t\}(z)$. Conservation of energy gives $\{F \circ \phi_t, H \circ \phi_t\}(z) = \{F \circ \phi_t, H\}(z)$ which proves (2.7). $\blacksquare$

Equation (2.7) is often written in the compact form

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

Corollary 2.13 *The smooth function F is a constant of motion for X_H if and only if $\{F, H\} = 0$.*

2.4 Cotangent Bundles

In mechanics, the phase space is very often the cotangent bundle T^*Q of a configuration space Q . A cotangent bundle has a weak canonical symplectic structure and so any cotangent bundle is symplectic. The *canonical two-form* Ω on T^*Q is constructed from the *canonical one-form* Θ in the following way. Define $\Omega = -\mathbf{d}\Theta$ where Θ is the one-form given by

$$\Theta_{\alpha_q}(w_{\alpha_q}) := \langle \alpha_q, T_{\alpha_q} \pi(w_{\alpha_q}) \rangle, \quad (2.8)$$

where $\alpha_q \in T_q^*Q$, $w_{\alpha_q} \in T_{\alpha_q}(T^*Q)$, $\pi : T^*Q \rightarrow Q$ is the projection, and $T\pi : T(T^*Q) \rightarrow TQ$ is the tangent map of π . The structure is called “canonical” because all the elements occurring in the definition are naturally associated to any cotangent bundle.

In local coordinates (q^i, p_i) on T^*Q the expressions of the canonical one- and two-forms are

$$\Theta = p_i dq^i, \quad \Omega = -d\Theta = dq^i \wedge dp_i.$$

Thus the natural cotangent bundle coordinates are Darboux coordinates for the canonical symplectic form.

Given a diffeomorphism $f : Q \rightarrow S$ define the **cotangent lift** of f as the map $T^*f : T^*S \rightarrow T^*Q$ given by

$$T^*f(\alpha_s)(v_{f^{-1}(s)}) := \alpha_s(Tf(v_{f^{-1}(s)})), \quad (2.9)$$

where $\alpha_s \in T_s^*S$ and $v_{f^{-1}(s)} \in T_{f^{-1}(s)}Q$.

Theorem 2.14 *The cotangent lift T^*f of a diffeomorphism $f : Q \rightarrow S$ preserves the canonical one-form Θ .*

Proof For simplicity, we give the proof in finite dimensions. Denote by $(q^1, \dots, q^n, p_1, \dots, p_n)$ and $(s^1, \dots, s^n, r_1, \dots, r_n)$ canonical coordinates on T^*Q and T^*S , respectively. If $(s^1, \dots, s^n) = f(q^1, \dots, q^n)$, the cotangent lift T^*f is given by

$$(s^1, \dots, s^n, r_1, \dots, r_n) \mapsto (q^1, \dots, q^n, p_1, \dots, p_n), \quad (2.10)$$

where

$$p_j = \frac{\partial s^i}{\partial q^j} r_i. \quad (2.11)$$

To see that the lift (2.10) of f preserves the one-form is only a matter of using the chain rule and (2.11), that is,

$$r_i ds^i = r_i \frac{\partial s^i}{\partial q^j} dq^j = p_j dq^j.$$

Therefore $(T^*f)^*\Theta_Q = \Theta_S$ where Θ_Q and Θ_S are the canonical one-forms on Q and S respectively. ■

The converse of theorem 2.14 is also true, that is, if $\varphi : T^*S \rightarrow T^*Q$ is a diffeomorphism such that $\varphi^*\Theta_Q = \Theta_S$ then there exists a diffeomorphism $f : Q \rightarrow S$ such that $\phi = T^*f$. The proof of this fact is more involved; see Proposition 6.3.2 in [MaRa94].

2.5 Magnetic Terms

If $A \in \Omega^1(Q)$ define the **fiber translation map** $t_A : T^*Q \rightarrow T^*Q$ by

$$t_A(\alpha_q) := \alpha_q + A(q) \quad \alpha_q \in T_q^*Q.$$

Proposition 2.15 *Let Θ be the canonical one-form on T^*Q and $t_A : T^*Q \rightarrow T^*Q$ the fiber translation by $A \in \Omega^1(Q)$. Then*

$$t_A^*(\Theta) = \Theta + \pi^*A$$

where $\pi : T^*Q \rightarrow Q$ is the canonical projection. Hence

$$t_A^*\Omega = \Omega - \pi^*\mathbf{d}A,$$

where $\Omega = -\mathbf{d}\Theta$ is the canonical symplectic form on T^*Q . Thus t_A is a symplectic transformation if and only if $\mathbf{d}A = 0$.

Proof We do the proof in finite dimensions. In local coordinates we have

$$t_A(q^i, p_j) = (q^i, p_j + A_j(q))$$

so

$$t_A^*\Theta = t_A^*(p_i \mathbf{d}q^i) = (p_i + A_i(q)) \mathbf{d}q^i = p_i \mathbf{d}q^i + A_i(q) \mathbf{d}q^i,$$

which is the coordinate expression of $\Theta + \pi^*A$. ■

Let Ω be the canonical two-form on T^*Q . If $B \in \Omega^2(Q)$ is closed (i.e. $\mathbf{d}B = 0$) then

$$\Omega_B := \Omega - \pi^*B$$

is a (weak) symplectic form on T^*Q . Indeed, the matrix form of Ω_B is

$$\begin{bmatrix} -B & \mathbf{I} \\ -\mathbf{I} & 0 \end{bmatrix}, \quad (2.12)$$

which shows that Ω_B is nondegenerate since (2.12) is nonsingular. The extra term $-B$ in the symplectic form Ω_B is called a **magnetic term**.

Proposition 2.16 *Let B and B' be closed two-forms on Q such that $B - B' = \mathbf{d}A$. Then the map*

$$t_A : (T^*Q, \Omega_B) \rightarrow (T^*Q, \Omega_{B'})$$

*is a symplectic diffeomorphism of (T^*Q, Ω_B) with $(T^*Q, \Omega_{B'})$.*

Proof As

$$t_A^* \Omega = \Omega - \pi^* \mathbf{d}A = \Omega - \pi^* B + \pi^* B'$$

and $\pi \circ t_A = \pi$ we have

$$t_A^* (\Omega - \pi^* B') = \Omega - \pi^* B$$

which proves the statement. ■

Example: A particle in a magnetic field. Consider the motion of a charged particle in a given time independent magnetic field $\mathbf{B} := B_x \mathbf{i} + B_y \mathbf{j} + B_z \mathbf{k}$, where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ is the usual orthonormal basis of $\mathbb{R}^3$. As customary in electromagnetism, we assume that $\operatorname{div} \mathbf{B} = 0$. The divergence free vector field $\mathbf{B}$ uniquely defines a closed two-form B on $\mathbb{R}^3$ by

$$B := \mathbf{i}_B(\mathbf{d}x \wedge \mathbf{d}y \wedge \mathbf{d}z) = B_x \mathbf{d}y \wedge \mathbf{d}z + B_y \mathbf{d}z \wedge \mathbf{d}x + B_z \mathbf{d}x \wedge \mathbf{d}y.$$

The Lorentz force law for a particle with charge e and mass m is

$$m \frac{d\mathbf{v}}{dt} = \frac{e}{c} \mathbf{v} \times \mathbf{B}, \quad (2.13)$$

where $\mathbf{q} = (x, y, z)$ is the position of the particle, $\mathbf{v} = (\dot{x}, \dot{y}, \dot{z})$ its velocity, e its charge, m its mass, and c the speed of light. What is the Hamiltonian formulation of these equations? As we shall see, there are two different answers.

First, let $(p_x, p_y, p_z) := \mathbf{p} := m\mathbf{v}$, consider on $(\mathbf{q}, \mathbf{p})$ -space, that is, $T^*\mathbb{R}^3$, the magnetic symplectic form

$$\begin{aligned} \Omega_B &= \mathbf{d}x \wedge \mathbf{d}p_x + \mathbf{d}y \wedge \mathbf{d}p_y + \mathbf{d}z \wedge \mathbf{d}p_z - \frac{e}{c} B \\ &= m(\mathbf{d}x \wedge \mathbf{d}\dot{x} + \mathbf{d}y \wedge \mathbf{d}\dot{y} + \mathbf{d}z \wedge \mathbf{d}\dot{z}) - \frac{e}{c} B, \end{aligned}$$

and the Hamiltonian given by the kinetic energy of the particle

$$H = \frac{1}{2m} \|\mathbf{p}\|^2 = \frac{m}{2} (\dot{x}^2 + \dot{z}^2 + \dot{y}^2).$$

Let us show that (2.13) is given by the Hamiltonian vector field X_H determined by the equation

$$\mathbf{d}H = \mathbf{i}_{X_H} \Omega_B. \quad (2.14)$$

If $X_H(\mathbf{q}, \mathbf{p}) = (u, v, w, \dot{u}, \dot{v}, \dot{w})$, then (2.14) is equivalent to

$$\begin{aligned} m(\dot{x}\mathbf{d}\dot{x} + \dot{y}\mathbf{d}\dot{y} + \dot{z}\mathbf{d}\dot{z}) &= m(u\mathbf{d}\dot{x} - \dot{u}\mathbf{d}x + v\mathbf{d}\dot{y} - \dot{v}\mathbf{d}y + w\mathbf{d}\dot{z} - \dot{w}\mathbf{d}z) \\ &\quad - \frac{e}{c} [B_x v \mathbf{d}z - B_x w \mathbf{d}y - B_y u \mathbf{d}z + B_y w \mathbf{d}x + B_z u \mathbf{d}y - B_z v \mathbf{d}x]. \end{aligned}$$

Comparing coefficients we get

$$\begin{aligned} u &= \dot{x} \\ v &= \dot{y} \\ w &= \dot{z} \\ m\dot{u} &= \frac{e}{c}(B_z v - B_y w) \\ m\dot{v} &= \frac{e}{c}(B_x w - B_z u) \\ m\dot{w} &= \frac{e}{c}(B_y u - B_x v), \end{aligned}$$

which are equivalent to (2.13). Thus Hamilton's equations for the Lorentz force law need to be taken relative to the magnetic symplectic form and the kinetic energy of the particle.

Second, write $B = \mathbf{d}A$ (always possible on $\mathbb{R}^3$), or equivalently, $\mathbf{B} = \nabla \times \mathbf{A}$, where $\mathbf{A}^\flat = A$, that is $\mathbf{A} = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}$ and $A = A_x dx + A_y dy + A_z dz$. Then the map $t_A : (\mathbf{x}, \mathbf{p}) \mapsto (\mathbf{x}, \mathbf{p} + \frac{e}{c} \mathbf{A})$, pulls back the canonical two-form Ω of $T^*\mathbb{R}^3$ to Ω_B by Proposition 2.15. So, the Lorentz equations are also Hamiltonian relative to the canonical two-form Ω and Hamilton's function

$$H_A(\mathbf{x}, \mathbf{p}) = \frac{1}{2m} \left\| \mathbf{p} - \frac{e}{c} \mathbf{A} \right\|^2.$$

(See also [MaRa94] and [Jost]). It is not always possible to write $B = \mathbf{d}A$ if the domain on which one considers these equations is not simply connected, the magnetic monopole being such a case. We shall return to this example in the context of the Lagrangian formalism in §3.3.

3 Lagrangian Formalism

This lecture describes the local geometry of the Lagrangian formalism and ends with some remarks on global Lagrangian variational principles. The references for this material continue to be [Arnold79], [AbMa78], [MaRa94], [LiMa87].

3.1 Lagrangian Systems

The Legendre transformation. A smooth function $L : TQ \rightarrow \mathbb{R}$ is called a **Lagrangian**. The **Legendre transformation** $\mathbb{F}L : TQ \rightarrow T^*Q$ is the smooth vector bundle map covering the identity defined by

$$\langle \mathbb{F}L(v_q), w_q \rangle := \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} L(v_q + \epsilon w_q).$$

In the finite dimensional case, the local expression of $\mathbb{F}L$ is

$$\mathbb{F}L(q^i, \dot{q}^i) = \left(q^i, \frac{\partial L}{\partial \dot{q}^i} \right) = (q^i, p_i).$$

Given a Lagrangian L , the **action** of L is the map $A : TQ \rightarrow \mathbb{R}$ given by

$$A(v) := \langle \mathbb{F}L(v), v \rangle, \quad (3.1)$$

and the **energy** of L is

$$E(v) := A(v) - L(v). \quad (3.2)$$

Let Θ and Ω be the canonical one and two-forms on T^*Q , respectively. The Legendre transformation $\mathbb{F}L$ induces a one-form Θ_L and a closed two-form Ω_L on TQ by

$$\begin{aligned} \Theta_L &= (\mathbb{F}L)^*\Theta \\ \Omega_L &= -\mathbf{d}\Theta_L = (\mathbb{F}L)^*\Omega. \end{aligned}$$

In finite dimensions, the local expressions are

$$\begin{aligned} \Theta_L &:= \frac{\partial L}{\partial \dot{q}^i} \mathbf{d}q^i \\ \Omega_L &:= \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \mathbf{d}q^i \wedge \mathbf{d}q^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \mathbf{d}q^i \wedge \mathbf{d}\dot{q}^j. \end{aligned}$$

The closed two-form Ω_L can be written as the $2n \times 2n$ skew-symmetric matrix

$$\Omega_L = \begin{pmatrix} \mathcal{A} & \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \\ -\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} & 0 \end{pmatrix} \quad (3.3)$$

where $\mathcal{A}$ is the skew-symmetrization of the $n \times n$ matrix $\left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \right)$. The non-degeneracy of Ω_L is equivalent to the invertibility of the matrix $\left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \right)$. If (3.3) is invertible, the Lagrangian L is said to be **regular**. In this case, by the implicit function theorem, $\mathbb{F}L$ is locally invertible. If $\mathbb{F}L$ is a diffeomorphism, L is called **hyperregular**.

Lagrangian systems. Lagrangian systems are special vector fields on the tangent bundle that are naturally associated to a function $L : TQ \rightarrow \mathbb{R}$.

Definition 3.1 A vector field Z on TQ is called a **Lagrangian vector field** if

$$\Omega_L(v)(Z(v), w) = \langle \mathbf{d}E(v), w \rangle,$$

for all $v \in T_q Q$, $w \in T_v(TQ)$.

Proposition 3.2 *The energy is conserved along the flow of a Lagrangian vector field Z .*

Proof Let $v(t) \in TQ$ be an integral curve of Z . Skew-symmetry of Ω_L implies

$$\begin{aligned} \frac{d}{dt}E(v(t)) &= \langle \mathbf{d}E(v(t)), \dot{v}(t) \rangle = \langle \mathbf{d}E(v(t)), Z(v(t)) \rangle \\ &= \Omega_L(v(t))(Z(v(t)), Z(v(t))) = 0. \end{aligned}$$

Thus $E(v(t))$ is constant in t . ■

Second order equations. Lagrangian vector fields are intimately tied to second order equations.

Definition 3.3 Let $\tau_Q : TQ \rightarrow Q$ be the canonical projection. A vector field $Z \in \mathfrak{X}(TQ)$ is called a **second order equation** if

$$T\tau_Q \circ Z = \text{id}_{TQ}.$$

If $c(t)$ is an integral curve of Z , the curve $\tau_Q(c(t))$ on Q is called a **base integral curve** of Z .

Let $\mathbf{V}$ be the model of Q . In a canonical tangent bundle chart $U \times \mathbf{V}$ of TQ , U open in $\mathbf{V}$, the vector field Z can be written as

$$Z(\mathbf{q}, \dot{\mathbf{q}}) = (\mathbf{q}, \dot{\mathbf{q}}, Z_1(\mathbf{q}, \dot{\mathbf{q}}), Z_2(\mathbf{q}, \dot{\mathbf{q}})),$$

for some smooth maps $Z_1, Z_2 : U \times \mathbf{V} \rightarrow \mathbf{V}$. Thus,

$$T\tau_Q(Z(\mathbf{q}, \dot{\mathbf{q}})) = (\mathbf{q}, Z_1(\mathbf{q}, \dot{\mathbf{q}})).$$

Since Z is a second order equation, this implies that $Z_1(\mathbf{q}, \dot{\mathbf{q}}) = \dot{\mathbf{q}}$. That is, in a local chart, the vector field is given by

$$Z(\mathbf{q}, \dot{\mathbf{q}}) = (\mathbf{q}, \dot{\mathbf{q}}, \dot{\mathbf{q}}, Z_2(\mathbf{q}, \dot{\mathbf{q}})).$$

So the dynamics is determined by

$$\begin{aligned} \frac{d\mathbf{q}}{dt} &= Z_1(\mathbf{q}, \dot{\mathbf{q}}) = \dot{\mathbf{q}} \\ \frac{d\dot{\mathbf{q}}}{dt} &= Z_2(\mathbf{q}, \dot{\mathbf{q}}) \end{aligned}$$

which is equivalent to the usual second order equation

$$\frac{d^2\mathbf{q}}{dt^2} = Z_2(\mathbf{q}, \dot{\mathbf{q}}).$$

The relationship between second order equations and Lagrangian vector fields is given by the following statement, whose proof is a straightforward (but somewhat lengthy) computation.

Theorem 3.4 *Let Z be the Lagrangian vector field on Q defined by L and suppose that Z is a second order equation. Then, in a canonical chart $U \times V$ of TQ , an integral curve $(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ of Z satisfies the **Euler-Lagrange equations***

$$\begin{aligned} \frac{dq^i}{dt} &= \dot{q}^i \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} &= \frac{\partial L}{\partial q^i}, \quad i = 1, \dots, n. \end{aligned}$$

Furthermore, if L is regular then Z is always second order.

3.2 Geodesics

An important example of a Lagrangian vector field is the geodesic spray of a pseudo-Riemannian metric. Recall that a **pseudo-Riemannian manifold** is a smooth manifold Q endowed with a symmetric nondegenerate covariant tensor g . Thus, on each tangent space $T_q Q$ there is a nondegenerate (but indefinite, in general) inner product $g(q)$. If g is positive definite, then the pair (Q, g) is called a **Riemannian manifold**.

If (Q, g) is a pseudo-Riemannian manifold, there is a natural Lagrangian on it given by the **kinetic energy** K of the metric g , namely,

$$K(v) := \frac{1}{2} g(q)(v_q, v_q),$$

for $q \in Q$ and $v_q \in T_q Q$. In finite dimensions, in a local chart,

$$K(q^k, \dot{q}^l) = \frac{1}{2} g_{ij}(q^k) \dot{q}^i \dot{q}^j.$$

The Legendre transformation is in this case $\mathbb{F}K(v_q) = g(q)(v_q, \cdot)$, for $v_q \in T_q Q$. The Euler-Lagrange equations become the **geodesic equations** for the metric g , given (for finite dimensional Q in a local chart) by

$$\ddot{q}^i + \Gamma_{jk}^i \dot{q}^j \dot{q}^k = 0, \quad i = 1, \dots, n,$$

where

$$\Gamma_{jk}^h = \frac{1}{2} g^{hl} \left(\frac{\partial g_{jl}}{\partial q^k} + \frac{\partial g_{kl}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^l} \right)$$

are the **Christoffel symbols**. The Lagrangian vector field associated to K is

called the **geodesic spray**. Since the Legendre transformation is a diffeomorphism (in finite dimensions or in infinite dimensions if the metric is assumed to be strong), the geodesic spray is always a second order equation.

Let us link this approach to geodesics to the classical formulation using covariant derivatives. The covariant derivative $\nabla : \mathfrak{X}(Q) \times \mathfrak{X}(Q) \rightarrow \mathfrak{X}(Q)$, $(X, Y) \mapsto \nabla_X(Y)$, of the Levi-Civita connection on (Q, g) is given in local charts by

$$\nabla_X(Y) = X^i Y^j \Gamma_{ij}^k \frac{\partial}{\partial q^k} + X^i \frac{\partial Y^j}{\partial q^i} \frac{\partial}{\partial q^j}.$$

If $c(t)$ is a curve on Q and $X \in \mathfrak{X}(Q)$, the covariant derivative of X along $c(t)$ is defined by

$$\frac{DX}{Dt} := \nabla_{\dot{c}} X,$$

or locally,

$$\left(\frac{DX}{Dt} \right)^i = \Gamma_{jk}^i(c(t)) \dot{c}^j(t) X^k(c(t)) + \frac{d}{dt} X^i(c(t)).$$

A vector field is said to be **parallel transported** along $c(t)$ if $\frac{DX}{Dt} = 0$. Thus $\dot{c}(t)$ is parallel transported along $c(t)$ if and only if

$$\ddot{c}^i + \Gamma_{jk}^i \dot{c}^j \dot{c}^k = 0.$$

In classical differential geometry a **geodesic** is defined to be a curve $c(t)$ in Q whose tangent vector $\dot{c}(t)$ is parallel transported along $c(t)$. As the expression above shows, geodesics are base integral curves of the Lagrangian vector field defined by the kinetic energy of g .

A **classical mechanical system** is given by a Lagrangian of the form $L(v_q) = K(v_q) - V(q)$, for $v_q \in T_q Q$. The smooth function $V : Q \rightarrow \mathbb{R}$ is called the **potential energy**. The total energy of this system is given by $E = K + V$ and the Euler-Lagrange equations (which are always second order) are

$$\ddot{q}^i + \Gamma_{jk}^i \dot{q}^j \dot{q}^k + g^{il} \frac{\partial V}{\partial q^l} = 0, \quad i = 1, \dots, n,$$

where g^{ij} are the entries of the inverse matrix of (g_{ij}) . If $Q = \mathbb{R}^3$ and the metric is given by $g_{ij} = \delta_{ij}$, these equations are Newton's equations of motion (1.3) of a particle in a potential field with which we began our discussion in the Introduction.

3.3 Hyperregular Lagrangians

We shall summarize here the precise equivalence between the Lagrangian and Hamiltonian formulation for hyperregular Lagrangians and Hamiltonians. The proofs are easy lengthy verifications; see [AbMa78] or [MaRa94].

- (a) Let L be a hyperregular Lagrangian on TQ and $H = E \circ (\mathbb{F}L)^{-1}$, where E is the energy of L . Then the Lagrangian vector field Z on TQ and the Hamiltonian vector field X_H on T^*Q are related by the identity

$$(\mathbb{F}L)^* X_H = Z.$$

Furthermore, if $c(t)$ is an integral curve of Z and $d(t)$ an integral curve of X_H with $\mathbb{F}L(c(0)) = d(0)$, then $\mathbb{F}L(c(t)) = d(t)$ and their base integral curves coincide, that is, $\tau_Q(c(t)) = \pi_Q(d(t)) = \gamma(t)$, where $\tau_Q : TQ \rightarrow Q$ and $\pi_Q : T^*Q \rightarrow Q$ are the canonical bundle projections.

- (b) A Hamiltonian $H : T^*Q \rightarrow \mathbb{R}$ is said to be **hyperregular** if the smooth vector bundle map covering the identity, $\mathbb{F}H : T^*Q \rightarrow TQ$, defined by

$$\langle \mathbb{F}H(\alpha_q), \beta_q \rangle := \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} H(\alpha_q + \epsilon\beta_q), \quad \alpha_q, \beta_q \in T_q^*Q,$$

is a diffeomorphism. Define the **action** of H by $G := \langle \Theta, X_H \rangle$. If H is a hyperregular Hamiltonian then the energies of L and H and the actions of L and H are related by

$$E = H \circ (\mathbb{F}H)^{-1}, \quad A = G \circ (\mathbb{F}H)^{-1}.$$

In addition, the Lagrangian $L = A - E$ is hyperregular and $\mathbb{F}L = \mathbb{F}H^{-1}$.

- (c) The constructions above define a bijective correspondence between hyperregular Lagrangians and Hamiltonians.

Example. Let us return to the example in §2.5 of a particle of charge e and mass m moving in a magnetic field $\mathbf{B}$. Recall that if $\mathbf{B} = \nabla \times \mathbf{A}$ is a given magnetic field on $\mathbb{R}^3$, then with respect to the canonical variables $(\mathbf{q}, \mathbf{p})$ on $T^*\mathbb{R}^3$, the canonical symplectic form $d\mathbf{q} \wedge d\mathbf{p}$, and the Hamiltonian

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2m} \left\| \mathbf{p} - \frac{e}{c} \mathbf{A} \right\|^2, \quad (3.4)$$

Hamilton's equations coincide with Newton's equations for the Lorentz force law (2.13). It is obvious that H is hyperregular since

$$\dot{\mathbf{q}} = \mathbb{F}H(\mathbf{q}, \mathbf{p}) = \frac{1}{m} \left(\mathbf{p} - \frac{e}{c} \mathbf{A} \right)$$

has the inverse

$$\mathbf{p} = m\dot{\mathbf{q}} + \frac{e}{c}\mathbf{A}.$$

The Lagrangian is given by

$$\begin{aligned} L(\mathbf{q}, \dot{\mathbf{q}}) &= \mathbf{p} \cdot \dot{\mathbf{q}} - H(\mathbf{q}, \mathbf{p}) \\ &= \left(m\dot{\mathbf{q}} + \frac{e}{c}\mathbf{A}\right) \cdot \dot{\mathbf{q}} - \frac{1}{2m} \left\| \mathbf{p} - \frac{e}{c}\mathbf{A} \right\|^2 \\ &= m\|\dot{\mathbf{q}}\|^2 + \frac{e}{c}\mathbf{A} \cdot \dot{\mathbf{q}} - \frac{1}{2m} \|m\dot{\mathbf{q}}\|^2 \\ &= \frac{m}{2} \|\dot{\mathbf{q}}\|^2 + \frac{e}{c}\mathbf{A} \cdot \dot{\mathbf{q}}. \end{aligned} \quad (3.5)$$

Lagrange's equations for this L can be directly verified to yield $m\ddot{\mathbf{q}} = \frac{e}{c}\dot{\mathbf{q}} \times \mathbf{B}$, that is, Newton's equations for the Lorentz force (2.13).

These equations are not geodesic because the Lagrangian is not given by the kinetic energy of a metric. Can one enlarge the space such that these equations are induced by some geodesic equations on a higher dimensional space? The answer is positive and is given by the **Kaluza-Klein construction**. We shall return once more to this example after we have presented the basic elements of reduction theory. The direct construction proceeds as follows. Define the manifold $Q_{KK} := \mathbb{R}^3 \times S^1$ with variables $(\mathbf{q}, \theta)$. On Q_{KK} introduce the one-form $A + d\theta$ (a connection one-form on the trivial circle bundle $\mathbb{R}^3 \times S^1 \rightarrow \mathbb{R}^3$) and define the **Kaluza-Klein Lagrangian**

$$\begin{aligned} L_{KK}(\mathbf{q}, \theta, \dot{\mathbf{q}}, \dot{\theta}) &= \frac{1}{2}m\|\dot{\mathbf{q}}\|^2 + \frac{1}{2} \left\| \left\langle A + d\theta, (\dot{\mathbf{q}}, \dot{\theta}) \right\rangle \right\|^2 \\ &= \frac{1}{2}m\|\dot{\mathbf{q}}\|^2 + \frac{1}{2} \left(\mathbf{A} \cdot \dot{\mathbf{q}} + \dot{\theta} \right)^2. \end{aligned} \quad (3.6)$$

Note that L_{KK} is positive definite in $(\dot{\mathbf{q}}, \dot{\theta})$ so it is the kinetic energy of a metric, the **Kaluza-Klein metric** on Q_{KK} . Thus the Euler-Lagrange equations for L_{KK} are the geodesic equations of this metric on $\mathbb{R}^3 \times S^1$. (For the readers who know a little about connections and Lie algebras, it is obvious that this construction can be generalized to a principal bundle with compact structure group endowed with a connection; one gets in this way Yang-Mills theory.) The Legendre transformation for L_{KK} gives the momenta

$$\mathbf{p} = m\dot{\mathbf{q}} + (\mathbf{A} \cdot \dot{\mathbf{q}} + \dot{\theta})\mathbf{A} \quad \text{and} \quad \pi = \mathbf{A} \cdot \dot{\mathbf{q}} + \dot{\theta}. \quad (3.7)$$

Since L_{KK} does not depend on θ , the Euler-Lagrange equation

$$\frac{d}{dt} \frac{\partial L_{KK}}{\partial \dot{\theta}} = \frac{\partial L_{KK}}{\partial \theta} = 0$$

shows that $\pi = \partial L_{KK}/\partial \dot{\theta}$ is conserved. The **charge** is now defined by $e := c\pi$. The Hamiltonian H_{KK} associated to L_{KK} by the Legendre transformation (3.7) is

$$\begin{aligned}
 H_{KK}(\mathbf{q}, \theta, \mathbf{p}, \pi) &= \mathbf{p} \cdot \dot{\mathbf{q}} + \pi \dot{\theta} - L_{KK}(\mathbf{q}, \dot{\mathbf{q}}, \theta, \dot{\theta}) \\
 &= \mathbf{p} \cdot \frac{1}{m} (\mathbf{p} - \pi \mathbf{A}) + \pi (\pi - \mathbf{A} \cdot \dot{\mathbf{q}}) \\
 &\quad - \frac{1}{2} m \|\dot{\mathbf{q}}\|^2 - \frac{1}{2} \pi^2 \\
 &= \mathbf{p} \cdot \frac{1}{m} (\mathbf{p} - \pi \mathbf{A}) + \frac{1}{2} \pi^2 \\
 &\quad - \pi \mathbf{A} \cdot \frac{1}{m} (\mathbf{p} - \pi \mathbf{A}) - \frac{1}{2m} \|\mathbf{p} - \pi \mathbf{A}\|^2 \\
 &= \frac{1}{2m} \|\mathbf{p} - \pi \mathbf{A}\|^2 + \frac{1}{2} \pi^2. \tag{3.8}
 \end{aligned}$$

Since $\pi = e/c$ is a constant, this Hamiltonian, regarded as a function of only the variables $(\mathbf{q}, \mathbf{p})$, is up to a constant equal to the Lorentz force Hamiltonian (3.4). This fact is not an accident and is due to the geometry of reduction. This example will be redone in §7.3 from the point of view of reduction theory.

3.4 Variational Principles

We now sketch the variational approach to mechanics.

Theorem 3.5 (Variational Principle of Hamilton) *Let L be a Lagrangian on TQ . A C^2 curve $c : [a, b] \rightarrow Q$ joining $q_1 = c(a)$ to $q_2 = c(b)$ satisfies the Euler-Lagrange equations if and only if*

$$\delta \int_a^b L(c(t), \dot{c}(t)) dt = 0.$$

Proof The meaning of the variational derivative in the statement is the following. Consider a family of C^2 curves $c_\lambda(t)$ for $|\lambda| < \varepsilon$ satisfying $c_0(t) = c(t)$, $c_\lambda(a) = q_1$, and $c_\lambda(b) = q_2$ for all $\lambda \in (-\varepsilon, \varepsilon)$. Then

$$\delta \int_a^b L(c(t), \dot{c}(t)) dt := \left. \frac{d}{d\lambda} \right|_{\lambda=0} \int_a^b L(c_\lambda(t), \dot{c}_\lambda(t)) dt.$$

Differentiating under the integral sign, working in local coordinates (covering the curve $c(t)$ by a finite number of coordinate charts), integrating by parts, denoting

$$v(t) := \left. \frac{d}{d\lambda} \right|_{\lambda=0} c_\lambda(t),$$

and taking into account that $v(a) = v(b) = 0$, yields

$$\int_a^b \left(\frac{\partial L}{\partial q^i} v^i + \frac{\partial L}{\partial \dot{q}^i} \dot{v}^i \right) dt = \int_a^b \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \right) v^i dt.$$

This vanishes for any C^1 function $v(t)$ if and only if the Euler-Lagrange equations hold. ■

The integral appearing in this theorem

$$\mathcal{A}(c(\cdot)) := \int_a^b L(c(t), \dot{c}(t)) dt$$

is called the **action integral**. It is defined on C^2 curves $c : [a, b] \rightarrow Q$ with fixed endpoints, $c(a) = q_1$ and $c(b) = q_2$.

The next theorem emphasizes the role of the Lagrangian one- and two-forms in the variational principle. It is a direct corollary of the previous theorem.

Theorem 3.6 *Given a C^k Lagrangian $L : TQ \rightarrow \mathbb{R}$ for $k \geq 2$, there exists a unique C^{k-2} map $\mathcal{EL}(L) : \ddot{Q} \rightarrow T^*Q$, where*

$$\ddot{Q} := \left\{ \frac{d^2 q}{dt^2}(0) \in T(TQ) \mid q(t) \text{ is a } C^2 \text{ curve in } Q \right\}$$

is a submanifold of $T(TQ)$ (second order submanifold), and a unique C^{k-1} one-form $\Theta_L \in \Omega^1(TQ)$, such that for all C^2 variations $q_\varepsilon(t)$ (defined on a fixed t -interval) of $q_0(t) := q(t)$, we have

$$\mathbf{d}\mathcal{A}[q(\cdot)] \cdot \delta q(\cdot) = \int_a^b \mathcal{EL}(L) \left(\frac{d^2 q}{dt^2} \right) \cdot \delta q \, dt + \Theta_L \left(\frac{dq}{dt} \right) \widehat{\delta q} \Big|_a^b$$

where

$$\delta q(t) = \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} q_\varepsilon(t), \quad \widehat{\delta q}(t) = \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \frac{dq_\varepsilon(t)}{dt}.$$

The map $\mathcal{EL} : \ddot{Q} \rightarrow T^*Q$ is called the **Euler-Lagrange operator** and its expression in local coordinates is

$$\mathcal{EL}(q^j, \dot{q}^j, \ddot{q}^j)_i = \frac{\partial L}{\partial \dot{q}^i} - \frac{d}{dt} \frac{\partial L}{\partial \ddot{q}^i},$$

where it is understood that the formal time derivative is taken in the second summand and everything is expressed as a function of $(q^j, \dot{q}^j, \ddot{q}^j)$. The one-form Θ_L , whose existence and uniqueness is guaranteed by this theorem, appears as the boundary term of the derivative of the action integral if the endpoints of the curves on the configuration manifold are free; it coincides with

$(\mathbb{F}L)^*\Theta$, where Θ is the canonical one-form on T^*Q , defined in §3.1 as an easy verification in coordinates shows.

Remark. From the variational principle one can recover well-known results for regular Lagrangians; for proofs of these statements see e.g. [MaRa94], §8.2.

- (i) If F_t is the flow of the Lagrangian vector field, then $F_t^*\Omega_L = \Omega_L$, where $\Omega_L = -d\Theta_L$.
- (ii) If L is time dependent, for $|t - t_0|$ small, and $q^i(s)$ the solution of the Euler-Lagrange equation subject to the condition $q^i(t_0) = \bar{q}^i$, the convex neighborhood theorem guarantees that the action integral

$$S(q^i, \bar{q}^i, t) = \int_{t_0}^t L(q^i(s), \dot{q}^i(s), s) ds,$$

satisfies the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}, t\right) = 0. \quad (3.9)$$

There is another classical variational principle involving the Hamiltonian, known under the name of **Hamilton's phase space principle**. Denote in what follows by $\pi_Q : T^*Q \rightarrow Q$ the cotangent bundle projection.

Theorem 3.7 *Let $H : T^*Q \rightarrow \mathbb{R}$ a smooth Hamiltonian. A C^1 curve $z : [a, b] \rightarrow T^*Q$ joining $z(a)$ to $z(b)$ satisfies Hamilton's equations if and only if*

$$\delta \int_a^b (\langle \Theta, \dot{z}(t) \rangle - H(z(t))) dt = 0,$$

where the variations δz satisfy $T\pi_Q(\delta\alpha(a)) = T\pi_Q(\delta\alpha(b)) = 0$.

Proof One follows the method of proof of Theorem 3.5. The meaning of the variational derivative is the following. Consider a family of C^1 curves $z_\lambda(t)$ for $|\lambda| < \varepsilon$ satisfying $z_0(t) = z(t)$, $z_\lambda(a) = z_1$, and $z_\lambda(b) = z_2$ for all $\lambda \in (-\varepsilon, \varepsilon)$. Then

$$\delta \int_a^b (\langle \Theta, \dot{z}(t) \rangle - H(z(t))) dt := \left. \frac{d}{d\lambda} \right|_{\lambda=0} \int_a^b (\langle \Theta, \dot{z}_\lambda(t) \rangle - H(z_\lambda(t))) dt.$$

Differentiating under the integral sign, working in local coordinates (covering the curve $z(t)$ by a finite number of coordinate charts), integrating by parts, denoting

$$\delta z(t) := \left. \frac{d}{d\lambda} \right|_{\lambda=0} z_\lambda(t),$$

writing in coordinates $\delta z(t) = (\delta q^i(t), \delta p_i(t))$, and taking into account that $\delta q^i(a) = \delta q^i(b) = 0$, yields

$$\begin{aligned} & \frac{d}{d\lambda} \Big|_{\lambda=0} \int_a^b ((p_\lambda)_i(t) (\dot{q}_\lambda)^i - H((q_\lambda)^i(t), (p_\lambda)_i(t))) dt \\ &= \int_a^b \left((\delta p)_i(t) \dot{q}^i(t) + p_i(t) (\delta \dot{q})^i(t) - \frac{\partial H(q^i(t), p_i(t))}{\partial q^i} (\delta q)^i(t) \right. \\ & \quad \left. - \frac{\partial H(q^i(t), p_i(t))}{\partial p_i} (\delta p)_i(t) \right) dt \\ &= \int_a^b \left[\left(\dot{q}^i(t) - \frac{\partial H(q^i(t), p_i(t))}{\partial p_i} \right) (\delta p)_i(t) \right. \\ & \quad \left. - \left(\dot{p}_i + \frac{\partial H(q^i(t), p_i(t))}{\partial q^i} \right) (\delta q)^i(t) \right] dt. \end{aligned}$$

This vanishes for any functions $\delta q^i, \delta p_i$ if and only if Hamilton's equations hold. $\blacksquare$

Critical point theory. We have seen that under appropriate regularity assumptions, the vanishing of the first variation of the action integral is equivalent to the Euler-Lagrange equations. That is, critical points of the action integral are solutions of the equations of motion. Now, critical points of a function are connected to the topology of its level sets and are described by Morse and Ljusternik-Schirelman theory. For functions on finite dimensional spaces see e.g. [DFN95] or [Milnor] for an exposition of this theory. For the action integral, which is defined on the infinite dimensional manifold of admissible motions, see e.g. [MawWil989] for the development of the relevant Morse and Ljusternik-Schirelman theory.

Here is an example of a strategy for the search of periodic orbits in Lagrangian systems. Let Q be a n -dimensional configuration manifold. We search a trajectory $\gamma(t) \in Q$, such that $\gamma(t) = \gamma(T + t)$ for a certain number $T > 0$ (the period), where $\gamma(t)$ is a solution of the Euler-Lagrange equations. To do this, we study the action integral $\mathcal{A}$, where

$$\mathcal{A}[q] = \int_0^T L(\dot{q}(t), q(t)) dt,$$

is defined on a space of T -periodic trajectories lying in Q of a certain differentiability class.

Since we want to consider continuous paths, by the Sobolev Embedding Theorem, we take

$$H^1([0, T]; Q) := \{q(\cdot) : [0, T] \rightarrow Q \mid q(\cdot) \text{ of class } H^1, q(0) = q(T)\}$$

which is an infinite dimensional Hilbert manifold; it is a simple example of a manifold of maps from a compact manifold with boundary to an arbitrary smooth paracompact manifold (see [Palais68] or [EbMa70] for example). However, in this case, there is simpler more direct construction (see e.g. [Ben86]). Embed Q into some $\mathbb{R}^N$ and consider the space of loops $H^1([0, T]; \mathbb{R}^N)$ of period T with values in $\mathbb{R}^N$, obtained as the completion of the space of smooth loops $C^\infty([0, T]; \mathbb{R}^n)$ relative to the norm

$$\|q\|_1^2 := \int_0^T \|\dot{q}(t)\|^2 dt + \int_0^T \|q(t)\|^2 dt.$$

Therefore, $q(\cdot) \in H^1([0, T]; \mathbb{R}^n)$ is an absolutely continuous loop with L^2 derivative; $H^1([0, T]; \mathbb{R}^N)$ is compactly immersed in $C^0([0, T]; \mathbb{R}^N)$ (see e.g. [BlBr92]). Then $H^1([0, T]; Q)$ is the submanifold of $H^1([0, T]; \mathbb{R}^N)$ consisting of all loops with values in Q .

Critical points of the action integral $\mathcal{A}$ are trajectories where the homology of the level sets of $\mathcal{A}$ changes. For the sake of simplicity we can consider them to be minimal. Therefore, given a subset $\mathcal{M} \subset H^1([0, T]; Q)$ we could look for $\gamma^*(\cdot) \in \mathcal{M}$ subject to the condition

$$\gamma^* = \min_{\gamma \in \mathcal{M}} \mathcal{A}[\gamma].$$

Typical conditions that guarantee the existence of a minimum γ^* for $\mathcal{A}$ are:

- (i) $\mathcal{M}$ is weakly closed in $H^1([0, T]; Q)$,
- (ii) $\mathcal{A}$ is C^1 on $\mathcal{M}$ and bounded from below,
- (iii) the set $\{\gamma \in H^1([0, T]; Q) \mid \mathcal{A}[\gamma] < \infty, \mathbf{d}\mathcal{A}[\gamma] = 0\}$ is compact in $H^1([0, T]; Q)$.

Points (i) and (iii) imply that γ^* solves the equations of motion, that is,

$$\mathbf{d}\mathcal{A}[\gamma^*](u) = 0,$$

for all $u(\cdot) \in T_{\gamma^*} H^1([0, T]; Q)$. We refer, as general references, to [BlBr92], [Jost], and [MawWil989], and for applications of this theory to the N -body problem to [AmCZ96].

The judicious choice of the domain $\mathcal{M}$, where one searches for critical points of $\mathcal{A}$, involves a lot of geometrical, analytical, and physical information, intimately tied to the problem under consideration. In fact, the form of the potential provides severe restrictions; for example, in the N -body problem one has to take into account all the trajectories leading to collisions of two or more bodies. Also, when a system is symmetric then the symmetry group gives additional structure to $\mathcal{M}$ and leads to topological constraints; for example, restrictions on the homotopy type for the trajectories enter in this case.

Regarding the connection between global variational methods and geometry we want to mention an interesting problem. We have seen that the Euler equations $\dot{\Pi} = \Pi \times \Omega$ for the free rigid body motion can be formulated in a variational setting assuming a particular class of variations. This construction can be generalized to any Lie group, as we shall see in Theorem 6.6. It would be very interesting to exploit the geometrical structure and to extend the powerful methods of the calculus of variations to formulate results about the existence of critical points of such types of restricted variational principles.

4 Poisson Manifolds

This lecture quickly reviews the basic theory of Poisson manifolds. Very little will be proved here and we refer to standard books (such as [GuSt84, LiMa87, Marsden92, MaRa94, McDsal95, V96]) and [W] for the detailed discussion of the topics below. Unless otherwise specified, all manifolds in this chapter are finite dimensional. Whenever infinite dimensional manifolds will be used the results presented here are formal. A theory for infinite dimensional Poisson manifolds generalizing that of strong symplectic manifolds can be found in [OdzRa03] (for more examples see [OdzRa004] and [BelRa04]). No general satisfactory theory for infinite dimensional general Poisson manifolds has been developed yet, even though there are many significant examples, some of which will be presented in these lectures. Remarks regarding the infinite dimensional situation will be made throughout the text.

4.1 Fundamental Concepts

Definition 4.1 A *Poisson bracket* on a manifold P is a bilinear operation $\{, \}$ on the space $\mathcal{F}(P) := \{F : P \rightarrow \mathbb{R} \mid F \text{ is smooth}\}$ verifying the following conditions:

- (i) $(\mathcal{F}(P), \{, \})$ is a Lie algebra, and
- (ii) $\{, \}$ satisfies the Leibniz identity on each factor.

A manifold endowed with a Poisson bracket is called a *Poisson manifold* and is denoted by $(P, \{, \})$. The elements of the center $\mathcal{C}(P)$ of the Poisson algebra are called *Casimir functions*.

A smooth map $\varphi : (P_1, \{, \}_1) \rightarrow (P_2, \{, \}_2)$ between the Poisson manifolds $(P_1, \{, \}_1)$ and $(P_2, \{, \}_2)$ is called a *canonical* or *Poisson map*, if $\varphi^*\{F, H\}_2 = \{\varphi^*F, \varphi^*H\}_1$ for any $F, H \in \mathcal{F}(P_2)$.

Note that (i) is equivalent to the statement that the Poisson bracket is real, bilinear, antisymmetric, and satisfies the Jacobi identity. Furthermore, (ii) states that the linear map $F \mapsto \{F, H\}$ (and $H \mapsto \{F, H\}$) is a **derivation**, that is,

$$\{FK, H\} = \{F, H\}K + F\{K, H\}$$

for all $F, H, K \in \mathcal{F}(P)$.

As in the case of strong symplectic manifolds, we can define Hamiltonian vector fields on a finite dimensional Poisson manifold.

Definition 4.2 Let $(P, \{, \})$ be a finite dimensional Poisson manifold and $H \in \mathcal{F}(P)$. The unique vector field X_H on P that satisfies $X_H[F] = \{F, H\}$ for all $F \in \mathcal{F}(P)$, is called the **Hamiltonian vector field** associated to the **Hamiltonian function** H .

Note that the Jacobi identity is equivalent to

$$[X_F, X_H] = -X_{\{F, H\}}$$

for all $F, H \in \mathcal{F}(P)$.

Any symplectic manifold (P, Ω) is Poisson. First, recall that there is a Poisson bracket naturally defined on P , namely, $\{F, H\} = \Omega(X_F, X_H)$, where X_H is defined by the identity $\mathbf{d}H = \Omega(X_H, \cdot)$. Second, the relation $X_H[F] = \langle \mathbf{d}F, X_H \rangle = \Omega(X_F, X_H)$, shows that the Hamiltonian vector field defined via the symplectic form coincides with the Hamiltonian vector field defined using the Poisson bracket.

If P is a Poisson manifold, note that $F \in \mathcal{C}(P)$ if and only if $X_F = 0$.

If ϕ_t is the flow of X_H , then $H \circ \phi_t = H$, that is, H is conserved. Indeed,

$$\frac{d}{dt}(H \circ \phi_t) = \left\langle \mathbf{d}H, \frac{d}{dt}\phi_t \right\rangle = \langle \mathbf{d}H, X_H \rangle = \{H, H\} = 0.$$

Thus $H \circ \phi_t$ is constant in t and since ϕ_0 is the identity, it follows that $H \circ \phi_t = H$, for all t .

Hamilton's equations $\dot{z} = X_H(z)$ for the function $H \in \mathcal{F}(P)$ can be equivalently written as

$$\frac{d}{dt}(F \circ \phi_t) = \{F, H\} \circ \phi_t \quad \text{or, in shorthand notation,} \quad \dot{F} = \{F, H\}$$

for any $F \in \mathcal{F}(P)$. To see this, note first that if ϕ_t is the flow of X_H , we have $\phi_t^* X_H = X_H$, or, equivalently, $T\phi_t \circ X_H = X_H \circ \phi_t$. Thus, for any $z \in P$

we have

$$\begin{aligned} \frac{d}{dt}(F \circ \phi_t)(z) &= \left\langle \mathbf{d}F(\phi_t(z)), \frac{d}{dt}\phi_t(z) \right\rangle = \langle \mathbf{d}F(\phi_t(z)), X_H(\phi_t(z)) \rangle \\ &= \langle \mathbf{d}F(\phi_t(z)), T_z\phi_t(X_H(z)) \rangle = \langle \mathbf{d}(F \circ \phi_t)(z), X_H(z) \rangle \\ &= \{F \circ \phi_t, H\}(z). \end{aligned}$$

Since $H \circ \phi_t = H$ we conclude that Hamilton's equations on a Poisson manifold can be written as

$$\frac{d}{dt}(F \circ \phi_t) = \{F, H\} \circ \phi_t = \{F \circ \phi_t, H\} = \{F \circ \phi_t, H \circ \phi_t\}. \quad (4.1)$$

Proposition 4.3 *The flows of Hamiltonian vector fields are Poisson diffeomorphisms.*

Proof Let ϕ_t be the flow of the Hamiltonian vector field X_H , that is, $\frac{d}{dt}\phi_t = X_H \circ \phi_t$. We need to prove $\{F, K\} \circ \phi_t = \{F \circ \phi_t, K \circ \phi_t\}$ for any $F, K \in \mathcal{F}(P)$. To see this, note that for any $z \in P$,

$$\begin{aligned} \frac{d}{dt}\{F, K\}(\phi_t(z)) &= \left\langle \mathbf{d}\{F, K\}(\phi_t(z)), \frac{d}{dt}\phi_t(z) \right\rangle \\ &= \langle \mathbf{d}\{F, K\}(\phi_t(z)), X_H(\phi_t(z)) \rangle = \{\{F, K\}, H\}(\phi_t(z)), \end{aligned}$$

that is,

$$\frac{d}{dt}(\{F, K\} \circ \phi_t) = \{\{F, K\}, H\} \circ \phi_t = \{\{F, K\} \circ \phi_t, H\} \quad (4.2)$$

by (4.1). On the other hand, the bilinearity of the Poisson bracket gives

$$\frac{d}{dt}\{F \circ \phi_t, K \circ \phi_t\} = \left\{ \frac{d}{dt}(F \circ \phi_t), K \circ \phi_t \right\} + \left\{ F \circ \phi_t, \frac{d}{dt}(K \circ \phi_t) \right\}.$$

This equality, the Jacobi identity, and (4.1) imply that

$$\begin{aligned} \frac{d}{dt}\{F \circ \phi_t, K \circ \phi_t\} &= \{\{F \circ \phi_t, H\}, K \circ \phi_t\} + \{F \circ \phi_t, \{K \circ \phi_t, H\}\} \\ &= \{\{F \circ \phi_t, K \circ \phi_t\}, H\}. \end{aligned} \quad (4.3)$$

Comparing (4.2) and (4.3) one sees that both $\{F, K\} \circ \phi_t$ and $\{F \circ \phi_t, K \circ \phi_t\}$ satisfy the same equation, namely, $\dot{L} = \{L, H\}$. Since for $t = 0$, the functions $\{F, K\} \circ \phi_t$ and $\{F \circ \phi_t, K \circ \phi_t\}$ are both equal to $\{F, K\}$, it follows that $\{F, K\} \circ \phi_t = \{F \circ \phi_t, K \circ \phi_t\}$ for all t . ■

The same strategy of proof is used to show the following statement.

Proposition 4.4 *A smooth map $\phi : (P_1, \{\cdot, \cdot\}_1) \rightarrow (P_2, \{\cdot, \cdot\}_2)$ is Poisson if and only if for any $H \in \mathcal{F}(V)$, V open in P_2 , we have $T\phi \circ X_{H \circ \phi} = X_H \circ \phi$ on $\phi^{-1}(V)$.*

Canonical maps are the key ingredient in the definition of the notion of a Poisson submanifold.

Definition 4.5 Let $(P_1, \{\cdot, \cdot\}_1)$ and $(P_2, \{\cdot, \cdot\}_2)$ be two Poisson manifolds, $P_1 \subset P_2$, such that the inclusion $i : P_1 \hookrightarrow P_2$ is an immersion. The Poisson manifold $(P_1, \{\cdot, \cdot\}_1)$ is called a **Poisson submanifold** of $(P_2, \{\cdot, \cdot\}_2)$ if i is a canonical map.

An immersed submanifold P_1 of P_2 is called a **quasi Poisson submanifold** of $(P_2, \{\cdot, \cdot\}_2)$ if for any $p \in P_1$, any open neighborhood U of p in P_2 , and any $F \in C^\infty(U)$ we have

$$X_F(i(p)) \in T_p i(T_p P_1),$$

where X_F is the Hamiltonian vector field of F on U with respect to the Poisson bracket $\{\cdot, \cdot\}_2$ restricted to smooth functions defined on U .

The proofs of the following statements can be found in, e.g., [OR04], §4.1.

- If $(P_1, \{\cdot, \cdot\}_1)$ is a Poisson submanifold of $(P_2, \{\cdot, \cdot\}_2)$ then there is no other bracket $\{\cdot, \cdot\}'$ on P_1 making the inclusion $i : P_1 \hookrightarrow P_2$ into a canonical map.
- If P_1 is a quasi Poisson submanifold of $(P_2, \{\cdot, \cdot\}_2)$ then there exists a unique Poisson bracket on P_1 making it into a Poisson submanifold of P_2 .
- Proposition 4.4 implies that any Poisson submanifold is quasi Poisson. However, a quasi Poisson submanifold P_1 of $(P_2, \{\cdot, \cdot\}_2)$ could carry a Poisson structure that has nothing to do with the one induced from P_2 , so it won't be a Poisson submanifold of P_2 relative to this a priori given structure.
- If (P_1, Ω_1) and (P_2, Ω_2) are two symplectic manifolds such that $P_1 \subset P_2$, then P_1 is said to be a **symplectic submanifold** of P_2 if $i^* \Omega_2 = \Omega_1$, where $i : P_1 \hookrightarrow P_2$ is the inclusion. In this case, the inclusion is necessarily an immersion (P_1 and P_2 are assumed finite dimensional).
- Symplectic submanifolds of a symplectic manifold are, in general, neither Poisson nor quasi Poisson submanifolds.
- The only quasi Poisson submanifolds of a symplectic manifold are its open sets which are, in fact, Poisson submanifolds.

4.2 Structure Theorems

The derivation property of a Poisson bracket $\{\cdot, \cdot\}$ implies that the value $X_F(z)$ of the Hamiltonian vector field of F at $z \in P$ depends on F only through

$\mathbf{d}F(z)$. Thus there is a contravariant antisymmetric two-tensor Λ on P , called the **Poisson tensor**, defined by

$$\Lambda(z)(\mathbf{d}F(z), \mathbf{d}H(z)) = \{F, H\}(z)$$

for any $F, H \in \mathcal{F}(U)$, U open in P , $z \in U$.

In finite dimensions, if $(z^1, z^2, \dots, z^n)$ are local coordinates, Λ is determined by the matrix $[\Lambda^{ij}]$, where $\Lambda^{ij} = \{z^i, z^j\}$, and hence the expression of the Poisson bracket of the two functions $F, H \in \mathcal{F}(P)$ is given in terms of Λ by

$$\{F, H\} = \Lambda^{ij} \frac{\partial F}{\partial z^i} \frac{\partial H}{\partial z^j}.$$

The **rank** of the Poisson structure at a point $z \in P$ is defined to be the rank of the matrix $[\Lambda^{ij}]$.

The Poisson tensor defines a vector bundle map $\Lambda^\sharp : T^*P \rightarrow TP$ by

$$\Lambda(z)(\alpha_z, \beta_z) = \langle \alpha_z, \Lambda^\sharp(\beta_z) \rangle, \quad \alpha_z, \beta_z \in T_z^*P.$$

Since

$$\{F, H\}(z) = \Lambda(z)(\mathbf{d}F(z), \mathbf{d}H(z)) = \langle \mathbf{d}F(z), \Lambda_z^\sharp(\mathbf{d}H(z)) \rangle,$$

it follows that the Hamiltonian vector field X_H is given by

$$X_H(z) = \Lambda_z^\sharp(\mathbf{d}H(z)),$$

that is, $\Lambda_z^\sharp : \mathbf{d}H(z) \mapsto X_H(z)$. Thus the image of $\Lambda_z^\sharp$ is the set of all Hamiltonian vector fields evaluated at z .

Note that if the Poisson tensor is nondegenerate, that is $\Lambda^\sharp : T_z^*P \rightarrow T_zP$ is an isomorphism for all $z \in P$, then P is symplectic with the symplectic form $\Omega(X_F, X_H) := \{F, H\}$ for all locally defined Hamiltonian vector fields X_F, X_H (the closeness of Ω is equivalent to the Jacobi identity of the Poisson bracket).

The image $\text{Im}(\Lambda^\sharp) \subset TP$ of $\Lambda^\sharp$ defines a **smooth generalized distribution** on P , i.e., $\text{Im}(\Lambda_z^\sharp) \subset T_zP$ is a vector subspace for each $z \in P$ and for every point $z_0 \in P$ and every vector $v \in \text{Im}(\Lambda_{z_0}^\sharp)$, there exists an open neighborhood U of z_0 and smooth vector field $X \in \mathfrak{X}(U)$ such that $X(u) \in \text{Im}(\Lambda_u^\sharp)$ for all $u \in U$ and $X(z_0) = v$.

If the rank of the Poisson tensor is constant then $\text{Im}(\Lambda^\sharp)$ is a smooth vector subbundle of TP . Furthermore, the Jacobi identity gives $[X_F, X_H] = X_{-\{F, H\}}$ which shows that the distribution $\text{Im}(\Lambda^\sharp)$ is involutive. The Frobenius theorem guarantees then its integrability.

If $\text{Im}(\Lambda^\sharp)$ is not a subbundle then integrability and involutivity (as used in the Frobenius theorem) are not equivalent. Recall that if $D \subset TP$ is a smooth

generalized distribution, an immersed connected submanifold S of P , $S \subset P$, is said to be an **integral manifold** of D if for every $z \in S$, $T_z i(T_z S) \subset D(z)$, where $i : S \rightarrow P$ is the inclusion. The integral submanifold S is said to be of **maximal dimension** at a point $z \in S$ if $T_z i(T_z S) = D(z)$. The smooth generalized distribution D is **integrable** if for every point $z \in P$ there is an integral manifold of D everywhere of maximal dimension containing z . The smooth generalized distribution D is **involutive** if it is invariant under the (local) flows associated to differentiable sections of D . Note that this definition of involutivity is weaker than the one used in the Frobenius theorem and that it only coincides with it when the dimension of $D(z)$ is the same for any $z \in P$.

Theorem 4.6 (Stefan-Sussmann) *The smooth generalized distribution D on a finite dimensional manifold P is integrable if and only if it is involutive.*

Thus P carries a generalized foliation, all of whose integral manifolds are injectively immersed.

If P is a Poisson manifold, $\text{Im}(\Lambda^\#)$ is an involutive smooth generalized distribution so each of its integral submanifolds has its tangent space at every point equal to the vector space of all Hamiltonian vector fields evaluated at that point. Thus, these integral submanifolds are symplectic and the Poisson bracket defined by their symplectic structure coincides with the original Poisson bracket on P . These integral manifolds are called the **symplectic leaves** of P . It turns out that the symplectic leaves are the equivalence classes of the following equivalence relation: $z_1 \mathcal{R} z_2$ if and only if there is a piecewise smooth curve in P joining z_1 and z_2 each segment of which is an integral curve of some Hamiltonian vector field. The following theorem summarizes this discussion.

Theorem 4.7 (Symplectic Stratification Theorem) *Let P be a finite dimensional Poisson manifold. Then P is the disjoint union of its symplectic leaves. Each symplectic leaf in P is an injectively immersed Poisson submanifold and the induced Poisson structure on the leaf is symplectic. The dimension of a leaf through a point z equals the rank of the Poisson structure at that point.*

Note that if C is a Casimir function, then $\Lambda^\#(\mathbf{d}C) = 0$, which shows that Casimir functions are constant on symplectic leaves. However, one should not conclude that the symplectic leaves are level sets of Casimir functions. This is not even true for the maximal dimensional ones, which are generic. For example, symplectic leaves may be open or they may all have a common accumulation point. Worse, there are Poisson manifolds with no global Casimir

functions. Locally, Casimir functions always exist generically as the next local structure theorem shows.

Theorem 4.8 (Weinstein) *Let P be a finite dimensional Poisson manifold and $z \in P$. There exists a neighborhood U of z and an isomorphism $\phi = \phi_S \times \phi_N : U \rightarrow S \times N$ where S is symplectic, N is Poisson and the rank of N at $\phi_N(z)$ is zero. The factors S, N are unique up to a local isomorphism.*

In this theorem S can be chosen to be an open set in the symplectic leaf through z and N any submanifold of P transverse to it such that $S \cap N = \{z\}$. While there is no canonical choice of N in general, the Poisson structure on it is uniquely determined up to a Poisson isomorphism. This Poisson structure is called the **transverse Poisson structure** at z .

Assume that the Poisson structure has rank $0 \leq 2k \leq \dim(P) = 2k + l$ at the point $z \in P$. Then there are coordinates $(q^1, \dots, q^k, p_1, \dots, p_k, y^1, \dots, y^l)$ in a chart around z such that

$$\{q^i, q^j\} = \{p_i, p_j\} = \{q^i, y^j\} = \{p_i, y^j\} = 0, \{q^i, p_j\} = \delta_j^i$$

and the brackets $\{y^i, y^j\}$ depend only on $y^1, \dots, y^l$ and vanish at the point z . In this chart, the transverse Poisson structure is given by the subspace defining these coordinates $y^1, \dots, y^l$. If, in addition, there is a neighborhood of z such that the rank of the Poisson structure is constant on it then, shrinking if necessary the above chart such that it lies in this neighborhood, the coordinates $y^1, \dots, y^l$ can be chosen such that $\{y^i, y^j\} = 0$ for all $i, j = 1, \dots, l$. In this case, the y^i are the local Casimir functions in a chart about z .

Comments on Banach Poisson manifolds. Definition 4.1 presents several problems in the infinite dimensional case that will be briefly reviewed here. Let $(P, \{, \})$ be a Banach Poisson manifold. The Leibniz property insures, as in finite dimensions, that the value of the Poisson bracket at $z \in P$ depends only on the differentials $\mathbf{d}F(z), \mathbf{d}H(z) \in T_z^*P$ which implies that there is a smooth section Λ of the vector bundle $\bigwedge^2 T^{**}P$ satisfying $\{F, H\} = \Lambda(\mathbf{d}F, \mathbf{d}H)$.

If the Poisson tensor is strongly nondegenerate, that is $\Lambda^\sharp : T_z^*P \rightarrow T_zP$ is an isomorphism of Banach spaces for all $z \in P$, then P is strong symplectic with the symplectic form $\Omega(X_F, X_H) := \{F, H\}$ for all locally defined Hamiltonian vector fields X_F, X_H (the closedness of Ω is equivalent to the Jacobi identity of the Poisson bracket). However if $\Lambda^\sharp$ is one-to-one but not surjective, that is, the Poisson tensor is weakly nondegenerate, then P is not, in general, symplectic (see example of page 344 of [MaRa94]). Worse, a weak symplectic manifold is not a Poisson manifold since not every locally defined function defines a Hamiltonian vector field.

On a Banach Poisson manifold the rule $X_F := \Lambda^\sharp(\mathbf{d}F)$ defines a smooth section of $T^{**}P$ and hence is not, in general, a vector field on P . In analogy with the finite dimensional and the strong symplectic case, we need to require that X_F be a Hamiltonian vector field. In order to achieve this, we are forced to make the assumption that the Poisson bracket on P satisfies the condition $\Lambda^\sharp(T^*P) \subset TP \subset T^{**}P$. The study of such Banach Poisson manifolds was begun in [OdzRa03] with special emphasis on the Lie-Poisson case. See [OdzRa004] and [BelRa04] for further examples of such Banach Poisson manifolds.

While these manifolds are important in quantum mechanics, most infinite dimensional examples in classical continuum mechanics do not satisfy this hypothesis on $\Lambda^\sharp$; in fact, most of them have weak symplectic phase spaces. In this case, the beginning of a systematic theory of weak symplectic manifolds and the associated Hamiltonian dynamics can be found in [ChMa74]. For (weak) Poisson manifolds, not even a proposal of a theory is available today and the rigorous study of several examples coming from fluid dynamics, elasticity theory, and plasma physics should shed light on the general abstract case.

4.3 Examples of Poisson Brackets

1. **Symplectic Bracket.** As we mentioned before, a strong symplectic form Ω on a manifold P gives the Poisson bracket

$$\{F, H\} := \Omega(X_F, X_H).$$

If C is a Casimir function on a connected strong symplectic manifold P , i.e., $\{C, F\} = 0$ for all $F \in \mathcal{F}(P)$, then $X_C = 0$. Strong nondegeneracy of Ω implies then that $\mathbf{d}C = 0$, which in turn shows that C is constant on the connected manifold P . Thus, on a connected strong symplectic manifold the Casimir functions are the constants, i.e., the center of $\mathcal{F}(P)$ is $\mathbb{R}$.

2. **Lie-Poisson Bracket.** Let $\mathfrak{g}$ be a Lie algebra with Lie bracket $[\cdot, \cdot]$ and $\mathfrak{g}^*$ its dual. Define the functional derivative of the smooth function $F : \mathfrak{g}^* \rightarrow \mathbb{R}$ at $\mu \in \mathfrak{g}^*$ to be the unique element $\frac{\delta F}{\delta \mu} \in \mathfrak{g}$ given by

$$\mathbf{D}F(\mu) \cdot \delta\mu = \lim_{\epsilon \rightarrow 0} \frac{F(\mu + \epsilon \delta\mu) - F(\mu)}{\epsilon} = \left\langle \delta\mu, \frac{\delta F}{\delta \mu} \right\rangle$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$. Note that $\mathbf{D}F(\mu)$ is the usual Fréchet derivative, i.e., $\mathbf{D}F(\mu) \in L(\mathfrak{g}^*, \mathbb{R}) = \mathfrak{g}^{**}$. If $\mathfrak{g}$ is finite dimensional, then $\mathfrak{g}^{**} \cong \mathfrak{g}$ naturally and $\delta F / \delta \mu \in \mathfrak{g}$ is the element of $\mathfrak{g}$ representing the functional $\mathbf{D}F(\mu) \in \mathfrak{g}^{**}$ on $\mathfrak{g}^*$.

The Banach space $\mathfrak{g}^*$ is a Poisson manifold for each of the Lie-Poisson brackets $\{, \}_+$ and $\{, \}_-$ defined by

$$\{F, H\}_\pm(\mu) = \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right] \right\rangle \quad (4.4)$$

for all $\mu \in \mathfrak{g}^*$ and $F, H \in \mathcal{F}(\mathfrak{g}^*)$. The bilinearity and skew-symmetry are obvious from the definition. The derivation property follows from the Leibniz rule for functional derivatives. For the direct proof of the Jacobi identity see [MaRa94], pg. 329.

In general, if $\mathfrak{g}$ is a Banach Lie algebra, $\mathfrak{g}^*$ is not a Banach Poisson manifold that satisfies the condition $\Lambda^\sharp(T^*\mathfrak{g}^*) \subset T\mathfrak{g}^*$ discussed at the end of §4.2. It was shown in [OdzRa03] that a Banach space $\mathfrak{b}$ is a Banach Lie-Poisson space $(\mathfrak{b}, \{, \})$ if and only if its dual $\mathfrak{b}^*$ is a Banach Lie algebra $(\mathfrak{b}^*, [,])$ satisfying

$$\text{ad}_x^* \mathfrak{b} \subset \mathfrak{b} \subset \mathfrak{b}^{**} \quad \text{for all } x \in \mathfrak{b}^*, \quad (4.5)$$

where $\text{ad}_x : \mathfrak{b}^* \rightarrow \mathfrak{b}^*$ is the adjoint representation $\text{ad}_x y := [x, y]$ for any $x, y \in \mathfrak{b}^*$. Of course, the Poisson bracket of $F, H \in \mathcal{F}(\mathfrak{b})$ is given in this case by

$$\{F, H\}(b) = \langle [\mathbf{D}F(b), \mathbf{D}H(b)], b \rangle, \quad (4.6)$$

where $b \in \mathfrak{b}$ and $\mathbf{D}$ denotes the Fréchet derivative. If H is a smooth function on $\mathfrak{b}$, the associated Hamiltonian vector field is given by

$$X_H(b) = -\text{ad}_{\mathbf{D}H(b)}^* b. \quad (4.7)$$

Note that if the condition $\text{ad}_x^* \mathfrak{b} \subset \mathfrak{b}$ does not hold for every $x \in \mathfrak{b}^*$, then (4.7) does not make sense, in general. We shall encounter below such situations, in which case formulas (4.6) and (4.7) will be taken formally, or subject to the condition that only functions for which (4.7) makes sense will be used.

3. **Rigid body bracket.** As a particular case of the previous example, consider the Lie algebra $\mathfrak{g} = \mathfrak{so}(3)$ of the rotation group consisting of 3×3 skew-symmetric matrices. This Lie algebra is isomorphic to $(\mathbb{R}^3, \times)$, where $\times$ is the cross product of vectors, via the isomorphism (1.13), that is $(\mathbf{u} \times \mathbf{v})^\wedge = [\hat{\mathbf{u}}, \hat{\mathbf{v}}]$ for any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$.

We identify $\mathfrak{so}(3)^*$ with $\mathbb{R}^3$ using as pairing the Euclidean inner product. The Fréchet derivative and the functional derivative of a function defined on $\mathbb{R}^3$ coincide and are both equal to the usual gradient of the function. Thus the Lie-Poisson bracket (4.4) or (4.6) take the form

$$\{F, H\}_\pm(\mu) = \pm \mu \cdot (\nabla F \times \nabla H).$$

Let us show that $C(\mu) = \Phi\left(\frac{1}{2}\|\mu\|^2\right)$ is a Casimir function, where $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ is an arbitrary differentiable function. Indeed, since $\nabla C(\mu) = \Phi'\left(\frac{1}{2}\|\mu\|^2\right)\mu$, for any $F \in \mathcal{F}(\mathbb{R}^3)$ we have

$$\{C, F\}_{\pm}(\mu) = \pm \mu \cdot (\nabla C \times \nabla F) = \pm \Phi'\left(\frac{1}{2}\|\mu\|^2\right) \mu \cdot (\mu \times \nabla F) = 0.$$

4. **Frozen Lie-Poisson Bracket.** Let $\mathfrak{g}$ be a finite dimensional Lie algebra. For $\nu \in \mathfrak{g}^*$ define for any $F, H \in \mathcal{F}(\mathfrak{g}^*)$ the brackets

$$\{F, H\}_{\mathfrak{g}^*}^{\nu}(\mu) = \pm \left\langle \nu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right] \right\rangle. \quad (4.8)$$

A computation almost identical to the one needed to prove the Jacobi identity for the Lie-Poisson bracket shows that (4.8) also satisfies the Jacobi identity. This Poisson bracket on $\mathfrak{g}^*$ is called the **frozen Lie-Poisson bracket** at $\nu \in \mathfrak{g}^*$. A discussion similar to that at the end of the second example leads to a condition on (4.8) (like (4.5)) that makes it into a rigorous functional analytic Poisson bracket for any smooth functions F and H .

It is worth noting that $\{, \} + s\{, \}^{\nu}$ is also a Poisson bracket on $\mathfrak{g}^*$ for any $\nu \in \mathfrak{g}^*$ and any $s \in \mathbb{R}$. One says that these two Poisson brackets are **compatible**. The verification of this statement is direct, using the previously alluded proofs of the Jacobi identity for the Lie-Poisson and frozen Lie-Poisson brackets.

5. **Ideal Fluid Bracket in Velocity Representation.** In this infinite dimensional example we shall work formally or with the understanding that we restrict the class of functions to those admitting functional derivatives.

Let $\mathfrak{X}_{\text{div}}(D)$ be the Lie algebra of smooth divergence free vector fields tangent to the boundary ∂D defined on an oriented Riemannian manifold D with Riemannian metric g and Riemannian volume form μ . Consider the weakly nondegenerate pairing $\langle, \rangle : \mathfrak{X}_{\text{div}}(D) \times \mathfrak{X}_{\text{div}}(D) \rightarrow \mathbb{R}$ given by the L^2 pairing

$$\langle u, v \rangle = \int_D g(u, v) \mu.$$

Thus, we formally regard $\mathfrak{X}_{\text{div}}(D)^*$ as being $\mathfrak{X}_{\text{div}}(D)$ and apply the formulas from the general theory. The plus Lie-Poisson bracket is

$$\{F, H\}(v) = - \int_D g\left(v, \left[\frac{\delta F}{\delta v}, \frac{\delta H}{\delta v}\right]\right) \mu, \quad (4.9)$$

where the functional derivative $\frac{\delta F}{\delta v}$ is the element of $\mathfrak{X}_{\text{div}}(D)$ defined by

$$\lim_{\epsilon \rightarrow 0} \frac{F(v + \epsilon \delta v) - F(v)}{\epsilon} = \int_D g\left(\frac{\delta F}{\delta v}, v\right) \mu.$$

The minus sign in front of the right hand side of (4.9) appears because the usual Lie bracket of vector fields is the *negative* of the left Lie algebra bracket of the diffeomorphism group, whose Lie algebra has underlying vector space $\mathfrak{X}(D)$; this will be shown explicitly in Section 5.1, Example 2.

6. **Ideal Fluid Bracket in Vorticity Representation.** We continue the previous example realizing the dual of $\mathfrak{X}_{\text{div}}(D)$ in a different manner; this approach is due to [MaWei83]. The natural dual of $\mathfrak{X}(D)$ is the space of one-forms $\Omega^1(D)$ via the weak pairing $(u, \alpha) \in \mathfrak{X}(D) \times \Omega^1(D) \mapsto \int_D \alpha(u) \mu \in \mathbb{R}$. However, restricted to $\mathfrak{X}_{\text{div}}(D)$ this pairing is degenerate. By the Hodge decomposition theorem, the kernel of the linear map sending $\alpha \in \Omega^1(D)$ to the element of $\mathfrak{X}(D)^*$ given by $u \in \mathfrak{X}(D) \mapsto \int_D \alpha(u) \mu \in \mathbb{R}$ is $\mathbf{d}\mathcal{F}(D)$ and thus we can identify, formally, $\mathfrak{X}(D)^*$ with $\Omega^1(D)/\mathbf{d}\mathcal{F}(D)$.

Next, note that the linear map $[\alpha] \in \Omega^1(D)/\mathbf{d}\mathcal{F}(D) \mapsto \mathbf{d}\alpha \in \mathbf{d}\Omega^1(D)$ is well defined and has kernel the first de Rham cohomology group of D . Assuming that this first cohomology group is zero, the above map becomes an isomorphism. Thus, under this topological assumption on D , we can formally identify $\mathfrak{X}(D)^*$ with $\mathbf{d}\Omega^1(D)$. Summarizing, the weak pairing between $\mathfrak{X}_{\text{div}}(D)$ and $\mathbf{d}\Omega^1(D)$ is given by

$$(u, \omega) \in \mathfrak{X}_{\text{div}}(D) \times \mathbf{d}\Omega^1(D) \mapsto \int_D \alpha(u) \mu \in \mathbb{R}, \quad \text{for } \omega := \mathbf{d}\alpha.$$

Therefore, the functional derivative $\delta F/\delta \omega \in \mathfrak{X}_{\text{div}}(D)$ of $F : \mathfrak{X}(D)^* = \mathbf{d}\Omega^1(D) \rightarrow \mathbb{R}$ is defined by the identity

$$\mathbf{D}F(\omega) \cdot \delta \omega = \int_D \delta \alpha \left(\frac{\delta F}{\delta \omega} \right) \mu, \quad \text{for } \delta \omega := \mathbf{d}(\delta \alpha).$$

Thus, the plus Lie-Poisson bracket of $F, H : \mathbf{d}\Omega^1(D) \rightarrow \mathbb{R}$ has the expression

$$\{F, H\}(\omega) = - \int_D \alpha \left(\left[\frac{\delta F}{\delta \omega}, \frac{\delta H}{\delta \omega} \right] \right) \mu, \quad \text{where } \omega = \mathbf{d}\alpha.$$

However, since $\mathbf{d}\alpha(u, v) = u[\alpha(v)] - v[\alpha(u)] - \alpha([u, v])$, this formula becomes

$$\begin{aligned} \{F, H\}(\omega) &= \int_D \omega \left(\frac{\delta F}{\delta \omega}, \frac{\delta H}{\delta \omega} \right) \mu - \int_D \frac{\delta F}{\delta \omega} \left[\alpha \left(\frac{\delta H}{\delta \omega} \right) \right] \mu \\ &\quad + \int_D \frac{\delta H}{\delta \omega} \left[\alpha \left(\frac{\delta F}{\delta \omega} \right) \right] \mu. \end{aligned}$$

The following argument shows that the last two terms vanish. For $u \in$

$\mathfrak{X}_{\text{div}}(D)$ and $f \in \mathcal{F}(D)$ we have by the Stokes theorem

$$\int_D u[f]\mu = \int_D \mathcal{L}_u(f\mu) = \int_D \mathbf{d}\mathbf{i}_u(f\mu) = \int_{\partial D} \mathbf{i}_u(f\mu) = \int_{\partial D} f\mathbf{i}_u\mu.$$

The definition of the boundary volume form $\mu_{\partial D}$ induced by the volume form μ implies that $\mathbf{i}_u\mu = g(u, n)\mu_{\partial D}$, where n is the outward unit normal to the boundary ∂D (see e.g. [AMR88] §7.2). Since u is tangent to the boundary, this term is zero. Thus the plus Lie-Poisson bracket has the expression

$$\{F, H\}(\omega) = \int_D \omega \left(\frac{\delta F}{\delta \omega}, \frac{\delta H}{\delta \omega} \right) \mu.$$

The term “vorticity representation” appears because of the following interpretation of the variable ω . Define the **vorticity** of a vector field u by $\omega_u := \mathbf{d}u^\flat$, where $u^\flat := g(u, \cdot) \in \Omega^1(D)$. Regard now $u \mapsto \omega_u$ as a change of variables to be implemented in the velocity representation of the Lie-Poisson bracket. To this end, if $F : \mathbf{d}\Omega^1(D) \rightarrow \mathbb{R}$, define $\tilde{F} : \mathfrak{X}_{\text{div}}(D) \rightarrow \mathbb{R}$ by $\tilde{F}(u) := F(\omega_u)$. Then, if $\delta\omega := \mathbf{d}(\delta u^\flat)$, we get by the chain rule

$$\begin{aligned} \int_D g \left(\frac{\delta \tilde{F}}{\delta u}, \delta u \right) \mu &= \mathbf{D}\tilde{F}(u) \cdot \delta u = \mathbf{D}F(\omega) \cdot \delta\omega \\ &= \int_D \delta u^\flat \left(\frac{\delta F}{\delta \omega} \right) \mu = \int_D g \left(\delta u, \frac{\delta F}{\delta \omega} \right) \mu \end{aligned}$$

and thus $\delta\tilde{F}/\delta u = \delta F/\delta\omega$ which shows that the vorticity representation of the Lie-Poisson bracket is obtained from the velocity representation of the Lie-Poisson bracket by the linear change of variables $u \mapsto \omega_u$.

7. **Poisson-Vlasov Bracket.** For a Poisson manifold $(P, \{, \}_P)$ endowed with a volume form μ , the algebra $\mathcal{F}(P)$ is also a Lie algebra with Lie bracket the Poisson bracket. Consider the weak pairing between $\mathcal{F}(P)$ and the smooth densities $\mathcal{F}(P)^*$ on P given by

$$\langle \varphi, f \rangle = \int_P \varphi \bar{f} \mu, \quad \text{where} \quad \bar{f}\mu := f \in \mathcal{F}(P)^*, \quad \bar{f} \in \mathcal{F}(P).$$

The plus Lie-Poisson bracket in $\mathcal{F}(P)^*$ is thus given by

$$\{F, G\}(f) = \int_P f \left\{ \frac{\delta F}{\delta f}, \frac{\delta G}{\delta f} \right\}_P.$$

This formal Poisson bracket is known as the Poisson-Vlasov bracket since the Poisson-Vlasov equations form a Hamiltonian system for it if $P = T^*\mathbb{R}^3$

and the Hamiltonian is given by

$$H(f) = \frac{1}{2} \iint \|v\|^2 f(x, v, t) d^3x d^3v + \frac{1}{2m} \int \|\nabla \phi_f\|^2 d^3x.$$

This system describes the motion of a collisionless plasma consisting of a single species of particles with mass m and charge e in the electrostatic limit (that is, one lets the speed of light $c \rightarrow \infty$). The physical significance of $f \geq 0$ is the plasma density that depends on the position $x \in \mathbb{R}^3$, the velocity $v \in \mathbb{R}^3$, and evolves in time $t \in \mathbb{R}$. The charge density is given by $\rho_f(x) := e \int f(x, v) d^3v$ and the electric potential $\phi_f(x)$ by the Poisson equation $-\Delta \phi_f = \rho_f$.

Let us carry out a formal computation to determine $\delta H / \delta f$ assuming that correct decay at infinity conditions are put on the relevant functions so that all integration by parts below are justified. We have

$$\begin{aligned} \left\langle \frac{\delta H}{\delta f}, \delta f \right\rangle &= \iint \|v\|^2 \delta f(x, v) d^3x d^3v + \frac{1}{m} \int \nabla \phi_f(x) \cdot \nabla \phi_{\delta f}(x) d^3x \\ &= \iint \|v\|^2 \delta f(x, v) d^3x d^3v - \frac{1}{m} \int (\Delta \phi_{\delta f}(x)) \phi_f(x) d^3x \\ &= \iint \|v\|^2 \delta f(x, v) d^3x d^3v + \frac{1}{m} \int \rho_{\delta f}(x) \phi_f(x) d^3x \\ &= \iint \|v\|^2 \delta f(x, v) d^3x d^3v + \frac{e}{m} \iint \delta f(x, v) \phi_f(x) d^3x d^3v, \end{aligned}$$

which shows that

$$\frac{\delta H}{\delta f} = \|v\|^2 + \frac{e}{m} \phi_f.$$

Thus Hamilton's equations $\dot{F} = \{F, H\}$ for an arbitrary functional F of f become $\dot{f} + \{f, \frac{\delta H}{\delta f}\} = 0$ (where the Poisson bracket is now the one on $T^*\mathbb{R}^3$). Replacing here the formula for $\delta H / \delta f$ just found yields the Poisson-Vlasov equations

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f - \frac{e}{m} \nabla_x \phi_f \cdot \nabla_v f = 0.$$

8. **Korteweg-de Vries Bracket.** Let $\mathcal{F}(\mathbb{R})$ be a space of smooth functions on $\mathbb{R}$ that satisfy together with their derivatives all necessary decay conditions at infinity guaranteeing that all integrals as well as the integrations by parts that will be carried out below make sense and the boundary terms appearing in these computations vanish.

The KdV bracket on $\mathcal{F}(\mathbb{R})$ is given by

$$\{F, G\}(u) = \int_{-\infty}^{+\infty} \frac{\delta F}{\delta u} \frac{d}{dx} \left(\frac{\delta G}{\delta u} \right) dx,$$

where the functional derivatives are taken relative to the L^2 product on real valued functions. Let us work out Hamilton's equations $\dot{F} = \{F, H\}$ for this bracket. We have

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{\delta F}{\delta u} \dot{u} dx &= \mathbf{D}F(u(t)) \cdot \dot{u}(t) = \frac{d}{dt} F(u(t)) = \{F, H\}(u(t)) \\ &= \int_{-\infty}^{+\infty} \frac{\delta F}{\delta u} \frac{d}{dx} \frac{\delta H}{\delta u} dx, \end{aligned}$$

so Hamilton's equation are

$$u_t = \frac{d}{dx} \left(\frac{\delta H}{\delta u} \right) = \left(\frac{\delta H}{\delta u} \right)_x,$$

where $u_t := \frac{\partial u}{\partial t}$ and $u_x := \frac{\partial u}{\partial x}$.

In particular, if we take the Hamiltonian function

$$H_1 = -\frac{1}{6} \int_{-\infty}^{\infty} u^3 dx,$$

Hamilton's equations for the KdV bracket become the one dimensional transport equation $u_t + uu_x = 0$. Indeed, since

$$\int_{-\infty}^{+\infty} \frac{\delta H_1}{\delta u} \delta u dx = \mathbf{D}H_1(u) \cdot \delta u = -\frac{1}{2} \int_{-\infty}^{+\infty} u^2 \delta u dx,$$

it follows that $\frac{\delta H_1}{\delta u} = -\frac{1}{2} u^2$ and so the equation of motion is

$$u_t = \frac{\partial}{\partial x} \left(-\frac{1}{2} u^2 \right) = -u u_x.$$

If one takes the Hamiltonian equal to

$$H_2(u) = \int_{-\infty}^{\infty} \left(\frac{1}{2} u_x^2 - u^3 \right) dx$$

then $\frac{\delta H_2}{\delta u} = -u_{xx} - 3u^2$ and Hamilton's equation is the Korteweg-de Vries (KdV) equation

$$u_t + 6u u_x + u_{xxx} = 0. \quad (4.10)$$

The KdV equation has infinitely many independent integrals F_i in involution (that is, $\{F_i, F_j\} = 0$) and is a completely integrable system in infinite

dimensions. Here are the first integrals:

$$F_0(u) = \int_{-\infty}^{+\infty} u(x) dx$$

$$F_1(u) = \frac{1}{2} \int_{-\infty}^{+\infty} u^2(x) dx$$

$$F_2(u) = H_2(x) = \int_{-\infty}^{+\infty} \left(-u^3(x) + \frac{1}{2} u_x^2(x) \right) dx$$

$$F_3(u) = \int_{-\infty}^{+\infty} \left(\frac{1}{2} u_{xxx}^2(x) - 5u(x)u_x^2(x) + \frac{5}{2} u_x^4(x) \right) dx.$$

The existence of such integrals is believed to be closely related to the presence of solitons, that is, “solitary waves which interact pairwise by passing through each other without changing shape” (see e.g. [AbMa78] and references therein).

Let us look for traveling wave solutions of the KdV equation (4.10), that is, solutions of the form $u(t, x) = \phi(x - ct)$, for $c > 0$ and $\phi \geq 0$. Substituting into (4.10) we get

$$c\phi' - 6\phi\phi' - \phi''' = 0.$$

Integrating once this equation gives

$$c\phi - 3\phi^2 - \phi'' = C, \quad (4.11)$$

where $C \in \mathbb{R}$ is a constant. This equation is equivalent to

$$\begin{cases} \frac{d\phi}{dx} = \phi' = \frac{\partial h}{\partial \phi'} \\ \frac{d\phi'}{dx} = \phi'' = c\phi - 3\phi^2 - C = -\frac{\partial h}{\partial \phi} \end{cases} \quad (4.12)$$

where

$$h(\phi, \phi') = \frac{1}{2}(\phi')^2 - \frac{c}{2}\phi^2 + \phi^3 + C\phi. \quad (4.13)$$

Thus (4.12) is Hamiltonian in the variables $(\phi, \phi') \in \mathbb{R}^2$ with Hamiltonian function (4.13). In particular, h can be viewed to be of the form kinetic energy $K(\phi, \phi') = \frac{1}{2}(\phi')^2$ plus potential energy $V(\phi) = \frac{-c}{2}\phi^2 + \phi^3 + C\phi$. Since energy is conserved, we have $h(\phi, \phi') = D$, for some constant $D \in \mathbb{R}$, which implies

$$\phi' = \pm \sqrt{c\phi^2 - 2\phi^3 + 2C\phi + 2D}.$$

Integrating we have

$$s = \pm \int \frac{d\phi}{\sqrt{c\phi^2 - 2\phi^3 + 2C\phi + 2D}}$$

where $s = x - ct$.

We seek solutions which together with their derivatives vanish at $\pm\infty$. Then $D = 0$ by (4.13) and $C = 0$ by (4.11). Thus we get

$$s = \pm \int \frac{d\phi}{\sqrt{c\phi^2 - \phi^3}} = \pm \frac{1}{\sqrt{c}} \log \left| \frac{\sqrt{c - 2\phi} - \sqrt{c}}{\sqrt{c - 2\phi} + \sqrt{c}} \right| + K. \quad (4.14)$$

The Hamiltonian system (4.12) has two equilibria when $C = D = 0$, namely $(\phi_e, \phi'_e) = (0, 0)$ and $(\phi_e, \phi'_e) = (c/3, 0)$. The matrix of the linearized system at $(\phi_e, \phi'_e) = (0, 0)$ and $(\phi_e, \phi'_e) = (c/3, 0)$ is

$$\begin{bmatrix} 0 & 1 \\ c & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 \\ -c & 0 \end{bmatrix},$$

respectively. So $(0, 0)$ is a saddle point while $(c/3, 0)$ is spectrally stable (the eigenvalues are on the imaginary axis). To see if $(c/3, 0)$ is stable or not we can use the second variation of the potential energy criterion. Since

$$\frac{\delta^2 V}{\delta \phi^2}(c/3, 0) = c > 0$$

it follows that $(c/3, 0)$ is a Lyapunov stable point.

Consider $(\phi(s), \phi'(s))$ to be a homoclinic orbit emanating and ending at $(0, 0)$ to which $(c/3, 0)$ belongs. Both equilibria belong to the zero level set of the energy and the homoclinic orbit is given by (4.14). Furthermore when $C = 0$ we have $h(c/2, 0) = 0$ that is $(c/2, 0)$ belongs also to this homoclinic orbit. Let us take $(c/2, 0)$ as initial condition $(\phi(0), \phi'(0))$. Then by (4.14) we get $K = 0$ and the homoclinic orbit is given

$$\pm \sqrt{c}s = \log \left| \frac{\sqrt{c - 2\phi(s)} - \sqrt{c}}{\sqrt{c - 2\phi(s)} + \sqrt{c}} \right|.$$

As $\phi > 0$ the value inside the modulus is negative and the homoclinic orbit has the expression

$$\begin{aligned} e^{\pm \sqrt{c}s} &= - \frac{\sqrt{c - 2\phi(s)} - \sqrt{c}}{\sqrt{c - 2\phi(s)} + \sqrt{c}} \iff \phi(s) = \frac{2ce^{\pm \sqrt{c}s}}{(1 + e^{\pm \sqrt{c}s})^2} \\ &= \frac{c}{2} \operatorname{sech}^2(\sqrt{c}s/2), \end{aligned}$$

which gives the soliton: $u(x, t) = \phi(x - ct) = \frac{c}{2} \operatorname{sech}^2(\frac{\sqrt{c}}{2}(x - ct))$.

9. **Operator Algebra Brackets.** This example is taken from the papers [Bo00] and [OdzRa03]. Let $\mathcal{H}$ be a complex Hilbert space. Denote by $\mathfrak{S}(\mathcal{H})$, $\mathfrak{H}\mathfrak{S}(\mathcal{H})$, and $\mathfrak{B}(\mathcal{H})$ the involutive Banach algebras of the trace class operators, the Hilbert-Schmidt operators, and the bounded operators on $\mathcal{H}$,

respectively. Recall that $\mathfrak{S}(\mathcal{H})$ and $\mathfrak{H}\mathfrak{S}(\mathcal{H})$ are self adjoint ideals in $\mathfrak{B}(\mathcal{H})$. Let $\mathfrak{K}(\mathcal{H}) \subset \mathfrak{B}(\mathcal{H})$ denote the ideal of all compact operators on $\mathcal{H}$. Then

$$\mathfrak{S}(\mathcal{H}) \subset \mathfrak{H}\mathfrak{S}(\mathcal{H}) \subset \mathfrak{K}(\mathcal{H}) \subset \mathfrak{B}(\mathcal{H})$$

and the following remarkable dualities hold:

$$\mathfrak{K}(\mathcal{H})^* \cong \mathfrak{S}(\mathcal{H}), \quad \mathfrak{H}\mathfrak{S}(\mathcal{H})^* \cong \mathfrak{H}\mathfrak{S}(\mathcal{H}), \quad \text{and} \quad \mathfrak{S}(\mathcal{H})^* \cong \mathfrak{B}(\mathcal{H});$$

the right hand sides are all Banach Lie algebras. These dualities are implemented by the strongly nondegenerate pairing

$$\langle x, \rho \rangle = \text{trace}(x\rho)$$

where $x \in \mathfrak{S}(\mathcal{H})$, $\rho \in \mathfrak{K}(\mathcal{H})$ for the first isomorphism, $\rho, x \in \mathfrak{H}\mathfrak{S}(\mathcal{H})$ for the second isomorphism, and $x \in \mathfrak{B}(\mathcal{H})$, $\rho \in \mathfrak{S}(\mathcal{H})$ for the third isomorphism. Thus condition (4.5) holds and hence the Banach spaces $\mathfrak{S}(\mathcal{H})$, $\mathfrak{H}\mathfrak{S}(\mathcal{H})$, and $\mathfrak{K}(\mathcal{H})$ are Banach Lie-Poisson spaces in a rigorous functional analytic sense (see the discussion at the end of §4.2). The Lie-Poisson bracket (4.6) becomes in this case

$$\{F, H\}(\rho) = \pm \text{trace}([\mathbf{D}F(\rho), \mathbf{D}H(\rho)]\rho) \quad (4.15)$$

where ρ is an element of $\mathfrak{S}(\mathcal{H})$, $\mathfrak{H}\mathfrak{S}(\mathcal{H})$, or $\mathfrak{K}(\mathcal{H})$, respectively. The bracket $[\mathbf{D}F(\rho), \mathbf{D}H(\rho)]$ denotes the commutator bracket of operators. The Hamiltonian vector field associated to H is given by

$$X_H(\rho) = \pm[\mathbf{D}H(\rho), \rho]. \quad (4.16)$$

4.4 Generalities on Lie-Poisson Structures

We shall collect here some of the most important properties of Lie-Poisson structures. Let $\mathfrak{g}^*$ be the dual of the finite dimensional Lie algebra $\mathfrak{g}$ of a Lie group G , $\text{ad}_\xi \eta := [\xi, \eta]$ for $\xi, \eta \in \mathfrak{g}$, and $[\cdot, \cdot]$ the Lie bracket on $\mathfrak{g}$.

Proposition 4.9 *The equations of motion for the Hamiltonian $H : \mathfrak{g}^* \rightarrow \mathbb{R}$ with respect to the $(\pm)$ Lie-Poisson bracket on $\mathfrak{g}^*$ are*

$$\frac{d\mu}{dt} = X_H(\mu) = \mp \text{ad}_{\delta H / \delta \mu}^* \mu. \quad (4.17)$$

Proof For an arbitrary function $F \in \mathcal{F}(\mathfrak{g}^*)$ and $\mu \in \mathfrak{g}^*$ we have:

$$\frac{dF}{dt} = \mathbf{D}F(\mu) \cdot \dot{\mu} = \left\langle \frac{\delta F}{\delta \mu}, \dot{\mu} \right\rangle. \quad (4.18)$$

On the other hand,

$$\begin{aligned}\{F, H\}_{\pm}(\mu) &= \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right] \right\rangle = \pm \left\langle \mu, -\text{ad}_{\delta H / \delta \mu} \frac{\delta F}{\delta \mu} \right\rangle \\ &= \mp \left\langle \text{ad}_{\delta H / \delta \mu}^* \mu, \frac{\delta F}{\delta \mu} \right\rangle.\end{aligned}$$

This last equality and (4.18) gives the result. $\blacksquare$

Let $\{\xi_a\}$, $a = 1, 2, \dots, n$, be a basis of $\mathfrak{g}$ and $\{\xi^a\}$ its dual basis. The structure constants C_{ab}^d of $\mathfrak{g}$ are defined by $[\xi_a, \xi_b] = \sum_d C_{ab}^d \xi_d$.

For $\mu = \sum_a \mu_a \xi^a$ the Lie-Poisson brackets become

$$\{F, G\}_{\pm} = \pm \mu_d \frac{\delta F}{\delta \mu_a} \frac{\delta G}{\delta \mu_b} C_{ab}^d,$$

where summation on repeated indices is understood. In particular,

$$\{\mu_a, \mu_b\}_{\pm} = \pm \mu_d C_{ab}^d.$$

So the equations of motion for H are

$$\dot{\mu}_a = \mp \mu_d C_{ab}^d \frac{\delta H}{\delta \mu_b}.$$

Next we study linear Poisson maps.

Proposition 4.10 *Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras and $\alpha : \mathfrak{g} \rightarrow \mathfrak{h}$ a linear map. Then α is a homomorphism of Lie algebras if and only if its dual $\alpha^* : \mathfrak{h}^* \rightarrow \mathfrak{g}^*$ is a Poisson map.*

Proof By definition of the Lie-Poisson bracket on $\mathfrak{h}^*$ we have

$$\{F \circ \alpha^*, H \circ \alpha^*\}_{\pm}(\mu) = \pm \left\langle \mu, \left[\frac{\delta(F \circ \alpha^*)}{\delta \mu}, \frac{\delta(H \circ \alpha^*)}{\delta \mu} \right] \right\rangle \quad (4.19)$$

for any $F, H \in \mathcal{F}(\mathfrak{g}^*)$, $\mu \in \mathfrak{h}$.

By definition of the functional derivative, the chain rule, and the definition of the dual map one has

$$\begin{aligned}\left\langle \frac{\delta(F \circ \alpha^*)}{\delta \mu}, \delta \mu \right\rangle &= \mathbf{D}(F \circ \alpha^*) \cdot \delta \mu = \mathbf{D}F(\alpha^*(\mu)) \cdot \alpha^*(\delta \mu) \\ &= \left\langle \frac{\delta F}{\delta \nu}, \alpha^*(\delta \mu) \right\rangle = \left\langle \alpha \left(\frac{\delta F}{\delta \nu} \right), \delta \mu \right\rangle,\end{aligned}$$

where $\nu := \alpha^*(\mu) \in \mathfrak{g}^*$. Thus equation (4.19) becomes

$$\begin{aligned} \{F \circ \alpha^*, H \circ \alpha^*\}_\pm(\mu) &= \pm \left\langle \mu, \left[\frac{\delta(F \circ \alpha^*)}{\delta \mu}, \frac{\delta(H \circ \alpha^*)}{\delta \mu} \right] \right\rangle \\ &= \pm \left\langle \mu, \left[\alpha \left(\frac{\delta F}{\delta \nu} \right), \alpha \left(\frac{\delta H}{\delta \nu} \right) \right] \right\rangle. \end{aligned} \quad (4.20)$$

If α is a Lie algebra homomorphism then (4.20) is equal to

$$\pm \left\langle \mu, \alpha \left(\left[\frac{\delta F}{\delta \nu}, \frac{\delta H}{\delta \nu} \right] \right) \right\rangle = \pm \left\langle \alpha^*(\mu), \left[\frac{\delta F}{\delta \nu}, \frac{\delta H}{\delta \nu} \right] \right\rangle = \{F, H\}_\pm(\alpha^*(\mu)),$$

which shows that α^* is a Poisson map.

Conversely, if α^* is a Poisson map, by (4.20) we get

$$\begin{aligned} \pm \left\langle \mu, \left[\alpha \left(\frac{\delta F}{\delta \nu} \right), \alpha \left(\frac{\delta H}{\delta \nu} \right) \right] \right\rangle &= \{F, H\}_\pm(\alpha^*(\mu)) \\ &= \pm \left\langle \alpha^*(\mu), \left[\frac{\delta F}{\delta \nu}, \frac{\delta H}{\delta \nu} \right] \right\rangle = \pm \left\langle \mu, \alpha \left(\left[\frac{\delta F}{\delta \nu}, \frac{\delta H}{\delta \nu} \right] \right) \right\rangle \end{aligned}$$

for any $F, H \in \mathcal{F}(\mathfrak{g}^*)$. In particular, taking $F(\rho) := \langle \rho, \xi \rangle$ and $H(\rho) := \langle \rho, \eta \rangle$ for arbitrary $\xi, \eta \in \mathfrak{g}$, we get $\langle \mu, [\alpha(\xi), \alpha(\eta)] \rangle = \langle \mu, \alpha([\xi, \eta]) \rangle$ for any $\mu \in \mathfrak{h}$, which implies $[\alpha(\xi), \alpha(\eta)] = \alpha([\xi, \eta])$, that is, α is a Lie algebra homomorphism. $\blacksquare$

The last key property of Lie-Poisson brackets is that they are the only linear ones in the following sense. Consider a finite dimensional vector space V , V^* its dual, and let $\langle \cdot, \cdot \rangle$ be the duality pairing between V^* and V . One can then think of elements of V to be the linear functionals on V^* . A Poisson bracket in V^* is said to be **linear** if the bracket of two any linear functionals on V^* is a linear functional. Thus, if X', Y' are functionals on V^* then, as the pairing $\langle \cdot, \cdot \rangle$ is nondegenerate, there exist unique elements $X, Y \in V$ such that

$$X'(\mu) = \langle \mu, X \rangle \quad \text{and} \quad Y'(\mu) = \langle \mu, Y \rangle.$$

The linearity assumption on the Poisson bracket in V^* implies the existence of a unique element of V , say $[X, Y]$, such that

$$\{X', Y'\}(\mu) = [X, Y]'(\mu) = \langle \mu, [X, Y] \rangle.$$

It is easy to prove that $[\cdot, \cdot]$ so defined on V is a Lie algebra bracket and so the given linear Poisson bracket is the Lie-Poisson bracket for the Lie algebra V .

In the infinite dimensional case the same proof works if $\langle \cdot, \cdot \rangle : V^* \times V \rightarrow \mathbb{R}$ is a weak pairing between Banach spaces and one makes the extra hypothesis that the Poisson bracket of any two linear functionals on V^* belongs to the range of the Poisson tensor $\Lambda(\mu) : V \rightarrow V^*$ for all $\mu \in V^*$.

Proposition 4.11 *Let V and V^* be two Banach spaces, $\langle \cdot, \cdot \rangle$ a weak nondegenerate pairing of V^* with V , and assume that V^* has a linear Poisson bracket. If the Poisson bracket of any two linear functionals in V^* belongs to the range of $\langle \mu, \cdot \rangle$ for all $\mu \in V^*$, then V is a Banach Lie algebra and the Poisson bracket on V^* is the corresponding Lie-Poisson bracket.*

5 Momentum Maps

In this lecture we shall introduce the concept of momentum map and study its properties. We shall address the issue of existence and equivariance of momentum maps, give an explicit formula for the case of a cotangent bundle, and present several basic examples. The full power of the momentum map will appear only in the next chapter when dealing with reduction.

5.1 Actions and Infinitesimal Generators

Let $\Phi : G \times P \rightarrow P$ be a smooth left action of the Lie group G on the Poisson manifold P . The action is **canonical** if the map $\Phi_g : z \mapsto \Phi(g, z) = \Phi_g(z)$ is a Poisson map for all $g \in P$, that is

$$\{F \circ \Phi_g, G \circ \Phi_g\} = \{F, G\} \circ \Phi_g \quad \text{for all } F, G \in \mathcal{F}(P). \quad (5.1)$$

Denote by $\mathfrak{g}$ the Lie algebra of G . Let us recall the following standard facts.

- (i) $\text{Ad}_g : G \rightarrow G$ given for each $g \in G$ by

$$\text{Ad}_g(h) = ghg^{-1} = (L_g \circ R_{g^{-1}})(h) \quad (5.2)$$

is an inner automorphism. L_g and R_g denote, respectively, the left and right translations of G on itself.

- (ii) Differentiating Ad_g with respect to h at $h = e$ we get the **adjoint representation** $\text{Ad}_g = T_{g^{-1}}L_g \circ T_eR_{g^{-1}} : \mathfrak{g} \rightarrow \mathfrak{g}$ of G on $\mathfrak{g}$. The inverse of the dual map defines the **coadjoint representation** $\text{Ad}_{g^{-1}}^*$ of G on $\mathfrak{g}^*$.
- (iii) Differentiating Ad_g with respect to g at e in the direction ξ yields

$$T_e\text{Ad}(\xi) = \left. \frac{d}{dt} \right|_{t=0} \text{Ad}_{\exp t\xi} = [\xi, \cdot] =: \text{ad}_\xi : \mathfrak{g} \rightarrow \mathfrak{g}, \quad (5.3)$$

the **adjoint representation** of $\mathfrak{g}$ on $\mathfrak{g}$.

- (iv) In these lectures, all Lie algebras are *left* Lie algebras. This means that $\mathfrak{g} = T_eG$ as a vector space and the Lie bracket on $\mathfrak{g}$ is given by the identity $[\xi, \eta] = [X_\xi, X_\eta](e)$, where $X_\xi, X_\eta \in \mathfrak{X}(G)$ are the left

invariant vector fields whose values at the identity are ξ and η , respectively, that is, $X_\xi(g) = T_e L_g(\xi)$ and $X_\eta(g) = T_e L_g(\eta)$ for any $g \in G$. The vector fields X_ξ are complete, that is, each integral curve exists for all time.

- (v) If $\xi \in \mathfrak{g}$, there is a unique integral curve $\gamma_\xi(t)$ of X_ξ with initial condition $\gamma_\xi(0) = e$. Then $\gamma_\xi(s+t) = \gamma_\xi(s)\gamma_\xi(t)$ for any $s, t \in \mathbb{R}$, that is, $\gamma_\xi : \mathbb{R} \rightarrow G$ is a **one-parameter subgroup** of G . Conversely, any Lie group homomorphism $\gamma : \mathbb{R} \rightarrow G$ is of the form γ_ξ , where $\xi = \gamma'(0)$. The **exponential map** is defined by $\exp(\xi) := \gamma_\xi(1)$. Then $\exp(t\xi) = \gamma_\xi(t)$ and the flow of X_ξ is given by $(t, g) \mapsto g \exp(t\xi)$.
- (vi) If $F : G \rightarrow H$ is a homomorphism of Lie groups then $T_e F : \mathfrak{g} \rightarrow \mathfrak{h}$ is a Lie algebra homomorphism, that is,

$$T_e F([\xi, \eta]) = [T_e F(\xi), T_e F(\eta)]. \quad (5.4)$$

In particular, $\text{Ad}_g : \mathfrak{g} \rightarrow \mathfrak{g}$ is a Lie algebra isomorphism for every $g \in G$. In addition,

$$\exp_H \circ T_e F = F \circ \exp_G, \quad (5.5)$$

where $\exp_G : \mathfrak{g} \rightarrow G$ and $\exp_H : \mathfrak{h} \rightarrow H$ are the exponential maps. If one takes $G = H$ and $F = \text{Ad}_g$, this identity becomes

$$\exp(\text{Ad}_g \xi) = \text{Ad}_g(\exp \xi) = g \exp(\xi) g^{-1}. \quad (5.6)$$

If one takes $H = \text{Iso}(\mathfrak{g})$, the Lie group of Lie algebra isomorphisms of $\mathfrak{g}$, then its Lie algebra is the Lie algebra $\text{Der}(\mathfrak{g})$ of derivations relative to the bracket $[\cdot, \cdot]$. Choosing in (5.5) $F = \text{Ad} : G \rightarrow H = \text{Iso}(\mathfrak{g})$, we get for any $\xi \in \mathfrak{g}$

$$e^{\text{ad}_\xi} = \text{Ad}_{\exp \xi}. \quad (5.7)$$

- (vii) Given a Lie algebra element $\xi \in \mathfrak{g}$, $\exp(t\xi)$ defines a one-parameter subgroup in G and hence $t \mapsto \Phi_{\exp t\xi}$ is a flow on the manifold P . The vector field defined by this flow is denoted by ξ_P and is called the **infinitesimal generator** of the action determined by ξ . Thus we have

$$\xi_P(z) = \left. \frac{d}{dt} \right|_{t=0} \Phi_{\exp(t\xi)}(z).$$

The infinitesimal generator has the following properties

$$(\text{Ad}_g \xi)_P = \Phi_{g^{-1}}^* \xi_P \quad \text{and} \quad [\xi_P, \eta_P] = -[\xi, \eta]_P \quad (5.8)$$

for any $g \in G$ and $\xi, \eta \in \mathfrak{g}$.

Examples 1. We deduce the formulas stated in §1.2. Let $G = SO(3)$ acting on $\mathbb{R}^3$ by matrix multiplication. The Lie algebra $\mathfrak{so}(3)$ is the set of 3×3 skew symmetric matrices with Lie bracket the commutator. It is isomorphic to $(\mathbb{R}^3, \times)$ via the map $\mathbf{u} \in \mathbb{R}^3 \mapsto \hat{\mathbf{u}} \in \mathfrak{so}(3)$ given by (1.13). The adjoint action is hence

$$\text{Ad}_A \hat{\mathbf{u}} = A \hat{\mathbf{u}} A^{-1} = (A\mathbf{u})^\wedge.$$

Therefore,

$$[\hat{\mathbf{u}}, \hat{\mathbf{v}}] = \text{ad}_{\hat{\mathbf{u}}} \hat{\mathbf{v}} = \left. \frac{d}{dt} \right|_{t=0} (e^{t\hat{\mathbf{u}}} \hat{\mathbf{v}})^\wedge = (\hat{\mathbf{u}} \hat{\mathbf{v}})^\wedge = (\mathbf{u} \times \mathbf{v})^\wedge.$$

The dual $\mathfrak{so}(3)^*$ is identified with $\mathbb{R}^3$ by the isomorphism $\mathbf{\Pi} \in \mathbb{R}^3 \mapsto \tilde{\mathbf{\Pi}} \in \mathfrak{so}(3)^*$ given by $\tilde{\mathbf{\Pi}}(\hat{\mathbf{u}}) := \mathbf{\Pi} \cdot \mathbf{u}$ for any $\mathbf{u} \in \mathbb{R}^3$. Then the coadjoint action of $SO(3)$ on $\mathfrak{so}(3)^*$ is given by

$$\begin{aligned} (\text{Ad}_{A^{-1}}^* \tilde{\mathbf{\Pi}})(\hat{\mathbf{u}}) &= \tilde{\mathbf{\Pi}} \cdot \text{Ad}_{A^{-1}} \hat{\mathbf{u}} = \tilde{\mathbf{\Pi}} \cdot (A^{-1} \mathbf{u})^\wedge = \mathbf{\Pi} \cdot A^T \mathbf{u} \\ &= A \mathbf{\Pi} \cdot \mathbf{u} = (A \mathbf{\Pi})^\sim(\hat{\mathbf{u}}), \end{aligned}$$

that is, $\text{Ad}_{A^{-1}}^* \tilde{\mathbf{\Pi}} = (A \mathbf{\Pi})^\sim$, thereby recovering formula (1.20) in §1.2.

The infinitesimal generator corresponding to $\mathbf{u} \in \mathbb{R}^3$ has the expression

$$\mathbf{u}_{\mathbb{R}^3}(\mathbf{x}) = \left. \frac{d}{dt} \right|_{t=0} e^{t\hat{\mathbf{u}}} \mathbf{x} = \hat{\mathbf{u}} \mathbf{x} = \mathbf{u} \times \mathbf{x}. \quad (5.9)$$

2. Let $G = \text{Diff}_{\text{vol}}(D)$ be the group of volume preserving diffeomorphisms of the oriented Riemannian manifold (D, g) . On D there is a unique volume form μ which equals 1 on all positively oriented g -orthonormal bases of tangent vectors at all points of D ; this volume form μ is called the **Riemannian volume** of (D, g) and we shall assume from now on that the orientation of D is given by μ .

Let us show formally that the Lie algebra of $\text{Diff}_{\text{vol}}(D)$ is $\mathfrak{X}_{\text{div}}(D)$ endowed with the negative of the bracket of vector fields. First, as a vector space, the Lie algebra of the group $\text{Diff}(D)$ equals the space $\mathfrak{X}(D)$ of vector fields on D . Indeed, the flow of an arbitrary vector field is a smooth path in $\text{Diff}(D)$ whose tangent vector at time equal to zero is the given vector field.

Second, if $\eta_t \in \text{Diff}_{\text{vol}}(D)$ is the flow of the vector field v , then $\eta_t^* \mu = \mu$, so taking the derivative of this identity at $t = 0$ yields $(\text{div } v) \mu = \mathcal{L}_v \mu = 0$, that is, $v \in \mathfrak{X}_{\text{div}}(D)$.

Third, since $\text{Ad}_\eta \varphi = \eta \circ \varphi \circ \eta^{-1}$, letting φ_t be the flow of v , we get

$$\begin{aligned} \text{Ad}_\eta v &= T_e \text{Ad}_\eta(v) = \left. \frac{d}{dt} \right|_{t=0} \text{Ad}_\eta \varphi_t = \left. \frac{d}{dt} \right|_{t=0} (\eta \circ \varphi_t \circ \eta^{-1}) \\ &= T\eta \circ v \circ \eta^{-1} = \eta_* v. \end{aligned}$$

Fourth, if $u, v \in \text{Diff}_{\text{vol}}(D)$ and φ_t is the flow of u , the Lie algebra bracket of u and v in $\mathfrak{X}_{\text{div}}(D)$ is given by

$$\left. \frac{d}{dt} \right|_{t=0} \text{Ad}_{\varphi_t} v = \left. \frac{d}{dt} \right|_{t=0} (\varphi_t)_* v = \left. \frac{d}{dt} \right|_{t=0} (\varphi_{-t})^* v = -\mathcal{L}_u v = -[u, v].$$

Thus the left Lie algebra bracket on the space of vector fields equals the negative of the usual Jacobi-Lie bracket of vector fields.

Identify the dual $\mathfrak{X}_{\text{div}}(D)^*$ with $\mathfrak{d}\Omega^1(D)$ (assuming that the first cohomology group of D is zero). The coadjoint action of $\text{Diff}_{\text{vol}}(D)$ on $\mathfrak{d}\Omega^1(D)$ is computed in the following way. Let $\omega = \mathfrak{d}\alpha \in \mathfrak{d}\Omega^1(D)$ and $u \in \mathfrak{X}_{\text{div}}(D)$. Then

$$\langle u, \text{Ad}_{\eta^{-1}}^* \omega \rangle = \langle \text{Ad}_{\eta^{-1}} u, \omega \rangle = \int_D \alpha(\eta^* u) \mu = \int_D (\eta_* \alpha)(u) \mu = \langle u, \eta_* \omega \rangle$$

by the change of variables formula, taking into account that the Jacobian of η is one, and noting that $\mathfrak{d}\eta_* \alpha = \eta_* \omega$. Therefore $\text{Ad}_{\eta^{-1}}^* \omega = \eta_* \omega$.

5.2 Momentum Maps

Let the Lie group G with Lie algebra $\mathfrak{g}$ act on the Poisson manifold P in a canonical way, that is, (5.1) holds. Differentiating (5.1) with respect to g at the identity in the direction of $\xi \in \mathfrak{g}$ shows that the infinitesimal generator ξ_P is an infinitesimal Poisson automorphism, i.e.,

$$\xi_P [\{F_1, F_2\}] = \{\xi_P[F_1], F_2\} + \{F_1, \xi_P[F_2]\}$$

for any $F_1, F_2 \in \mathcal{F}(P)$. Denote by $\mathcal{P}(P)$ the set of all vector fields satisfying this relation and call this Lie subalgebra of $\mathfrak{X}(P)$ the Lie algebra of **Poisson bracket derivations** or of **infinitesimal Poisson automorphisms**.

If $\xi \in \mathfrak{g}$ we ask if the infinitesimal generator ξ_P is globally Hamiltonian. That is, we seek a Hamiltonian function $J^\xi \in \mathcal{F}(P)$ such that $X_{J^\xi} = \xi_P$ for every $\xi \in \mathfrak{g}$. Since the right hand side of this equation is linear in ξ , we shall require that the map $\xi \in \mathfrak{g} \mapsto J^\xi \in \mathcal{F}(P)$ be also linear.

Definition 5.1 Let G be a Lie group acting canonically on the Poisson manifold P . Suppose that there is a linear map $J : \mathfrak{g} \rightarrow \mathcal{F}(P)$ such that

$$X_{J^\xi} = \xi_P, \tag{5.10}$$

for all $\xi \in \mathfrak{g}$, where ξ_P is the infinitesimal generator corresponding to ξ for then G -action on P . Then the map $\mathbf{J} : P \rightarrow \mathfrak{g}^*$ defined by

$$\langle \mathbf{J}(z), \xi \rangle = J^\xi(z),$$

for all $\xi \in \mathfrak{g}$ and $z \in P$, is called a **momentum map** of the G -action.

One of the first questions that arise is whether or not equation (5.10) determines $\mathbf{J}$. Note that if $\mathbf{J}_1$ and $\mathbf{J}_2$ are functions verifying (5.10) then $X_{J_1^\xi - J_2^\xi} = 0$, which is equivalent with the statement that $J_1^\xi - J_2^\xi$ is a Casimir function. If P is symplectic and connected then the Casimirs are the constants and so equation (5.10) determines $\mathbf{J}$ only up to an element of $\mathfrak{g}^*$.

From the definition of $\mathbf{J}$ it follows that there is an isomorphism between the set of maps $P \rightarrow \mathfrak{g}^*$ and the set of maps $\mathfrak{g} \rightarrow \mathcal{F}(P)$. The collection of functions J^ξ as ξ varies on $\mathfrak{g}$ are the **components of the momentum map**.

To give a momentum map is therefore equivalent to specifying a linear map $J : \mathfrak{g} \rightarrow \mathcal{F}(P)$ making the following diagram

$$\begin{array}{ccc} \mathcal{F}(P) & \xrightarrow{F \mapsto X_F} & \mathcal{P}(P) \\ & \nwarrow J & \uparrow \xi \mapsto \xi_P \\ & & \mathfrak{g} \end{array}$$

commutative. Two natural questions arise:

- (A) What are the obstructions to the existence of a momentum map?
- (B) If the G -action admits a momentum map, under what conditions is it a Lie algebra homomorphism?

Let us give some answers to these questions.

(A) The map $\mathcal{H} : \mathcal{F}(P) \rightarrow \mathcal{P}(P)$ given by $F \mapsto X_F$ is a Lie algebra anti-homomorphism. Denote by $\mathcal{H}(P)$ the Lie algebra of globally Hamiltonian vector fields. The existence of a momentum map is equivalent to be able to lift the anti-homomorphism of Lie algebras $\rho : \xi \in \mathfrak{g} \mapsto \xi_P \in \mathcal{P}(P)$ through $\mathcal{H}$ to a linear map $J : \mathfrak{g} \rightarrow \mathcal{F}(P)$. So consider the following diagram where i is the inclusion and π the projection

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathcal{C}(P) & \xrightarrow{i} & \mathcal{F}(P) & \xrightarrow{\mathcal{H}} & \mathcal{P}(P) \xrightarrow{\pi} \mathcal{P}(P)/\mathcal{H}(P) \longrightarrow 0 \\ & & & & \nwarrow J & \uparrow \rho & \\ & & & & & \mathfrak{g} & \end{array}$$

If the linear map $J : \mathfrak{g} \rightarrow \mathcal{F}(P)$ is such that $\mathcal{H} \circ J = \rho$, then $\pi \circ \rho =$

$\pi \circ \mathcal{H} \circ J = 0$ by the exactness of the sequence. Conversely, if $\pi \circ \rho = 0$, then $\rho(\mathfrak{g}) \subset \mathcal{H}(P)$, that is, each ξ_P is globally Hamiltonian, so there exists a function $J^\xi \in \mathcal{F}(P)$ such that $\xi_P = X_{J^\xi}$. Requiring that $\xi \mapsto J^\xi$ be linear, yields the existence of a momentum map. So under what conditions do we have that $\pi \circ \rho = 0$?

- (i) If P is symplectic then $\mathcal{P}(P)$ coincides with the Lie algebra of locally Hamiltonian vector fields and $\mathcal{P}(P)/\mathcal{H}(P)$ is isomorphic to the first cohomology group $H^1(P)$, which is an Abelian Lie algebra. Thus, in the symplectic case, $\pi \circ \rho = 0$ if and only if the induced map $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}] \rightarrow H^1(P)$ vanishes. This happens, for instance, if $\mathfrak{g}$ is a semisimple Lie algebra because in that case, $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$.
- (ii) If $\mathcal{P}(P)/\mathcal{H}(P) = 0$ then clearly $\pi \circ \rho = 0$. If P is symplectic this is equivalent to the vanishing of the first cohomology group $H^1(P)$.
- (iii) If P is exact symplectic, i.e., the symplectic form is $\Omega = -d\Theta$, and Θ is a $\mathfrak{g}$ -invariant one-form, which means that $\mathcal{L}_{\xi_P} \Theta = 0$ for all $\xi \in \mathfrak{g}$. Indeed, $d\mathbf{i}_{\xi_P} \Theta + \mathbf{i}_{\xi_P} d\Theta = 0$, implies that $\mathbf{i}_{\xi_P} \Omega = d(\mathbf{i}_{\xi_P} \Theta)$, that is, the momentum map is given by $J^\xi = \mathbf{i}_{\xi_P} \Theta$.
- (iv) An important special case of the previous situation is $P = T^*Q$ and the G -action on P is lifted from a G -action on Q , that is,

$$g \cdot \alpha_q = T_{g \cdot q}^* \Phi_{g^{-1}}(\alpha_q),$$

for $\alpha_q \in T_q^*Q$, $g \in G$, and $\Phi : G \times Q \rightarrow Q$ an action. By theorem 2.14 a cotangent lift preserves the canonical one-form Θ on T^*Q . Therefore, by the previous case, this action admits a momentum map which is given by $\langle \mathbf{J}, \xi \rangle = \mathbf{i}_{\xi_P} \Theta$. This expression can be further simplified using (2.8) and the equivariance of the projection $\pi : T^*Q \rightarrow Q$, that is, $\pi \circ T^* \Phi_{g^{-1}} = \Phi_g \circ \pi$ for all $g \in G$. The derivative of this relation relative to g at the identity in the direction $\xi \in \mathfrak{g}$ yields $T\pi \circ \xi_P = \xi_Q \circ \pi$. Therefore

$$\begin{aligned} \langle \mathbf{J}(\alpha_q), \xi \rangle &= \mathbf{i}_{\xi_P} \Theta(\alpha_q) = \langle \Theta(\alpha_q), \xi_P(\alpha_q) \rangle = \langle \alpha_q, (T\pi \circ \xi_P)(\alpha_q) \rangle \\ &= \langle \alpha_q, (\xi_Q \circ \pi)(\alpha_q) \rangle = \langle \alpha_q, \xi_Q(q) \rangle. \end{aligned} \quad (5.11)$$

(B) To say that $J : \mathfrak{g} \rightarrow \mathcal{F}(P)$ is a Lie algebra homomorphism is equivalent to the identity

$$J^{[\xi, \eta]} = \{J^\xi, J^\eta\} \quad (5.12)$$

for all $\xi, \eta \in \mathfrak{g}$. How far are we from such a relation? To see this, note that

$$\begin{aligned} X_{J^{[\xi, \eta]}} &= [\xi, \eta]_P && \text{(by definition of } \mathbf{J}) \\ &= -[\xi_P, \eta_P] && \text{(by (5.8))} \\ &= -[X_{J^\xi}, X_{J^\eta}] && \text{(by definition of } \mathbf{J}) \\ &= X_{\{J^\xi, J^\eta\}} && (\mathcal{H} \text{ is an antihomomorphism}). \end{aligned} \quad (5.13)$$

Equation (5.13) shows that $J^{[\xi, \eta]} - \{J^\xi, J^\eta\}$ is a Casimir function, which we shall denote by $\Sigma(\xi, \eta)$. Thus $J : \mathfrak{g} \rightarrow \mathcal{F}(P)$ is a Lie algebra homomorphism if and only if $\Sigma(\xi, \eta) = 0$ for all $\xi, \eta \in \mathfrak{g}$.

The map $\Sigma : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathcal{C}(P)$ has remarkable properties, easily deduced from the definition: it is bilinear, antisymmetric, and satisfies the *cocycle identity*,

$$\Sigma(\xi, [\eta, \zeta]) + \Sigma(\eta, [\zeta, \xi]) + \Sigma(\zeta, [\xi, \eta]) = 0$$

for all $\xi, \eta, \zeta \in \mathfrak{g}$, that is, Σ is a $\mathcal{C}(P)$ -valued 2-cocycle of $\mathfrak{g}$. So J is a Lie algebra homomorphism if and only if $[\Sigma] = 0$ in $H^2(\mathfrak{g}, \mathcal{C}(P))$, the second $\mathcal{C}(P)$ -valued Lie algebra cohomology group of $\mathfrak{g}$.

When $\mathbf{J}$ verifies (5.12) we say that it is *infinitesimally equivariant*. This terminology is justified in the following way. The momentum map $\mathbf{J} : P \rightarrow \mathfrak{g}^*$ is said to be *equivariant*, if

$$\text{Ad}_{g^{-1}}^* \circ \mathbf{J} = \mathbf{J} \circ \Phi_g$$

for all $g \in G$. Pairing this relation with $\eta \in \mathfrak{g}$, putting $g = \exp t\xi$, and taking the derivative of the resulting relation at $t = 0$, yields (5.12). Thus equivariance implies infinitesimal equivariance. The converse is also true if G is connected (see [MaRa94], Theorem 12.3.2).

Here are two classes of equivariant momentum maps that appear often in applications.

- (i) The momentum map in point (A)(iii) is equivariant. Thus momentum maps of cotangent lifted actions (see (5.11)) are always equivariant. To see this, use (5.8) and G -invariance of Θ to get

$$\begin{aligned} \langle \mathbf{J}(g \cdot z), \xi \rangle &= \mathbf{i}_{\xi_P} \Theta(g \cdot z) = \Theta(g \cdot z)(\xi_P(g \cdot z)) \\ &= \Theta(g \cdot z)(T_z \Phi_g(\Phi_g^* \xi_P)(z)) \\ &= (\Phi_g^* \Theta)(z)((\text{Ad}_{g^{-1}} \xi)_P(z)) \\ &= \Theta(z)((\text{Ad}_{g^{-1}} \xi)_P(z)) = \langle \mathbf{J}(z), \text{Ad}_{g^{-1}} \xi \rangle \\ &= \langle \text{Ad}_{g^{-1}}^* \mathbf{J}(z), \xi \rangle, \end{aligned}$$

which shows that $\mathbf{J}(g \cdot z) = \text{Ad}_{g^{-1}}^* \mathbf{J}(z)$.

- (ii) For compact groups one can always choose the momentum map to be equivariant. More precisely, if the canonical G -action on the Poisson manifold P admits a momentum map $\mathbf{J} : P \rightarrow \mathfrak{g}^*$ and G is a compact Lie group, then $\mathbf{J}$ can be changed by the addition of an element in $L(\mathfrak{g}, \mathcal{C}(P))$ such that the resulting map is an equivariant momentum map for the same action. In particular, if P is symplectic, $\mathbf{J}$ can be changed by the addition of an element of $\mathfrak{g}^*$ on each connected component of P so that the resulting map is an equivariant momentum map.

To prove this statement, define for each $g \in G$

$$\mathbf{J}^g(z) := \text{Ad}_{g^{-1}}^* \mathbf{J}(g^{-1} \cdot z)$$

or equivalently,

$$(J^g)^\xi := J^{\text{Ad}_{g^{-1}} \xi} \circ \Phi_{g^{-1}}.$$

Then $\mathbf{J}^g$ is also a momentum map for the G -action on P . Indeed, if $z \in P$, $\xi \in \mathfrak{g}$, and $F : P \rightarrow \mathbb{R}$, we have by (5.8)

$$\begin{aligned} \{F, (J^g)^\xi\}(z) &= -\mathbf{d}(J^g)^\xi(z) \cdot X_F(z) \\ &= -\mathbf{d}J^{\text{Ad}_{g^{-1}} \xi}(g^{-1} \cdot z) \cdot T_z \Phi_{g^{-1}} \cdot X_F(z) \\ &= -\mathbf{d}J^{\text{Ad}_{g^{-1}} \xi}(g^{-1} \cdot z) \cdot (\Phi_g^* X_F)(g^{-1} \cdot z) \\ &= -\mathbf{d}J^{\text{Ad}_{g^{-1}} \xi}(g^{-1} \cdot z) \cdot X_{\Phi_g^* F}(g^{-1} \cdot z) \\ &= \{\Phi_g^* F, J^{\text{Ad}_{g^{-1}} \xi}\}(g^{-1} \cdot z) \\ &= X_{J^{\text{Ad}_{g^{-1}} \xi}}[\Phi_g^* F](g^{-1} \cdot z) \\ &= (\text{Ad}_{g^{-1}} \xi)_P[\Phi_g^* F](g^{-1} \cdot z) \\ &= (\Phi_g^* \xi_P)[\Phi_g^* F](g^{-1} \cdot z) \\ &= \mathbf{d}F(z) \cdot \xi_P(z) \\ &= \{F, J^\xi\}(z). \end{aligned}$$

Therefore, $\{F, (J^g)^\xi - J^\xi\} = 0$ for every $F : P \rightarrow \mathbb{R}$, that is, $(J^g)^\xi - J^\xi$ is a Casimir function on P for every $g \in G$ and every $\xi \in \mathfrak{g}$. Therefore, since $\mathbf{J}$ is a momentum map, so is $\mathbf{J}^g$ for every $g \in G$. Now define

$$\langle \mathbf{J} \rangle := \int_G \mathbf{J}^g dg$$

where dg denotes the normalized Haar measure on G , that is, the volume of G is one. Equivalently, this definition states that

$$\langle J \rangle^\xi := \int_G (J^g)^\xi dg$$

for every $\xi \in \mathfrak{g}$. By linearity of the Poisson bracket in each factor, it follows that

$$\{F, \langle J \rangle^\xi\} = \int_G \{F, (J^g)^\xi\} dg = \int_G \{F, J^\xi\} dg = \{F, J^\xi\}$$

for every $F \in \mathcal{F}(P)$. Thus $\langle J \rangle^\xi - J^\xi$ is a Casimir on P for every $\xi \in \mathfrak{g}$ which shows that $\langle \mathbf{J} \rangle - \mathbf{J} \in L(\mathfrak{g}, \mathcal{C}(P))$ and that $\langle \mathbf{J} \rangle : P \rightarrow \mathfrak{g}^*$ is also a momentum map for the G -action.

Finally we show that the momentum map $\langle \mathbf{J} \rangle$ is equivariant. Indeed, begin by noting that

$$\mathbf{J}^g(h \cdot z) = \text{Ad}_{h^{-1}}^* \mathbf{J}^{h^{-1}g}(z)$$

for every $g, h \in G$. Using invariance of the Haar measure on G under translations and inversion, we have for any $h \in G$

$$\begin{aligned} \langle \mathbf{J} \rangle(h \cdot z) &= \int_G \text{Ad}_{h^{-1}}^* \mathbf{J}^{h^{-1}g}(z) dg = \text{Ad}_{h^{-1}}^* \int_G \mathbf{J}^{h^{-1}g}(z) dg \\ &= \text{Ad}_{h^{-1}}^* \int_G \mathbf{J}^k(z) dk = \text{Ad}_{h^{-1}}^* \langle \mathbf{J} \rangle(z), \end{aligned}$$

where in the third equality we made the change of variables $g = hk$.

A crucial property of infinitesimally equivariant momentum maps is given in the following statement.

Theorem 5.2 *If $\mathbf{J}$ is an infinitesimally equivariant momentum map for the canonical G -action on the Poisson manifold P then $\mathbf{J}$ is a Poisson map, that is,*

$$\mathbf{J}^* \{F_1, F_2\}_+ = \{\mathbf{J}^* F_1, \mathbf{J}^* F_2\}$$

for all $F_1, F_2 \in \mathcal{F}(\mathfrak{g}^*)$, where $\{, \}_+$ denotes the $+$ Lie-Poisson bracket on $\mathfrak{g}^*$.

Proof For $F_1, F_2 : \mathfrak{g}^* \rightarrow \mathbb{R}$, $z \in P$, and $\mu = \mathbf{J}(z) \in \mathfrak{g}^*$, let $\xi := \frac{\delta F_1}{\delta \mu}$ and $\eta := \frac{\delta F_2}{\delta \mu}$. Then

$$\begin{aligned} \mathbf{J}^* \{F_1, F_2\}_+(z) &= \{F_1, F_2\}_+(\mathbf{J}(z)) = \left\langle \mu, \left[\frac{\delta F_1}{\delta \mu}, \frac{\delta F_2}{\delta \mu} \right] \right\rangle \\ &= \langle \mathbf{J}(z), [\xi, \eta] \rangle = J^{[\xi, \eta]}(z) \\ &= \{J^\xi, J^\eta\}, \end{aligned}$$

where the last equality follows by infinitesimal equivariance.

But, for $z \in P$ and $v_z \in T_z P$, we have

$$\begin{aligned} \mathbf{d}(F_1 \circ \mathbf{J})(z)(v_z) &= \mathbf{d}F_1(\mu)(T_z \mathbf{J}(v_z)) = \left\langle T_z \mathbf{J}(v_z), \frac{\delta F_1}{\delta \mu} \right\rangle \\ &= \langle T_z \mathbf{J}(v_z), \xi \rangle = \mathbf{d}J^\xi(z)(v_z). \end{aligned}$$

Thus $\mathbf{d}(F_1 \circ \mathbf{J})(z) = \mathbf{d}J^\xi(z)$. So, as the Poisson bracket on P depends only on the point values of the first derivatives, we have

$$\{F_1 \circ \mathbf{J}, F_2 \circ \mathbf{J}\}(z) = \{J^\xi, J^\eta\}(z)$$

which proves the theorem. $\blacksquare$

Remark 5.3 The same result holds if G acts on the right, provided that we consider on $\mathfrak{g}^*$ the minus Lie-Poisson structure.

Theorem 5.4 (Noether's Theorem) *Let P be a Poisson manifold, G a Lie group acting canonically on P admitting a momentum map $\mathbf{J}$ and $H : P \rightarrow \mathbb{R}$ a G -invariant function. Then $\mathbf{J}$ is a constant of motion for H . That is, if ϕ_t is the flow of X_H then $\mathbf{J} \circ \phi_t = \mathbf{J}$.*

Proof If H is G -invariant then $\xi_P[H] = 0$ which implies

$$0 = \xi_P[H] = X_{J^\xi}[H] = \{H, J^\xi\} = -X_H[J^\xi].$$

So J^ξ is constant on the flow of X_H for every $\xi \in \mathfrak{g}$. $\blacksquare$

5.3 Examples of Momentum Maps

1. The Hamiltonian

The flow ϕ_t of a complete vector field on a manifold P defines an $\mathbb{R}$ -action on P given by $\phi(t, z) := \phi_t(z)$.

Consider the $\mathbb{R}$ -action on a Poisson manifold P given by the flow of a complete Hamiltonian vector field X_H . Since the flow of X_H is canonical, this action preserves the Poisson bracket. Let us show that $H : P \rightarrow \mathbb{R}$ is an equivariant momentum map for this action. Indeed, if $s \in \mathbb{R}$, its infinitesimal generator is

$$s_P(z) = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \phi_{\epsilon s}(z) = sX_H(z) = X_{sH}(z)$$

which shows that $J^s = sH$. Identifying $\mathbb{R}$ with $\mathbb{R}^*$ using the product of elements in $\mathbb{R}$, we get hence $\mathbf{J} = H$. Invariance of H is equivalent to the conservation of energy.

2. Linear momentum

Let $N \in \mathbb{N}$ and consider the N -particle system, with configuration space $Q = \mathbb{R}^{3N}$. Let $\mathbb{R}^3$ act on Q by translations, i.e., $\Phi : \mathbb{R}^3 \times Q \rightarrow Q$ is given by

$$(x, (\mathbf{q}_1, \dots, \mathbf{q}_N)) \mapsto (\mathbf{q}_1 + x, \dots, \mathbf{q}_N + x).$$

The infinitesimal generator corresponding to $\xi \in \mathbb{R}^3$ is:

$$\begin{aligned} \xi_Q(\mathbf{q}_1, \dots, \mathbf{q}_N) &= \left. \frac{d}{dt} \right|_{t=0} \Phi(t\xi, (\mathbf{q}_1, \dots, \mathbf{q}_N)) \\ &= \left. \frac{d}{dt} \right|_{t=0} (\mathbf{q}_1 + t\xi, \dots, \mathbf{q}_N + t\xi) \\ &= (\xi, \dots, \xi) \in T_{(\mathbf{q}_1, \dots, \mathbf{q}_N)} \mathbb{R}^{3N}. \end{aligned}$$

Thus, by (5.11), the lifted $\mathbb{R}^{3N}$ -action to $T^*\mathbb{R}^{3N}$ admits an invariant momentum map given by

$$\begin{aligned} \langle \mathbf{J}(\mathbf{q}_1, \dots, \mathbf{q}_N, \mathbf{p}^1, \dots, \mathbf{p}^N), \xi \rangle &= \langle (\mathbf{p}^1, \dots, \mathbf{p}^N), \xi_Q(\mathbf{q}_1, \dots, \mathbf{q}_N) \rangle \\ &= \mathbf{p}^1 \cdot \xi + \dots + \mathbf{p}^N \cdot \xi = (\mathbf{p}^1 + \dots + \mathbf{p}^N) \cdot \xi, \end{aligned}$$

that is, $\mathbf{J}(\mathbf{q}_1, \dots, \mathbf{q}_N, \mathbf{p}^1, \dots, \mathbf{p}^N) = \mathbf{p}^1 + \dots + \mathbf{p}^N$, which is the classical linear momentum.

3. Angular momentum

Let $Q = \mathbb{R}^3$ and $\Phi : SO(3) \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be the standard action $\Phi(A, \mathbf{q}) := A\mathbf{q}$. Using the isomorphism of the Lie algebras $(\mathbb{R}^3, \times)$ and $(\mathfrak{so}(3), [\cdot, \cdot])$ given by (1.13) and the expression (5.9) of the infinitesimal generator, the equivariant momentum map (5.11) for the lifted action of $SO(3)$ to T^*Q is given by

$$\langle \mathbf{J}(\mathbf{q}, \mathbf{p}), \xi \rangle = \langle (\mathbf{q}, \mathbf{p}), \xi_Q(\mathbf{q}) \rangle = \mathbf{p} \cdot (\xi \times \mathbf{q}) = (\mathbf{q} \times \mathbf{p}) \cdot \xi,$$

where $\xi \in \mathbb{R}^3$. Thus $\mathbf{J}(\mathbf{q}, \mathbf{p}) = \mathbf{q} \times \mathbf{p}$ which is the classical angular momentum.

4. Momentum map for matrix groups

Denote by $GL(n, \mathbb{R})$ the group of linear isomorphisms of $\mathbb{R}^n$ to $\mathbb{R}^n$, that is the general linear group. Let G be a Lie subgroup of $GL(n, \mathbb{R})$, with Lie algebra $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{R})$. Consider the action of G on $Q := \mathbb{R}^n$ to be by matrix multiplication on the left, that is, $\Phi : (A, \mathbf{q}) \in G \times \mathbb{R}^n \mapsto A\mathbf{q} \in \mathbb{R}^n$. For $\xi \in \mathfrak{g}$ the corresponding infinitesimal generator is given by:

$$\xi_Q(\mathbf{q}) = \left. \frac{d}{dt} \right|_{t=0} \exp(t\xi)\mathbf{q} = \xi\mathbf{q}.$$

Identify $\mathfrak{gl}(n, \mathbb{R})^*$ with $\mathfrak{gl}(n, \mathbb{R})$ via the positive definite inner product on $\mathfrak{gl}(n, \mathbb{R})$

given by

$$\langle a, b \rangle = \text{trace}(a^T b), \quad (5.14)$$

where a^T is the transpose of a . By (5.11), the momentum map for the cotangent lifted action is given by:

$$\begin{aligned} \langle \mathbf{J}(\mathbf{q}, \mathbf{p}), \xi \rangle &= \mathbf{p} \cdot \xi_Q(\mathbf{q}) = \mathbf{p} \cdot \xi \mathbf{q} = \text{trace}(\mathbf{p}^T \xi \mathbf{q}) \\ &= \text{trace}(\mathbf{q} \mathbf{p}^T \xi) = \langle \mathbf{p} \mathbf{q}^T, \xi \rangle. \end{aligned}$$

for any $\xi \in \mathfrak{g}$. Now write $\mathfrak{gl}(n, \mathbb{R}) = \mathfrak{g} \oplus \mathfrak{g}^\perp$, where the perpendicular is taken relative to the inner product (5.14) and let $\Pi_{\mathfrak{g}} : \mathfrak{gl}(n, \mathbb{R}) \rightarrow \mathfrak{g}$ be the corresponding orthogonal projection. Then $\mathfrak{g}$ is identified with $\mathfrak{g}^*$ and we get, $\mathbf{J}(\mathbf{q}, \mathbf{p}) = \Pi_{\mathfrak{g}}(\mathbf{p} \mathbf{q}^T)$.

5. Canonical momentum map on $\mathfrak{g}^*$

The Lie group G acts on the dual $\mathfrak{g}^*$ of its Lie algebra $\mathfrak{g}$ by the coadjoint action. Since $\text{Ad}_g : \mathfrak{g} \rightarrow \mathfrak{g}$ is a Lie algebra isomorphism for every $g \in G$, Proposition 4.10 insures that the coadjoint action is canonical relative to the Lie-Poisson bracket on $\mathfrak{g}^*$. The infinitesimal generator corresponding to $\xi \in \mathfrak{g}$ for the coadjoint action is, for $\mu \in \mathfrak{g}^*$ and $\eta \in \mathfrak{g}$, given by:

$$\begin{aligned} \langle \xi_{\mathfrak{g}^*}(\mu), \eta \rangle &= \left. \frac{d}{dt} \right|_{t=0} \langle \text{Ad}_{\exp(-t\xi)}^* \mu, \eta \rangle = \left. \frac{d}{dt} \right|_{t=0} \langle \mu, \text{Ad}_{\exp(-t\xi)} \eta \rangle \\ &= \langle \mu, [-\xi, \eta] \rangle = \langle \mu, -\text{ad}_{\xi} \eta \rangle = \langle -\text{ad}_{\xi}^* \mu, \eta \rangle, \end{aligned} \quad (5.15)$$

so $\xi_{\mathfrak{g}^*} = -\text{ad}_{\xi}^*$ for every $\xi \in \mathfrak{g}$.

By Proposition 4.9, the Hamiltonian vector field for $H \in \mathcal{F}(\mathfrak{g}^*)$ has the expression

$$X_H(\mu) = \mp \text{ad}_{\delta H / \delta \mu}^* \mu.$$

Therefore, the momentum map for the coadjoint action, if it exists, must satisfy

$$\mp \text{ad}_{\delta J^{\xi} / \delta \mu}^* \mu = -\text{ad}_{\xi}^* \mu \quad \text{for all } \xi \in \mathfrak{g}, \quad \text{and } \mu \in \mathfrak{g}^*,$$

which shows that the momentum map for the coadjoint action exists and is given by $\langle \mathbf{J}(\mu), \xi \rangle = \pm \langle \mu, \xi \rangle$. Therefore $\mathbf{J} = \pm \text{id}_{\mathfrak{g}^*}$.

6. Momentum map for products

Let P_1 and P_2 be Poisson manifolds and $P_1 \times P_2$ be their product endowed with the product Poisson structure, that is, if $F, H : P_1 \times P_2 \rightarrow \mathbb{R}$, then

$$\{F, H\}_{P_1 \times P_2}(z_1, z_2) = \{F_{z_2}, H_{z_2}\}_{P_1}(z_1) + \{F_{z_1}, H_{z_1}\}_{P_2}(z_2),$$

where $F_{z_1} := F(z_1, \cdot) : P_2 \rightarrow \mathbb{R}$ and similarly for $F_{z_2} := F(\cdot, z_2) : P_1 \rightarrow \mathbb{R}$.

Let $\Phi : G \times P_1 \rightarrow P_1$ and $\Psi : G \times P_2 \rightarrow P_2$ be canonical G -actions

admitting (equivariant) momentum maps $\mathbf{J}_1 : P_1 \rightarrow \mathfrak{g}_1^*$ and $\mathbf{J}_2 : P_2 \rightarrow \mathfrak{g}_2^*$ respectively. Then the product action $\Pi : G \times P_1 \times P_2 \rightarrow P_1 \times P_2$ given by $\Pi(g, (z_1, z_2)) := (\Phi(g, z_1), \Psi(g, z_2))$ admits an (equivariant) momentum map $\mathbf{J} : P_1 \times P_2 \rightarrow \mathfrak{g}^*$ given by $\mathbf{J}(z_1, z_2) = \mathbf{J}_1(z_1) + \mathbf{J}_2(z_2)$.

To prove this statement, we begin by showing that the action Π is canonical. Indeed, for every $g \in G$ we get

$$\begin{aligned} \{F, H\}_{P_1 \times P_2}(\Pi_g(z_1, z_2)) &= \{F, H\}_{P_1 \times P_2}(\Phi(g, z_1), \Psi(g, z_2)) \\ &= \{F_{g \cdot z_2}, H_{g \cdot z_2}\}_{P_1}(\Phi(g, z_1)) + \{F_{g \cdot z_1}, H_{g \cdot z_1}\}_{P_2}(\Psi(g, z_2)) \\ &= \{F_{g \cdot z_2} \circ \Phi_g, H_{g \cdot z_2} \circ \Phi_g\}_{P_1}(z_1) + \{F_{g \cdot z_1} \circ \Psi_g, H_{g \cdot z_1} \circ \Psi_g\}_{P_2}(z_2) \\ &= \{(F \circ \Pi_g)_{z_2}, (H \circ \Pi_g)_{z_2}\}_{P_1}(z_1) + \{(F \circ \Pi_g)_{z_1}, (H \circ \Pi_g)_{z_1}\}_{P_2}(z_2) \\ &= \{F \circ \Pi_g, H \circ \Pi_g\}_{P_1 \times P_2}(z_1, z_2). \end{aligned}$$

For $\xi \in \mathfrak{g}$, the infinitesimal generator of Π corresponding to ξ is given by

$$\begin{aligned} \xi_{P_1 \times P_2}(z_1, z_2) &= \left. \frac{d}{dt} \right|_{t=0} \Pi(\exp(t\xi), (z_1, z_2)) \\ &= \left. \frac{d}{dt} \right|_{t=0} (\Phi(\exp(t\xi), z_1), \Psi(\exp(t\xi), z_2)) \\ &= (\xi_{P_1}(z_1), \xi_{P_2}(z_2)) = (X_{J_1^\xi}(z_1), X_{J_2^\xi}(z_2)) \\ &= X_{(J_1^\xi, J_2^\xi)}(z_1, z_2), \end{aligned}$$

where the Hamiltonian vector field in the last line is on $P_1 \times P_2$ for the function $(J_1^\xi, J_2^\xi)(z_1, z_2) := (J_1^\xi(z_1), J_2^\xi(z_2)) = \langle \mathbf{J}_1(z_1) + \mathbf{J}_2(z_2), \xi \rangle$. This shows that a momentum map for the product action is given indeed by $\mathbf{J}(z_1, z_2) = \mathbf{J}_1(z_1) + \mathbf{J}_2(z_2)$, as stated.

If $\mathbf{J}_1$ and $\mathbf{J}_2$ are equivariant so is $\mathbf{J}$, as an easy computation shows.

7. Momentum maps for the cotangent lift of the left and right translations of G to T^*G

Let G be a Lie group and denote by $L_g(h) := gh$ and $R_g(h) := hg$ the left and right translations of G on itself. Denote by $\mathbf{J}_L$ and $\mathbf{J}_R$ the corresponding equivariant momentum maps of the lifts of these actions to T^*G . To compute these momentum maps we use (5.11) to get for any $\alpha_g \in T_g^*G$ and $\xi \in \mathfrak{g}$

$$\begin{aligned} \langle \mathbf{J}_L(\alpha_g), \xi \rangle &= \left\langle \alpha_g, \left. \frac{d}{dt} \right|_{t=0} L_{\exp(t\xi)} g \right\rangle = \langle \alpha_g, T_e R_g \xi \rangle = \langle T_e^* R_g \alpha_g, \xi \rangle \\ \langle \mathbf{J}_R(\alpha_g), \xi \rangle &= \left\langle \alpha_g, \left. \frac{d}{dt} \right|_{t=0} R_{\exp(t\xi)} g \right\rangle = \langle \alpha_g, T_e L_g \cdot \xi \rangle = \langle T_e^* L_g \alpha_g, \xi \rangle, \end{aligned}$$

which shows that

$$\mathbf{J}_L(\alpha_g) = T_e^* R_g \alpha_g \quad \text{and} \quad \mathbf{J}_R(\alpha_g) = T_e^* L_g \alpha_g. \quad (5.16)$$

8. Momentum map in Maxwell's equations

Let $\mathcal{A}$ be the space of vector potentials $\mathbf{A}$ on $\mathbb{R}^3$, that is, smooth functions $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Let $P := T^*\mathcal{A}$, whose elements are denoted $(\mathbf{A}, -\mathbf{E})$ with $\mathbf{A}$ and $\mathbf{E}$ vector fields on $\mathbb{R}^3$. Let $G = \mathcal{F}(\mathbb{R}^3)$ act on $\mathcal{A}$ by

$$\phi \cdot \mathbf{A} = \mathbf{A} + \text{grad } \phi.$$

The Lie algebra $\mathfrak{g}$ of G coincides with $\mathcal{F}(\mathbb{R}^3)$ and we formally think of $\mathfrak{g}^*$ as $\mathfrak{g}$ via the weakly nondegenerate L^2 pairing. Thus given $\xi \in \mathfrak{g}$, the corresponding infinitesimal generator is:

$$\xi_{\mathcal{A}}(\mathbf{A}) = \text{grad } \xi.$$

Assuming that all computations below are justified by imposing the relevant decay conditions at infinity, the momentum map (5.11) becomes in this case

$$\langle \mathbf{J}(\mathbf{A}, -\mathbf{E}), \xi \rangle = \int -\mathbf{E} \cdot \text{grad } \xi \, d^3x = \int (\text{div } \mathbf{E}) \xi \, d^3x.$$

Thus the invariant momentum map $\mathbf{J} : T^*\mathcal{A} \rightarrow \mathcal{F}(\mathbb{R}^3)$ is $\mathbf{J}(\mathbf{A}, -\mathbf{E}) = \text{div } \mathbf{E}$.

9. Clairaut's Theorem

Let Q be a surface of revolution obtained by rotating the graph of the smooth function $r = f(z)$ about the z -axis. Pull back the usual Riemannian metric given by the Euclidean inner product on $\mathbb{R}^3$ to Q and identify T^*Q with TQ using this induced metric. The circle S^1 acts on Q and the Riemannian metric on Q is obviously invariant under this action. Consider the geodesic flow on Q , so the Hamiltonian of this vector field on TQ is given by the kinetic energy of the metric; thus it is also S^1 invariant. The infinitesimal generator of $\xi \in \mathbb{R}$, the Lie algebra of S^1 , is given by

$$\xi_Q(\mathbf{q}) = \left. \frac{d}{dt} \right|_{t=0} \begin{bmatrix} \cos t\xi & -\sin t\xi & 0 \\ \sin t\xi & \cos t\xi & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{q} = \xi \hat{\mathbf{k}} \mathbf{q} = \xi \mathbf{k} \times \mathbf{q}.$$

Therefore, the momentum map $\mathbf{J} : TQ \rightarrow \mathbb{R}$, given by (5.11), has the expression

$$\mathbf{J}(\mathbf{q}, \mathbf{v}) = \mathbf{v} \cdot \xi \mathbf{k} \times \mathbf{q} = \xi r \|\mathbf{v}\| \cos \theta$$

since r is the distance of $\mathbf{q}$ to the z -axis and where θ is the angle between $\mathbf{v}$

and the horizontal plane. Recall that $\|\mathbf{v}\|$ is conserved since the kinetic energy is constant on the geodesic flow. By Noether's theorem it follows that $\mathbf{J}$ is conserved which then implies that $r \cos \theta$ is conserved along any geodesic on Q . This is the statement of the classical Clairaut's theorem.

10. Momentum map for symplectic representations

Let (V, Ω) be a symplectic vector space and let G be a Lie group acting linearly and symplectically on V . This action admits an equivariant momentum map $\mathbf{J} : V \rightarrow \mathfrak{g}$ given by

$$J^\xi(v) = \langle \mathbf{J}(v), \xi \rangle = \frac{1}{2} \Omega(\xi \cdot v, v),$$

where $\xi \cdot v$ denotes the Lie algebra representation of the element $\xi \in \mathfrak{g}$ on the vector $v \in V$. To verify this, note that the infinitesimal generator $\xi_V(v) = \xi \cdot v$, by the definition of the Lie algebra representation induced by the given Lie group representation, and that $\Omega(\xi \cdot u, v) = -\Omega(u, \xi \cdot v)$ for all $u, v \in V$. Therefore

$$dJ^\xi(u)(v) = \frac{1}{2} \Omega(\xi \cdot u, v) + \frac{1}{2} \Omega(\xi \cdot v, u) = \Omega(\xi \cdot u, v).$$

Equivariance of $\mathbf{J}$ follows from the obvious relation $g^{-1} \cdot \xi \cdot g \cdot v = (\text{Ad}_{g^{-1}} \xi) \cdot v$ for any $g \in G$, $\xi \in \mathfrak{g}$, and $v \in V$.

11. Cayley-Klein parameters and the Hopf fibration

Consider the natural action of $SU(2)$ on $\mathbb{C}^2$. Since this action is by isometries of the Hermitian metric, it is automatically symplectic and therefore has a momentum map $\mathbf{J} : \mathbb{C}^2 \rightarrow \mathfrak{su}(2)^*$ given in example 10, that is,

$$\langle \mathbf{J}(z, w), \xi \rangle = \frac{1}{2} \Omega(\xi \cdot (z, w), (z, w)),$$

where $z, w \in \mathbb{C}$ and $\xi \in \mathfrak{su}(2)$. Now recall from §2.1 that the symplectic form on $\mathbb{C}^2$ is given by minus the imaginary part of the Hermitian inner product. The Lie algebra $\mathfrak{su}(2)$ of $SU(2)$ consists of 2×2 skew Hermitian matrices of trace zero. This Lie algebra is isomorphic to $\mathfrak{so}(3)$ and therefore to $(\mathbb{R}^3, \times)$ by the isomorphism given by

$$\mathbf{x} = (x^1, x^2, x^3) \in \mathbb{R}^3 \mapsto \tilde{\mathbf{x}} := \frac{1}{2} \begin{bmatrix} -ix^3 & -ix^1 - x^2 \\ -ix^1 + x^2 & ix^3 \end{bmatrix} \in \mathfrak{su}(2).$$

Thus we have $[\tilde{\mathbf{x}}, \tilde{\mathbf{y}}] = (\mathbf{x} \times \mathbf{y})^\sim$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$. Other useful relations are $\det(2\tilde{\mathbf{x}}) = \|\mathbf{x}\|^2$ and $\text{trace}(\tilde{\mathbf{x}}\tilde{\mathbf{y}}) = -\frac{1}{2}\mathbf{x} \cdot \mathbf{y}$. Identify $\mathfrak{su}(2)^*$ with $\mathbb{R}^3$ by the map $\mu \in \mathfrak{su}(2)^* \mapsto \tilde{\mu} \in \mathbb{R}^3$ defined by

$$\tilde{\mu} \cdot \mathbf{x} := -2\langle \mu, \tilde{\mathbf{x}} \rangle$$

for any $\mathbf{x} \in \mathbb{R}^3$. With these notations, the momentum map $\check{\mathbf{J}} : \mathbb{C}^2 \rightarrow \mathbb{R}^3$ can be explicitly computed in coordinates: for any $\mathbf{x} \in \mathbb{R}^3$ we have

$$\begin{aligned} \check{\mathbf{J}}(z, w) \cdot \mathbf{x} &= -2\langle \mathbf{J}(z, w), \tilde{\mathbf{x}} \rangle \\ &= \frac{1}{2} \operatorname{Im} \left(\begin{bmatrix} -ix^3 & -ix^1 - x^2 \\ -ix^1 + x^2 & ix^3 \end{bmatrix} \begin{bmatrix} z \\ w \end{bmatrix} \cdot \begin{bmatrix} z \\ w \end{bmatrix} \right) \\ &= -\frac{1}{2} (2 \operatorname{Re}(w\bar{z}), 2 \operatorname{Im}(w\bar{z}), |z|^2 - |w|^2) \cdot \mathbf{x}. \end{aligned}$$

Therefore

$$\check{\mathbf{J}}(z, w) = -\frac{1}{2} (2w\bar{z}, |z|^2 - |w|^2) \in \mathbb{R}^3.$$

By Theorem 5.2, $\check{\mathbf{J}}$ is a Poisson map from $\mathbb{C}^2$, endowed with the canonical symplectic structure, to $\mathbb{R}^3$, endowed with the + Lie Poisson structure. Therefore, $-\check{\mathbf{J}} : \mathbb{C}^2 \rightarrow \mathbb{R}^3$ is a canonical map, if $\mathbb{R}^3$ has the - Lie-Poisson bracket relative to which the free rigid body equations are Hamiltonian. Pulling back the Hamiltonian $H(\mathbf{\Pi}) = \mathbf{\Pi} \cdot \mathbb{I}^{-1} \mathbf{\Pi} / 2$ to $\mathbb{C}^2$ gives a Hamiltonian function (called collective) on $\mathbb{C}^2$. The classical Hamilton equations for this function are therefore projected by $-\check{\mathbf{J}}$ to the rigid body equations $\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbb{I}^{-1} \mathbf{\Pi}$. In this context, the variables (z, w) are called the **Cayley-Klein parameters**. They represent a first attempt to understand the rigid body equations as a Hamiltonian system, before the introduction of Poisson manifolds. In quantum mechanics, the same variables are called the **Kustaanheimo-Stiefel coordinates**. A similar construction was carried out in fluid dynamics making the Euler equations a Hamiltonian system relative to the so-called **Clebsch variables**.

Now notice that if $(z, w) \in S^3 := \{(z, w) \in \mathbb{C}^2 \mid |z|^2 + |w|^2 = 1\}$, then $\|-\check{\mathbf{J}}(z, w)\| = 1/2$, so that $-\check{\mathbf{J}}|_{S^3} : S^3 \rightarrow S_{1/2}^2$, where $S_{1/2}^2$ is the sphere in $\mathbb{R}^3$ of radius $1/2$. It is also easy to see that $-\check{\mathbf{J}}|_{S^3}$ is surjective and that its fibers are circles. Indeed, given $(x^1, x^2, x^3) = (x^1 + ix^2, x^3) = (re^{i\psi}, x^3) \in S_{1/2}^2$, the inverse image of this point is

$$\begin{aligned} -\check{\mathbf{J}}^{-1}(re^{i\psi}, x^3) &= \\ &= \left\{ \left(e^{i\theta} \sqrt{\frac{1}{2} + x^3}, e^{i\varphi} \sqrt{\frac{1}{2} - x^3} \right) \in S^3 \mid e^{i(\theta - \varphi + \psi)} = 1 \right\}. \end{aligned}$$

One recognizes now that $-\check{\mathbf{J}}|_{S^3} : S^3 \rightarrow S^2$ is the **Hopf fibration**. In other words, *the momentum map of the $SU(2)$ -action on $\mathbb{C}^2$, the Cayley-Klein parameters, the Kustaanheimo-Stiefel coordinates, and the family of Hopf fibrations on concentric three-spheres in $\mathbb{C}^2$ are the same map.*

6 Lie-Poisson and Euler-Poincaré Reduction

In this lecture we shall present the simplest case of reduction, namely the Lie-Poisson reduction theorem. It states that the quotient of a cotangent bundle of a Lie group by the lift of the left or right translation is Poisson isomorphic to the dual of the Lie algebra endowed with the $\pm$ Lie-Poisson bracket. The symplectic leaves of this Poisson structure are the connected components of the coadjoint orbits. The Lagrangian version of this result is a reduced constrained variational principle that is equivalent to first order equations on the dual of the Lie algebra, called Euler-Poincaré equations. We shall carry out in detail several examples both in finite and in infinite dimensions.

6.1 Lie-Poisson Reduction

One way to construct new Poisson manifolds out of known ones is by symmetry reduction.

Let G be a Lie group acting canonically on a Poisson manifold P . Assume that the orbit space P/G is a smooth manifold and the quotient projection $\pi : P \rightarrow P/G$ a surjective submersion. This is the case, for example, if the G -action is proper and free, or proper with all isotropy groups conjugate. Then there exists a unique Poisson bracket $\{\cdot, \cdot\}_{P/G}$ on P/G relative to which π is a Poisson map. The Poisson bracket on P/G is given in the following way. If $\hat{F}, \hat{H} \in \mathcal{F}(P/G)$, then $\hat{F} \circ \pi, \hat{H} \circ \pi \in \mathcal{F}(P)$ are G -invariant functions and, due to the fact that the action is canonical, their Poisson bracket $\{\hat{F} \circ \pi, \hat{H} \circ \pi\}$ is also G -invariant. Therefore, this function descends to a smooth function on the quotient P/G ; this is, by definition, $\{\hat{F}, \hat{H}\}_{P/G}$ and we have, by construction, $\{\hat{F} \circ \pi, \hat{H} \circ \pi\} = \{\hat{F}, \hat{H}\}_{P/G} \circ \pi$. It is easy to see that $\{\cdot, \cdot\}_{P/G}$ so defined satisfies all the axioms of a Poisson bracket. This proves in a constructive way the existence of the Poisson bracket on the quotient. In addition, because $\pi : P \rightarrow P/G$ is a surjective Poisson submersion, the bracket on the quotient is necessarily unique with the requirement that π is a Poisson map.

When the manifold P is the cotangent bundle T^*G of a Lie group G and the action of G on T^*G is by cotangent lift of the left (or right) translation of G on itself, the reduced space $(T^*G)/G$ is naturally diffeomorphic to the dual $\mathfrak{g}^*$ of the Lie algebra $\mathfrak{g}$ of G . The goal of this section is to show that *the quotient Poisson bracket is the minus (or plus) Lie-Poisson bracket*. To do this, we follow the presentation in [MaRa94], §13, and we will give two proofs.

First Proof. The left and right translations by $g \in G$ are denoted by $L_g(h) := gh$ and $R_g(h) = hg$. Let $\mathcal{F}_L(T^*G)$ be the space of smooth left-invariant functions on T^*G , that is, $F_L \in \mathcal{F}_L(T^*G)$ if and only if $F_L \circ T^*L_g = F_L$

for all $g \in G$, where T^*L_g is the cotangent lift of L_g . Similarly, a right-invariant function F_R verifies $F_R \circ T^*R_g = F_R$ and the space of all smooth right invariant functions on T^*G is denoted by $\mathcal{F}_R(T^*G)$. Note that $\mathcal{F}_L(T^*G)$ and $\mathcal{F}_R(T^*G)$ are closed under Poisson bracket.

Any $F \in \mathcal{F}(\mathfrak{g}^*)$ can be uniquely extended to a left (respectively right) invariant function F_L (respectively F_R) on T^*G by setting

$$F_L(\alpha_g) := F(T_e^*L_g\alpha_g) = (F \circ \mathbf{J}_L)(\alpha_g)$$

(respectively $F_R(\alpha_g) := F(T_e^*R_g\alpha_g) = (F \circ \mathbf{J}_R)(\alpha_g)$). Here $\mathbf{J}_L$ and $\mathbf{J}_R$ are the momentum maps for the left and right translations given by (5.16).

So, composition with $\mathbf{J}_R$ (respectively with $\mathbf{J}_L$) defines, by Theorem 5.2, an isomorphism of Poisson algebras $\mathcal{F}(\mathfrak{g}_-^*) \rightarrow \mathcal{F}_L(T^*G)$ (respectively, $\mathcal{F}(\mathfrak{g}_+^*) \rightarrow \mathcal{F}_R(T^*G)$) whose inverse is the restriction to the fiber $T_e^*G = \mathfrak{g}^*$:

$$\begin{aligned} \{F, H\}_- \circ \mathbf{J}_R &= \{F \circ \mathbf{J}_R, H \circ \mathbf{J}_R\} = \{F_L, H_L\}, \\ \{F, H\}_+ \circ \mathbf{J}_L &= \{F \circ \mathbf{J}_L, H \circ \mathbf{J}_L\} = \{F_R, H_R\}, \end{aligned}$$

$$\{F, H\}_- = \{F_L, H_L\}|_{\mathfrak{g}^*}, \quad \text{and} \quad \{F, H\}_+ = \{F_R, H_R\}|_{\mathfrak{g}^*},$$

where $\{\cdot, \cdot\}_\pm$ are the Lie-Poisson brackets on $\mathfrak{g}^*$ and $\{\cdot, \cdot\}$ is the Poisson bracket on T^*G . ■

While mathematically correct, this proof is unsatisfactory for it requires to know a priori that $\mathfrak{g}^*$ is a Poisson manifold. This is why we shall give below a second proof in which the Lie-Poisson bracket is *discovered* by carrying out the identification of $(T^*G)/G$ with $\mathfrak{g}^*$ explicitly.

Second Proof. This is done in several steps. We begin by noting that the map $\mathcal{P} : X \in \mathfrak{X}(Q) \mapsto \langle \cdot, X \rangle \in \mathcal{L}(T^*Q)$, where

$$\mathcal{L}(T^*Q) := \{f \in \mathcal{F}(T^*Q) \mid f \text{ linear on the fibers}\}$$

is a Poisson subalgebra of $\mathcal{F}(T^*Q)$, is a Lie algebra anti-isomorphism. To see this, work in coordinates and note that $F, H \in \mathcal{L}(T^*Q)$ if and only if $F(q, p) = X^i(q)p_i$, $H(q, p) = Y^i(q)p_i$ and hence

$$\{F, H\}(q, p) = \frac{\partial F}{\partial q^j} \frac{\partial H}{\partial p_j} - \frac{\partial H}{\partial q^j} \frac{\partial F}{\partial p_j} = \left(\frac{\partial X^i}{\partial q^j} Y^i - \frac{\partial Y^i}{\partial q^j} X^i \right) p_i.$$

Thus, $\{\mathcal{P}(X), \mathcal{P}(Y)\} = -\mathcal{P}([X, Y])$. This immediately implies that the linear isomorphism $Y \in \mathfrak{X}(Q) \mapsto X_{\mathcal{P}(Y)} \in \{X_F \mid F \in \mathcal{L}(T^*Q)\}$ preserves the Lie brackets. Indeed,

$$[X, Y] \mapsto X_{\mathcal{P}([X, Y])} = -X_{\{\mathcal{P}(X), \mathcal{P}(Y)\}} = [X_{\mathcal{P}(X)}, X_{\mathcal{P}(Y)}]$$

for any $X, Y \in \mathfrak{X}(Q)$. Thus $Y \in \mathfrak{X}(Q) \mapsto X_{\mathcal{P}(Y)} \in \{X_F \mid F \in \mathcal{L}(T^*Q)\}$ is a Lie algebra isomorphism.

Next we prove that if the flow of $X \in \mathfrak{X}(Q)$ is ϕ_t then its cotangent lift $T^*\phi_{-t}$ is the flow of $X_{\mathcal{P}(X)}$.

To see this, let $\pi : T^*Q \rightarrow Q$ be the canonical projection. Differentiating at $t = 0$ the equation $\pi \circ T^*\phi_{-t} = \phi_t \circ \tau_Q$, we get

$$T\pi \circ Y = X \circ \pi \quad \text{where} \quad Y(\alpha_q) = \left. \frac{d}{dt} \right|_{t=0} T^*\phi_{-t}(\alpha_q).$$

So $T^*\phi_{-t}$ is the flow of Y . As $T^*\phi_{-t}$ preserves the canonical one-form, it follows that $\mathcal{L}_Y\Theta = 0$ and hence $\mathbf{i}_Y\Omega = \mathbf{d}(\mathbf{i}_Y\Theta)$. This shows that Y is Hamiltonian with energy $\mathbf{i}_Y\Theta(\alpha_q) = \langle \alpha_q, (T\pi \circ Y)(\alpha_q) \rangle = \langle \alpha_q, X(q) \rangle = \mathcal{P}(X)(\alpha_q)$, that is, $Y = X_{\mathcal{P}(X)}$, which proves the statement.

Finally we shall implement the diffeomorphism between $(T^*G)/G$ and $\mathfrak{g}^*$ given by dropping to the quotient the left invariant map $\mathbf{J}_R : T^*G \rightarrow \mathfrak{g}^*$. Concretely, we shall prove that the push-forward by this diffeomorphism of the quotient Poisson bracket on $(T^*G)/G$ gives the known formula

$$\{F, H\}_-(\mu) = - \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right] \right\rangle.$$

To achieve this, we shall show that if $F, H \in \mathcal{F}(\mathfrak{g}^*)$, we get the identity $\{F_L, H_L\}|_{\mathfrak{g}^*} = \{F, H\}_-$.

This is done in the following way. Since the Poisson bracket of any $F, H \in \mathcal{F}(\mathfrak{g}^*)$ depends only on the differentials of F and H , it is enough to prove the statement for linear functions on $\mathfrak{g}^*$. So we can replace the general smooth function $F : \mathfrak{g}^* \rightarrow \mathbb{R}$ with its linear part

$$F^\ell(\mu) := \left\langle \mu, \frac{\delta F}{\delta \mu} \right\rangle.$$

Then, denoting by ξ_L the left invariant vector field on G whose value at the identity is $\xi \in \mathfrak{g}$, that is, $\xi_L(g) := T_e L_g \xi$, we get

$$\begin{aligned} F_L^\ell(\alpha_g) &= F^\ell(T_e L_g^*(\alpha_g)) = \left\langle T_e L_g^*(\alpha_g), \frac{\delta F}{\delta \mu} \right\rangle \\ &= \left\langle \alpha_g, T_e L_g \left(\frac{\delta F}{\delta \mu} \right) \right\rangle = \left\langle \alpha_g, \left(\frac{\delta F}{\delta \mu} \right)_L(g) \right\rangle \\ &= \mathcal{P} \left(\left(\frac{\delta F}{\delta \mu} \right)_L \right)(\alpha_g). \end{aligned}$$

Therefore if $\mu \in \mathfrak{g}^*$ we have

$$\begin{aligned} \{F_L^\ell, H_L^\ell\}|_{\mathfrak{g}^*}(\mu) &= \{F_L^\ell, H_L^\ell\}(\mu) = \left\{ \mathcal{P} \left(\left(\frac{\delta F}{\delta \mu} \right)_L \right), \mathcal{P} \left(\left(\frac{\delta H}{\delta \mu} \right)_L \right) \right\}(\mu) \\ &= -\mathcal{P} \left(\left[\left(\frac{\delta F}{\delta \mu} \right)_L, \left(\frac{\delta H}{\delta \mu} \right)_L \right] \right)(\mu) \\ &= -\mathcal{P} \left(\left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right]_L \right)(\mu) = -\left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right]_L(e) \right\rangle \\ &= -\left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right] \right\rangle = \{F, H\}_-(\mu) = \{F^\ell, H^\ell\}(\mu), \end{aligned}$$

which ends the proof. $\blacksquare$

Thus, the identification of the set of real-valued functions on $\mathfrak{g}^*$ with the left (respectively, right) invariant functions on T^*G endows $\mathfrak{g}^*$ with the minus (respectively the plus) Lie-Poisson bracket.

6.2 Lie-Poisson Reduction of Dynamics

In this section we shall discuss the Lie-Poisson reduction of dynamics. Since the momentum maps $\mathbf{J}_R : T^*G \rightarrow \mathfrak{g}_-^*$ and $\mathbf{J}_L : T^*G \rightarrow \mathfrak{g}_+^*$ are Poisson maps, they will map integral curves of left and right invariant Hamiltonian vector fields to integral curves of Lie-Poisson Hamiltonian systems. This immediately yields the following theorem.

Theorem 6.1 (Lie-Poisson reduction of dynamics) *If $H : T^*G \rightarrow \mathbb{R}$ is a left (respectively, right) G -invariant function its restriction $H^- := H|_{\mathfrak{g}^*}$ (respectively, $H^+ := H|_{\mathfrak{g}^*}$) to $\mathfrak{g}^*$ satisfies*

$$H = H^- \circ \mathbf{J}_R \quad (\text{respectively} \quad H^+ = H \circ \mathbf{J}_L),$$

where $\mathbf{J}_R = T^*L_g \alpha_g$ and $\mathbf{J}_L = T^*R_g \alpha_g$ for all $\alpha_g \in T_g^*G$.

The flow F_t of X_H on T^*G and the flow, F_t^- of X_{H^-} on $\mathfrak{g}_-^*$ (respectively, F_t^+ of X_{H^+} on $\mathfrak{g}_+^*$) are related by $\mathbf{J}_R \circ F_t = F_t^- \circ \mathbf{J}_R$ (respectively, $\mathbf{J}_L \circ F_t = F_t^+ \circ \mathbf{J}_L$).

As the original Hamiltonian and the reduced Hamiltonian are related by a momentum map we can get some additional information using the fact that the momentum map is a conserved quantity.

Proposition 6.2 *Let $H : T^*G \rightarrow \mathbb{R}$ be left-invariant, $H^- = H|_{\mathfrak{g}^*}$, $\alpha(t) \in T_{g(t)}^*G$ an integral curve of X_H , $\mu(t) = \mathbf{J}_R(\alpha(t))$, and $\nu = \mathbf{J}_L(\alpha(t))$. Then*

$$\nu = \text{Ad}_{g(t)}^* \mu(t).$$

Proof The curve $\nu(t) := \mathbf{J}_L(\alpha(t)) = T_e^* R_{g(t)} \alpha(t)$ is constant by the Noether theorem, say equal to ν . As $\mu(t) = \mathbf{J}_R(\alpha(t)) = T_e^* L_{g(t)} \alpha(t)$, we get

$$\nu = T_e^* R_{g(t)} \alpha(t) = \left(T_e^* R_{g(t)} \circ T_{g(t)}^* L_{g^{-1}(t)} \right) \mu(t) = \text{Ad}_{g^{-1}(t)}^* \mu(t)$$

which proves the statement. $\blacksquare$

It is interesting to relate the reduced dynamics to its left (right) trivialization. Explicitly, T^*G is diffeomorphic to $G \times \mathfrak{g}^*$ via the **left trivialization** diffeomorphism

$$\lambda : T^*G \rightarrow G \times \mathfrak{g}^*, \quad \lambda(\alpha_g) := (g, T_e^* L_g(\alpha_g)) = (g, \mathbf{J}_R(\alpha_g)).$$

Since $\mathbf{J}_R$ is equivariant, λ is an equivariant diffeomorphism for the cotangent lift of the left translation and the following action of G on $G \times \mathfrak{g}^*$

$$g \cdot (h, \mu) := (gh, \mu).$$

Thus $(T^*G)/G$ is diffeomorphic to $(G \times \mathfrak{g}^*)/G$. As G does not act on $\mathfrak{g}^*$, it follows that $(G \times \mathfrak{g}^*)/G$ is equal to $\mathfrak{g}^*$ and we see again that $(T^*G)/G$ is diffeomorphic to $\mathfrak{g}^*$.

If X_H is the Hamiltonian vector field on T^*G for a left invariant Hamiltonian H , a lengthy but elementary computation (see, e.g. [MaRa94], Proposition 13.4.3) shows that the left trivialization $\lambda_* X_H$ equals

$$(\lambda_* X_H) = \left(T_e L_g \frac{\delta H^-}{\delta \mu}, \mu, \text{ad}_{\delta H^- / \delta \mu}^* \mu \right) \in T_g G \times T_\mu \mathfrak{g}^*, \quad (6.1)$$

which says that Hamilton's equations on $G \times \mathfrak{g}^*$ for the push forward Hamiltonian function $\lambda_* H$ and the push forward symplectic form $\lambda_* \Omega$ are

$$\dot{\mu} = \text{ad}_{\delta H^- / \delta \mu}^* \mu, \quad \dot{g} = T_e L_g \frac{\delta H^-}{\delta \mu}. \quad (6.2)$$

Note that the first equation is just the Lie-Poisson reduced Hamiltonian vector field and hence does not depend on $g \in G$. Once the first equation is solved, the second one yields a *linear* equation with time dependent coefficients, that is, the second equation is what one usually calls a “quadrature”. We summarize these remarks in the following *Reconstruction Theorem*.

Theorem 6.3 *Let $H : T^*G \rightarrow \mathbb{R}$ be a left-invariant Hamiltonian, $H^- := H|_{\mathfrak{g}^*}$, and $\mu(t)$ the integral curve of the Lie-Poisson equations*

$$\frac{d\mu}{dt} = \text{ad}_{\delta H^- / \delta \mu}^* \mu$$

with initial condition $\mu(0) = T_e^* L_{g_0}(\alpha_{g_0})$. Then the integral curve $\alpha(t) \in T_{g(t)}^* G$ of X_H with initial condition $\alpha(0) := \alpha_{g_0}$ is given by

$$\alpha(t) = T_{g(t)}^* L_{g(t)^{-1}} \mu(t),$$

where $g(t)$ is the solution of the equation

$$\frac{dg(t)}{dt} = T_e L_{g(t)} \frac{\delta H^-}{\delta \mu},$$

with initial condition $g(0) = g_0$.

Proof A curve $\alpha(t)$ is the unique integral curve of X_H with initial condition $\alpha(0) = \alpha_{g_0}$ if and only if

$$\lambda(\alpha(t)) = (g(t), T_e^* L_{g(t)} \alpha(t)) = (g(t), J_R(\alpha(t))) = (g(t), \mu(t))$$

is the unique integral curve of $\lambda_* X_H$ with initial condition

$$\lambda(\alpha(0)) = (g_0, T_e^* L_{g_0} \alpha_{g_0}).$$

So, the result follows from equation (6.1). ■

A similar statement holds for right invariant Hamiltonians by replacing everywhere “left” by “right” and $-$ by $+$ in the Lie-Poisson equations.

6.3 Coadjoint Orbits

In §4.2 we studied the internal structure of a Poisson manifold, namely, its stratification into a disjoint union of symplectic leaves. In this section we shall see that the symplectic leaves of the Poisson manifold $\mathfrak{g}^*$ endowed with the Lie-Poisson bracket are the connected components of the coadjoint orbits.

The **coadjoint orbit** $\mathcal{O}(\mu)$ through $\mu \in \mathfrak{g}^*$ is the subset of $\mathfrak{g}^*$ defined by

$$\mathcal{O}(\mu) := G \cdot \mu := \{ \text{Ad}_{g^{-1}}^* (\mu) : g \in G \}.$$

Like the orbit of any Lie group, $\mathcal{O}(\mu)$ is an immersed submanifold of $\mathfrak{g}^*$ but is not, in general, a submanifold of $\mathfrak{g}^*$. If G is compact then $\mathcal{O}(\mu)$ is a closed embedded submanifold of $\mathfrak{g}^*$. This is, in general, not true for an arbitrary Lie group. Coadjoint orbits of algebraic groups are also embedded submanifolds.

For any smooth Lie group action $\Phi : G \times M \rightarrow M$ on a manifold M , the **orbit** through a point $m \in M$ is the set

$$\mathcal{O}_m = \{ \Phi(g, m) \mid g \in G \} \subset M.$$

For $m \in M$, the **isotropy subgroup** of Φ at m is

$$G_m = \{g \in G \mid \Phi(g, m) = m\} \subset G.$$

Since the map $\Phi^m : G \rightarrow M$, given by $\Phi^m(g) := \Phi(g, m)$, is smooth, $G_m = (\Phi^m)^{-1}(m)$ is a closed subgroup and hence a Lie subgroup of G . The bijective map $[g] \in G/G_m \mapsto \Phi(g, m) \in \mathcal{O}_m$ induces a manifold structure on $\mathcal{O}_m$ that makes it diffeomorphic to the smooth homogeneous manifold G/G_m .

Recall that for $\xi \in \mathfrak{g}$ the family of diffeomorphisms $t \mapsto \Phi_{\exp(t\xi)}$ on M defines a flow and the corresponding vector field $\xi_M \in \mathfrak{X}(M)$ is the **infinitesimal generator** of the action. Thus

$$\xi_M(m) := \left. \frac{d}{dt} \right|_{t=0} \Phi_{\exp(t\xi)}(m).$$

This definition shows that the tangent space to the orbit $\mathcal{O}_m$ is given by

$$T_m \mathcal{O}_m = \{\xi_M(m) \mid \xi \in \mathfrak{g}\}.$$

We apply these general considerations to $M = \mathfrak{g}^*$ and the G -action the coadjoint action. Then the orbit through $\mu \in \mathfrak{g}^*$ is diffeomorphic to G/G_μ , where G_μ is the isotropy subgroup of μ

$$G_\mu = \{g \in G \mid \text{Ad}_g^* \mu = \mu\}.$$

We recall that for $\xi \in \mathfrak{g}$, the infinitesimal generator for the coadjoint action corresponding to ξ is given by (5.15), that is,

$$\xi_{\mathfrak{g}^*}(\mu) = -\text{ad}_\xi^* \mu.$$

Therefore

$$T_\mu \mathcal{O}_\mu = \{-\text{ad}_\xi^* \mu \mid \xi \in \mathfrak{g}\} = \mathfrak{g}_\mu^\circ,$$

where $\mathfrak{g}_\mu^\circ := \{\nu \in \mathfrak{g}^* \mid \langle \nu, \eta \rangle = 0 \text{ for all } \eta \in \mathfrak{g}_\mu\}$, $\mathfrak{g}_\mu = \{\xi \in \mathfrak{g} \mid \text{ad}_\xi^* \mu = 0\}$, and $\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \rightarrow \mathbb{R}$ is a strongly nondegenerate pairing (see [MaRa94], Proposition 14.2.1).

The next theorems show that the symplectic leaves of the Poisson manifold $\mathfrak{g}^*$ are the connected components of the coadjoint orbits and give explicitly the symplectic form. The proofs can be found in [MaRa94].

Theorem 6.4 *Let G be a Lie group and let $\mathcal{O} \subset \mathfrak{g}^*$ be a coadjoint orbit. Then $\mathcal{O}$ is a symplectic manifold relative to the **orbit symplectic form***

$$\omega^\pm(\mu)(\xi_{\mathfrak{g}^*}(\mu), \eta_{\mathfrak{g}^*}(\mu)) := \pm \langle \mu, [\xi, \eta] \rangle \quad (6.3)$$

for all $\mu \in \mathcal{O}$ and $\xi, \eta \in \mathfrak{g}$.

The symplectic form (6.3) is also known as the *Kostant-Kirillov-Souriau symplectic form*.

Theorem 6.5 *The Lie-Poisson bracket and the coadjoint orbit symplectic structure are consistent in the following sense: for $F, H : \mathfrak{g}^* \rightarrow \mathbb{R}$ and $\mathcal{O}$ a coadjoint orbit in $\mathfrak{g}^*$, we have*

$$\{F, H\}_+|_{\mathcal{O}} = \{F|_{\mathcal{O}}, H|_{\mathcal{O}}\}^+,$$

where, $\{\cdot, \cdot\}_+$ is the + Lie-Poisson bracket, while $\{\cdot, \cdot\}^+$ is the Poisson bracket defined by the + coadjoint symplectic orbit structure ω^+ on $\mathcal{O}$. Similarly,

$$\{F, H\}_-|_{\mathcal{O}} = \{F|_{\mathcal{O}}, H|_{\mathcal{O}}\}^-.$$

We summarize below some results for coadjoint orbits:

- For $\mu, \nu \in \mathfrak{g}_-^*$ and $H : \mathfrak{g}^* \rightarrow \mathbb{R}$, the Hamiltonian vector field for H is $X_H(\nu) = \text{ad}_{\delta H / \delta \nu}^*(\nu)$. Therefore, if $\nu \in \mathcal{O}$ then $X_H(\nu)$ is tangent to $\mathcal{O}$. So the trajectory of X_H starting at $\mu \in \mathcal{O} \subset \mathfrak{g}^*$ stays in $\mathcal{O}$.
- Recall that a function $C \in \mathcal{F}(\mathfrak{g}^*)$ is a Casimir if and only if $0 = X_C(\mu) = \text{ad}_{\delta C / \delta \mu}^*(\mu)$. Thus if C is a Casimir of $\mathfrak{g}^*$ then $\delta C / \delta \mu \in \mathfrak{g}_\mu$ for all $\mu \in \mathfrak{g}^*$.
- **(Duflo-Vergne Theorem)** Let $\mathfrak{g}$ be a finite-dimensional Lie algebra with dual $\mathfrak{g}^*$ and let $r := \min\{\dim \mathfrak{g}_\mu \mid \mu \in \mathfrak{g}^*\}$. The set $\{\mu \in \mathfrak{g}^* \mid \dim \mathfrak{g}_\mu = r\}$ is Zariski open and thus open and dense in the usual topology of $\mathfrak{g}^*$. If $\dim \mathfrak{g}_\mu = r$, then $\mathfrak{g}_\mu$ is Abelian.
- If $C \in \mathcal{F}(\mathfrak{g}^*)$ is Ad^* -invariant, i.e $C(\text{Ad}_{g^{-1}}^* \mu) = C(\mu)$, then the differentiation of this equality with respect to g at $g = e$ shows that C is a Casimir function. Thus, a function that is constant on coadjoint orbits is necessarily a Casimir function.

In general Ad^* -invariance of C is a stronger condition than C being a Casimir. A theorem of Kostant gives a characterization of which Ad^* -invariant functions are Casimirs. Namely, an Ad^* -invariant function C is a Casimir if and only if $\delta C / \delta \mu$ lies in the center of $\mathfrak{g}_\mu$ for all $\mu \in \mathfrak{g}^*$ (see [MaRa94], Proposition 14.4.4).

The rest of this section is dedicated to working out a few examples.

(1) Rotation group. Recall from §5.1 that the coadjoint action of $SO(3)$ on $\mathfrak{so}(3)^*$ has the expression $\text{Ad}_{A^{-1}}^* \tilde{\Pi} = (A\Pi)^\sim$, where the isomorphism $\sim : \mathbb{R}^3 \rightarrow \mathfrak{so}(3)^*$ is given by $\tilde{\Pi}(\hat{\mathbf{u}}) := \Pi \cdot \mathbf{u}$ for any $\mathbf{u} \in \mathbb{R}^3$ and $\hat{\cdot} : (\mathbb{R}^3, \times) \rightarrow (\mathfrak{so}(3), [\cdot, \cdot])$ is the Lie algebra isomorphism (1.13). Therefore, the coadjoint orbit $\mathcal{O} = \{A\Pi \mid A \in SO(3)\} \subset \mathbb{R}^3$ of $SO(3)$ through $\Pi \in \mathbb{R}^3$ is a 2-sphere of radius $\|\Pi\|$.

To compute the coadjoint action of $\mathfrak{so}(3)$ on its dual, let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ and note that

$$\begin{aligned} \langle \text{ad}_{\hat{\mathbf{u}}}^* \tilde{\mathbf{\Pi}}, \hat{\mathbf{v}} \rangle &= \langle \tilde{\mathbf{\Pi}}, [\hat{\mathbf{u}}, \hat{\mathbf{v}}] \rangle = \langle \tilde{\mathbf{\Pi}}, (\mathbf{u} \times \mathbf{v})^\wedge \rangle = \mathbf{\Pi} \cdot (\mathbf{u} \times \mathbf{v}) \\ &= (\mathbf{\Pi} \times \mathbf{u}) \cdot \mathbf{v} = \langle \mathbf{\Pi} \times \mathbf{u} \rangle^\sim, \hat{\mathbf{v}} \rangle, \end{aligned}$$

which shows that $\text{ad}_{\hat{\mathbf{u}}}^* \tilde{\mathbf{\Pi}} = (\mathbf{\Pi} \times \mathbf{u})^\sim$, proving (1.21). Therefore, $T_{\mathbf{\Pi}}\mathcal{O} = \{\mathbf{\Pi} \times \mathbf{u} \mid \mathbf{u} \in \mathbb{R}^3\}$ as expected, since the plane perpendicular to $\mathbf{\Pi}$, that is, the tangent space to the sphere centered at the origin of radius $\|\mathbf{\Pi}\|$, is indeed given by $\{\mathbf{\Pi} \times \mathbf{u} \mid \mathbf{u} \in \mathbb{R}^3\}$.

The minus orbit symplectic structure on $\mathcal{O}$ is given hence by

$$\omega^-(\mathbf{\Pi})(\mathbf{\Pi} \times \mathbf{u}, \mathbf{\Pi} \times \mathbf{v}) = -\mathbf{\Pi} \cdot (\mathbf{u} \times \mathbf{v}).$$

How does this exactly relate to the area form on the sphere $\mathcal{O}$? To see this, recall that the oriented area of a planar parallelogram spanned by two vectors $\mathbf{a}, \mathbf{b}$ (in this order) is given by $\mathbf{a} \times \mathbf{b}$. Thus, the oriented area spanned by $\mathbf{\Pi} \times \mathbf{u}$ and $\mathbf{\Pi} \times \mathbf{v}$ is

$$(\mathbf{\Pi} \times \mathbf{u}) \times (\mathbf{\Pi} \times \mathbf{v}) = (\mathbf{\Pi} \cdot (\mathbf{u} \times \mathbf{v}))\mathbf{\Pi}.$$

The area element dS on the sphere assigns to each ordered pair of tangent vectors $\mathbf{a}, \mathbf{b}$ the number $dS(\mathbf{a}, \mathbf{b}) = \mathbf{n} \cdot (\mathbf{a} \times \mathbf{b})$, where $\mathbf{n}$ is the outward unit normal. Therefore

$$dS(\mathbf{\Pi} \times \mathbf{u}, \mathbf{\Pi} \times \mathbf{v}) = \frac{\mathbf{\Pi}}{\|\mathbf{\Pi}\|} \cdot ((\mathbf{\Pi} \times \mathbf{u}) \times (\mathbf{\Pi} \times \mathbf{v})) = \|\mathbf{\Pi}\| \mathbf{\Pi} \cdot (\mathbf{u} \times \mathbf{v}).$$

This shows that

$$\omega^-(\mathbf{\Pi}) = -\frac{1}{\|\mathbf{\Pi}\|} dS.$$

We have computed in §1.2 the kinetic energy of a heavy top. If the center of mass is the point of suspension of the top, that is $\ell = 0$, then one obtains the free rigid body and the total energy is the kinetic energy given by (1.38)

$$K(A, \dot{A}) = -\frac{1}{4} \text{trace}((JA^{-1}\dot{A} + A^{-1}\dot{A}J)A^{-1}\dot{A}).$$

This expression is on $T\mathcal{SO}(3)$ but using the Riemannian metric on $SO(3)$ obtained by left translating the inner product (1.34) on $\mathfrak{so}(3)$ one obtains a bundle metric on $T^*\mathcal{SO}(3)$ whose kinetic energy is the Hamiltonian of the free rigid body. Its restriction to $\mathfrak{so}(3)^* \cong \mathbb{R}^3$ is given by

$$H(\mathbf{\Pi}) = \frac{1}{2} \mathbf{\Pi} \cdot \mathbb{I}^{-1} \mathbf{\Pi}.$$

By (6.2) the geodesic equations in left trivialization are

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbb{I}^{-1} \mathbf{\Pi} \quad \text{and} \quad \dot{A} = A \mathbb{I}^{-1} \mathbf{\Pi}$$

since $\nabla H(\mathbf{\Pi}) = \mathbb{I}^{-1} \mathbf{\Pi}$. The first equation is the Lie-Poisson equation on $\mathbb{R}^3$ and the second one is a linear equation with time dependent coefficients and gives the attitude matrix of the body. Since the concentric spheres centered at the origin are the coadjoint orbits, the integral curves of the first equation necessarily lie on them. In addition, the first equation is Hamiltonian on these spheres relative to the orbit symplectic form and the Hamiltonian H . Thus the solutions of the Lie-Poisson equation $\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbb{I}^{-1} \mathbf{\Pi}$ are obtained by intersecting each sphere with the ellipsoids $H(\mathbf{\Pi}) = \text{constant}$. This shows that there are six equilibria four of which are stable (rotation about the long and short axes) and two of which are saddles (rotation about the middle axis).

(2) Affine group on $\mathbb{R}$. Consider the Lie group G of the transformations $T : \mathbb{R} \rightarrow \mathbb{R}$, $T(x) = ax + b$ with $a \neq 0$. We can identify G with the set of pairs (a, b) . As

$$(T_2 \circ T_1)(x) = T_2(a_1x + b_1) = (a_2a_1x + a_2b_1 + b_2),$$

then the group multiplication on $G = \{(a, b) \in \mathbb{R}^2 \mid a \neq 0\}$ is given by

$$(a_2, b_2)(a_1, b_1) = (a_2a_1, a_2b_1 + b_2).$$

The identity element of G is $e = (1, 0)$ and the inverse of (a, b) is $(a, b)^{-1} = (1/a, -b/a)$. The inner conjugation automorphism $\text{AD}_{(a,b)}$ is given by

$$\begin{aligned} \text{AD}_{(a,b)}(c, d) &= (a, b)(c, d)(a, b)^{-1} = (a, b)(c, d) \left(\frac{1}{a}, \frac{-b}{a} \right) \\ &= (c, -cb + ad + b). \end{aligned}$$

The adjoint action is obtained differentiating $\text{AD}_{(a,b)}(c, d)$ with respect to (c, d) at the identity $e = (1, 0)$ in the direction to (u, v) which gives

$$\begin{aligned} \text{Ad}_{(a,b)}(u, v) &= \left. \frac{d}{dt} \right|_{t=0} \text{AD}_{(a,b)}(c + tu, d + tv) \\ &= \left. \frac{d}{dt} \right|_{t=0} (c + tu, -(c + tu)b + a(d + tv) + b) \\ &= (u, av - bu). \end{aligned}$$

Thus, the adjoint orbit through (u, v) is $\{u\} \times \mathbb{R}$ if $(u, v) \neq (0, 0)$ and is the origin if $(u, v) = (0, 0)$. We shall see below that the coadjoint orbits are very different.

The underlying vector space of the Lie algebra $\mathfrak{g}$ of G is $\mathbb{R}^2$ since G is obviously open in $\mathbb{R}^2$. The Lie bracket is obtained by differentiating $\text{Ad}_{(a,b)}(u, v)$ with respect to (a, b) at the identity in the direction of (r, s) which gives

$$\left. \frac{d}{dt} \right|_{t=0} (u, (a + tr)v - (b + ts)u) = (0, rv - su) = [(r, s), (u, v)], \quad (6.4)$$

for $(r, s), (u, v) \in \mathfrak{g} = \mathbb{R}^2$.

Consider the pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$ to be the standard inner product in $\mathbb{R}^2$, that is, $\mathfrak{g}^* = \mathbb{R}^2$. So, for $(\alpha, \beta) \in \mathfrak{g}^*$ and $(u, v) \in \mathfrak{g}$, we have

$$\begin{aligned} \langle \text{Ad}_{(a,b)}^*(\alpha, \beta), (u, v) \rangle &= \langle (\alpha, \beta), \text{Ad}_{(a,b)}(u, v) \rangle \\ &= \langle (\alpha, \beta), (u, av - bu) \rangle = (\alpha - \beta b)u + \beta av, \end{aligned}$$

which shows that

$$\text{Ad}_{(a,b)}^*(\alpha, \beta) = (\alpha - \beta b, \beta a).$$

So, if $\beta = 0$, the coadjoint orbit through (α, β) is the single point $(\alpha, 0)$, while if $\beta \neq 0$, the coadjoint orbit through (α, β) is $\mathbb{R}^2$ minus the α -axis; this latter orbit is open in $\mathfrak{g}^*$. This example shows that the dimensions of the adjoint and coadjoint orbits can be different.

To compute the coadjoint action of $\mathfrak{g}$ on $\mathfrak{g}^*$ we use (6.4) and the Euclidean inner product for the duality pairing to get

$$\begin{aligned} \langle \text{ad}_{(u,v)}^*(\alpha, \beta), (r, s) \rangle &= \langle (\alpha, \beta), [(u, v), (r, s)] \rangle = \langle (\alpha, \beta), (0, su - rv) \rangle \\ &= su\beta - rv\beta = \langle (-v\beta, u\beta), (r, s) \rangle, \end{aligned}$$

that is,

$$\text{ad}_{(u,v)}^*(\alpha, \beta) = (-v\beta, u\beta).$$

This shows that $T_{(\alpha,\beta)}\mathcal{O}_{(\alpha,\beta)}$ equals $\{(0,0)\}$ if $\beta = 0$ or $\mathbb{R}^2$ if $\beta \neq 0$, as expected.

The minus Lie-Poisson bracket of $F, H : \mathfrak{g}^* \rightarrow \mathbb{R}$ is hence given by

$$\{F, H\}_-(\alpha, \beta) = -\beta \left(\frac{\partial F}{\partial \alpha} \frac{\partial H}{\partial \beta} - \frac{\partial F}{\partial \beta} \frac{\partial H}{\partial \alpha} \right).$$

The orbit symplectic structure (6.3) for the open orbit $\mathcal{O}$, that is, the orbit passing through (α, β) with $\beta \neq 0$, equals

$$\begin{aligned} \omega^-(\alpha, \beta) \left(\text{ad}_{(r,s)}^*(\alpha, \beta), \text{ad}_{(u,v)}^*(\alpha, \beta) \right) &= -\langle (\alpha, \beta), [(r, s), (u, v)] \rangle \\ &= -\beta(rv - su), \end{aligned}$$

or, in canonical coordinates (α, β) on $\mathcal{O}_{(\alpha, \beta)}$,

$$\omega^- = -\frac{1}{\beta} \mathbf{d}\alpha \wedge \mathbf{d}\beta.$$

Given a smooth function $H : \mathfrak{g}^* = \mathbb{R}^2 \rightarrow \mathbb{R}$, the Hamiltonian vector field relative to the minus Lie-Poisson bracket is given by

$$X_H(\alpha, \beta) = \beta(-\partial H / \partial \beta, \partial H / \partial \alpha).$$

As is obvious from this expression, an integral curve whose initial condition (α_0, β_0) satisfies $\beta_0 > 0$ (respectively $\beta_0 < 0$) will satisfy the same condition for all time. This verifies the standard fact that the open symplectic leaves are invariant under the flow. In addition, all points on the line $\beta = 0$ are equilibria, that is, the zero dimensional orbits are also invariant under the flow, as expected.

(3) The group $\text{Diff}_{\text{vol}}(D)$, vorticity representation. Let $G = \text{Diff}_{\text{vol}}(D)$ be the group of volume preserving diffeomorphisms of a k -dimensional oriented Riemannian manifold (D, g) with smooth boundary ∂D . The Riemannian volume form μ on D is the unique volume form on D which is equal to 1 on all positively oriented g -orthonormal bases of tangent vectors at all points of D .

The Riemannian volume μ naturally induces a Riemannian volume form on the boundary ∂D (relative to the induced metric) given in the following way. Let $i : \partial D \rightarrow D$ be the inclusion. If $v_1, \dots, v_{k-1} \in T_x(\partial D)$ is a basis such that $\mu(x)(n, v_1, \dots, v_{k-1}) > 0$ for n the outward pointing unit normal, define $\mu_{\partial D}(x)(v_1, \dots, v_{k-1}) := \mu(x)(n, v_1, \dots, v_{k-1})$ and extend it by skew symmetry and multilinearity to any other k -tuple of tangent vectors in $T_x \partial D$. Recall that the normal n is **pointing outward** if, in a (and hence any) chart on D intersecting ∂D whose image lies in the upper half space $\{(x^1, \dots, x^k) \in \mathbb{R}^k \mid x^k \geq 0\}$, the vector n is collinear with $-\partial/\partial x^k$. The key relation that relates g , μ , and $\mu_{\partial D}$ is

$$i^*(\mathbf{i}_v \mu) = g(v, n) \mu_{\partial D} \quad (6.5)$$

for any $v \in \mathfrak{X}(D)$.

As we have seen in §5.1, formally, the Lie algebra of G is the Lie algebra $\mathfrak{X}_{\text{div}}(D)$ of divergence free vector fields tangent to the boundary ∂D , endowed with minus the usual bracket of vector fields. We have identified the dual $\mathfrak{X}_{\text{div}}(D)^*$ with $\mathbf{d}\Omega^1(D)$, assuming that the first cohomology group of D is zero. The weak pairing (see §4.3) between $\mathfrak{X}_{\text{div}}(D)$ and $\mathbf{d}\Omega^1(D)$ was given by

$$(u, \omega) \in \mathfrak{X}_{\text{div}}(D) \times \mathbf{d}\Omega^1(D) \mapsto \int_D \alpha(u) \mu \in \mathbb{R}, \quad \text{for } \omega := \mathbf{d}\alpha \quad (6.6)$$

and the plus Lie-Poisson bracket by

$$\{F, H\}(\omega) = \int_D \omega \left(\frac{\delta F}{\delta \omega}, \frac{\delta H}{\delta \omega} \right) \mu, \quad (6.7)$$

for any $F, H \in \mathcal{F}(\mathfrak{X}_{\text{div}}(D))$.

The coadjoint action of the diffeomorphism $\eta \in \text{Diff}_{\text{vol}}(D)$ on $\omega \in \mathbf{d}\Omega^1(D)$ is given by $\text{Ad}_{\eta^{-1}}^* \omega = \eta_* \omega$ (see §5.1). Therefore, the coadjoint orbit passing through ω equals $\mathcal{O} = \{\eta_* \omega \mid \eta \in \text{Diff}_{\text{vol}}(D)\}$. The coadjoint action of $\mathfrak{X}_{\text{div}}(D)$ on $\mathbf{d}\Omega^1(D) \cong \mathfrak{X}_{\text{div}}(D)^*$ is hence given by

$$-\text{ad}_v^* \omega = -\mathcal{L}_v \omega. \quad (6.8)$$

Note the $-$ sign on the right hand side. Normally, one should expect a $+$ sign since

$$\left. \frac{d}{dt} \right|_{t=0} (\eta_t)_* \omega = -\mathcal{L}_v \omega,$$

where η_t is the flow of v . However, all formulas derived abstractly use the *left* Lie algebra and, as we have seen in Section 5.1, Example 2, the left Lie algebra bracket on vector fields is *minus* the usual Lie bracket. This is why one needs to change the sign in (6.8). One can easily derive (6.8) directly: for any $u, v \in \mathfrak{X}_{\text{div}}(D)$ and $\omega = \mathbf{d}\alpha \in \mathbf{d}\Omega^1(D)$, the identities $\text{ad}_v u = -[v, u]$ (note the minus sign), $\mathcal{L}_v(\alpha(u)) = (\mathcal{L}_v \alpha)(u) + \alpha([v, u])$, (6.5), and the Stokes theorem give

$$\begin{aligned} \langle \text{ad}_v^* \omega, u \rangle &= \langle \omega, \text{ad}_v u \rangle = \langle \omega, -[v, u] \rangle = - \int_D \alpha([v, u]) \mu \\ &= \int_D (\mathcal{L}_v \alpha)(u) \mu - \int_D \mathcal{L}_v(\alpha(u)) \mu \\ &= \int_D (\mathcal{L}_v \alpha)(u) \mu - \int_D \mathcal{L}_v(\alpha(u)) \mu \\ &= \int_D (\mathcal{L}_v \alpha)(u) \mu - \int_{\partial D} \alpha(u) \mathbf{i}_v \mu \\ &= \langle \mathbf{d}\mathcal{L}_v \alpha, u \rangle = \langle \mathcal{L}_v \omega, u \rangle, \end{aligned}$$

which, by weak non-degeneracy of the pairing, proves (6.8).

Thus, the tangent space to the orbit $\mathcal{O}$ is

$$T_\omega \mathcal{O} = \{ \mathcal{L}_v \omega = \mathbf{d}\mathbf{i}_v \omega \mid v \in \mathfrak{X}_{\text{div}}(D) \}.$$

The orbit symplectic structure (6.3) has therefore the expression

$$\omega^+(\omega)(\mathcal{L}_u \omega, \mathcal{L}_v \omega) = - \int_D \langle \alpha, [u, v] \rangle \mu \quad \text{for } \omega = \mathbf{d}\alpha.$$

However, if $\omega = \mathbf{d}\alpha$, for any $u, v \in \mathfrak{X}_{\text{div}}(D)$ we have

$$\langle \alpha, [u, v] \rangle + \omega(u, v) = u[\langle \alpha, v \rangle] - v[\langle \alpha, u \rangle]$$

so that by the Stokes theorem and (6.5)

$$\begin{aligned} \int_D (\langle \alpha, [u, v] \rangle + \omega(u, v)) \mu &= \int_D u[\langle \alpha, v \rangle] \mu - \int_D v[\langle \alpha, u \rangle] \mu \\ &= \int_D \mathcal{L}_u(\langle \alpha, v \rangle \mu) - \int_D \mathcal{L}_v(\langle \alpha, u \rangle \mu) \\ &= \int_D \mathbf{d}\mathbf{i}_u(\langle \alpha, v \rangle \mu) - \int_D \mathbf{d}\mathbf{i}_v(\langle \alpha, u \rangle \mu) \\ &= \int_{\partial D} i^*(\mathbf{i}_u \langle \alpha, v \rangle \mu) - \int_{\partial D} i^*(\mathbf{i}_v \langle \alpha, u \rangle \mu) \\ &= \int_{\partial D} (i^* \langle \alpha, v \rangle) i^*(\mathbf{i}_u \mu) - \int_{\partial D} (i^* \langle \alpha, u \rangle) i^*(\mathbf{i}_v \mu) \\ &= \int_{\partial D} (i^* \langle \alpha, v \rangle) g(u, n) \mu_{\partial D} - \int_{\partial D} (i^* \langle \alpha, u \rangle) g(v, n) \mu_{\partial D} = 0 \end{aligned}$$

since, by hypothesis, $g(u, n) = g(v, n) = 0$ on ∂D . Therefore, the orbit symplectic structure is given by

$$\omega^+(\omega)(\mathcal{L}_u \omega, \mathcal{L}_v \omega) = \int_D \omega(u, v) \mu. \quad (6.9)$$

Let us compute the plus Lie-Poisson equations for the geodesic flow, that is, the equations $\dot{F} = \{F, H\}$ for any $F \in \mathcal{F}(\text{Diff}_{\text{vol}}(D))$, where

$$H(\omega) = \frac{1}{2} \int_D \|v\|^2 \mu = \frac{1}{2} \int_D v^b(v) \mu = \frac{1}{2} \langle \mathbf{d}v^b, v \rangle \quad (6.10)$$

by (6.6). Define the **vorticity** associated to the spatial velocity vector field v of the incompressible perfect fluid by $\omega := \mathbf{d}v^b$. Note that the pairing (6.6) satisfies $\langle \mathbf{d}u^b, v \rangle = \langle \mathbf{d}v^b, u \rangle$ for any $u, v \in \mathfrak{X}_{\text{div}}(D)$, so that letting $\delta\omega := \mathbf{d}(\delta v)^b$ we get

$$\left\langle \delta\omega, \frac{\delta H}{\delta \omega} \right\rangle = \mathbf{D}H(\omega) \cdot \delta\omega = \langle \delta\omega, v \rangle,$$

that is, $\delta H / \delta \omega = v$. Thus, the plus Lie-Poisson equations (4.17) become

$$\frac{\partial \omega}{\partial t} + \mathcal{L}_v \omega = 0, \quad \text{where } \omega := \mathbf{d}v^b, \quad (6.11)$$

which are the *Euler equations for an incompressible homogeneous perfect fluid in vorticity formulation*. Therefore the geodesic $\eta_t \in \text{Diff}_{\text{vol}}(D)$ is given by solving the equation $\partial \eta_t / \partial t = v_t \circ \eta_t$ with the velocity v_t found after solving

for ω equation (6.11) and then inverting the relation $\omega = \mathbf{d}v^\flat$ with boundary condition $v \cdot n = 0$, where n is the outward unit normal to the boundary ∂D .

Equation (6.11) is equivalent to any of the following statements:

- (i) *The vorticity ω_t is transported by the flow.* Indeed, if ω_0 is an initial condition at $t = 0$ of (6.11), then

$$\left. \frac{d}{dt} \right|_{t=0} (\eta_t)_* \omega_0 = -(\eta_t)_* \mathcal{L}_v \omega_0 = -\mathcal{L}_v (\eta_t)_* \omega_0,$$

which shows that $\omega_t = (\eta_t)_* \omega_0$ solves (6.11). By uniqueness, this is the only solution of (6.11) with ω_0 as initial condition.

- (ii) *Solution curves of the vorticity equation (6.11) remain on coadjoint orbits in $\mathfrak{X}_{\text{div}}^*(D)$.* Indeed, since the solution of (6.11) is $\omega_t = (\eta_t)_* \omega_0$, where η_t is the flow of v , it follows that ω_t necessarily lies on the coadjoint orbit containing ω_0 .

- (iii) *Kelvin's circulation theorem: For any loop C in D bounding a surface S , the circulation*

$$\int_{C_t} v_t^\flat = \text{constant},$$

where $C_t := \eta_t(C)$ and η_t is the flow of v . Indeed, by change of variables and Stokes' theorem, for $S_t := \eta_t(S)$, we have

$$\int_{C_t} v_t^\flat = \int_{S_t} \mathbf{d}v_t^\flat = \int_{S_t} \omega_t = \int_{S_t} (\eta_t)_* \omega_0 = \int_S \omega_0 = \text{constant}.$$

(4) The group $\text{Diff}_{\text{vol}}(D)$, velocity representation. In §4.3 we have also identified $\mathfrak{X}_{\text{div}}(D)$ with itself by the weak L^2 pairing $\langle \cdot, \cdot \rangle : \mathfrak{X}_{\text{div}}(D) \times \mathfrak{X}_{\text{div}}(D) \rightarrow \mathbb{R}$ given by

$$\langle u, v \rangle = \int_D g(u, v) \mu.$$

The plus Lie-Poisson bracket is given by (4.9), namely

$$\{F, H\}(v) = - \int_D g \left(v, \left[\frac{\delta F}{\delta v}, \frac{\delta H}{\delta v} \right] \right) \mu,$$

for $F, H \in \mathcal{F}(\text{Diff}_{\text{vol}}(D))$. Using the change of variables formula and the fact

that η is volume preserving, the coadjoint action is computed to be

$$\begin{aligned} \langle \text{Ad}_{\eta^{-1}}^* u, v \rangle &= \langle u, \text{Ad}_{\eta^{-1}} v \rangle = \langle u, \eta^* v \rangle = \int_D g(u, \eta^* v) \mu \\ &= \int_D g(u, T\eta^{-1} \circ v \circ \eta) \mu \\ &= \int_D g((T\eta^{-1})^\dagger \circ u \circ \eta^{-1}, v) \mu = \langle \mathbb{P}((T\eta^{-1})^\dagger \circ u \circ \eta^{-1}), v \rangle, \end{aligned}$$

where $(T\eta)^\dagger$ is the pointwise adjoint relative to the metric g of the fiberwise linear map $T\eta : TD \rightarrow TD$ and $\mathbb{P} : \mathfrak{X}(D) \rightarrow \mathfrak{X}_{\text{div}}(D)$ is the Helmholtz projector. In defining $\mathbb{P}$ we used the *Helmholtz decomposition* which is the Hodge decomposition on forms for the special case of one-forms and formulated in terms of vector fields: *every vector field on D can be uniquely decomposed as an L^2 orthogonal sum of a divergence free vector field tangent to the boundary and the gradient of a function*. Thus, since the L^2 pairing is weakly nondegenerate on $\mathfrak{X}_{\text{div}}(D)$, we conclude

$$\text{Ad}_{\eta^{-1}}^* u = \mathbb{P}((T\eta^{-1})^\dagger \circ u \circ \eta^{-1})$$

for any $u \in \mathfrak{X}_{\text{div}}(D) \cong \mathfrak{X}_{\text{div}}(D)^*$. Therefore, the coadjoint orbit $\mathcal{O}$ passing through $w \in \mathfrak{X}_{\text{div}}(D)$ equals $\mathcal{O} = \{\mathbb{P}((T\eta)^\dagger \circ w \circ \eta) \mid \eta \in \text{Diff}_{\text{vol}}(D)\}$. Compared to the vorticity representation, the expression of the coadjoint action and of the coadjoint orbit are more complicated. This also shows that different identifications of the dual can give rise to different expressions for the coadjoint orbits. We shall remark below why it is important to work with both representations when considering the Euler equations.

The orbit symplectic structure (6.3) has hence the expression

$$\omega^+(w)(u_{\mathcal{O}}(w), v_{\mathcal{O}}(w)) = - \int_D g(w, [u, v]) \mu \quad (6.12)$$

for any $w \in \mathcal{O}$ and any $u, v \in \mathfrak{X}_{\text{div}}(D)$.

It is interesting to give the expression of the value of the infinitesimal generator $v_{\mathcal{O}}(w)$ at $w \in \mathcal{O}$. If η_t is the flow of w , we have (see [MaRa95])

$$\left. \frac{d}{dt} \right|_{t=0} \left[(T\eta_t)^\dagger \circ v \circ \eta_t \right] = \nabla_w v + (v \cdot \nabla w)^\sharp,$$

where $v \cdot \nabla w$ is the contraction of v with the upper index of ∇w , that is, if

$$(\nabla w)_k^j = \frac{\partial w^j}{\partial x^k} + \Gamma_{k\ell}^j w^\ell$$

then

$$(v \cdot (\nabla w))_k = g_{mj} v^m (\nabla w)_k^j.$$

Thus $v \cdot \nabla w \in \Omega^1(D)$ and its associated vector field using the metric is $(v \cdot \nabla w)^\sharp \in \mathfrak{X}(D)$. This shows that

$$v_{\mathcal{O}}(w) = -\text{ad}_v^* w = -\mathbb{P}(\nabla_w v + (v \cdot \nabla w)^\sharp). \quad (6.13)$$

We shall not compute here the plus Lie-Poisson equations for the Hamiltonian (6.10) because we shall carry out an identical computation in Section 6.4, Example 4, when dealing with the Euler-Poincaré equations. We mention only that they are the classical Euler equations for an incompressible homogeneous perfect fluid

$$\begin{cases} \frac{\partial v}{\partial t} + \nabla_v v = -\nabla p \\ \text{div } v = 0, \quad v \cdot n = 0, \end{cases} \quad (6.14)$$

where p is the **pressure** and n is the outward unit normal to the boundary ∂D . The pressure p exists and is determined up to a constant since it is a solution of the following Neumann problem

$$\Delta p = \text{div } \nabla_v v \quad \text{in } D, \quad \text{with} \quad \frac{\partial p}{\partial n} = (\nabla_v v) \cdot n \quad \text{on } \partial D$$

obtained by taking the divergence and the inner product with n on the boundary of the Euler equation; here $\Delta := -\text{div} \circ \text{grad}$ is the Laplacian on functions. Thus p is a nonlinear functional of the Eulerian velocity v .

Since the solutions of Lie-Poisson equations always lie on coadjoint orbits, we can conclude that the solution of the Euler equations with initial condition $v_0 \in \mathfrak{X}_{\text{div}}(D)$ necessarily lies on the coadjoint orbit $\{\mathbb{P}((T\eta)^\dagger \circ v_0 \circ \eta) \mid \eta \in \text{Diff}_{\text{vol}}(D)\}$.

6.4 Euler-Poincaré Reduction

A Hamiltonian H on T^*Q often arises from a Lagrangian L on TQ . Namely, as seen in section 3.3, the two formalisms are equivalent when the Legendre transform $\mathbb{F}L : TQ \rightarrow T^*Q$ is a diffeomorphism. In the previous sections we saw how to reduce a Poisson bracket on T^*G to one on $\mathfrak{g}^*$ via Lie-Poisson reduction and how Hamiltonian dynamics on T^*G induces Lie-Poisson dynamics on $\mathfrak{g}^*$. In this section we study the passage from TG to $\mathfrak{g}$ in a context appropriate for the Lagrangian formalism. As Lagrangian mechanics is based on variational principles, it is natural that the basic objects to be reduced here are the variational principles rather than the Poisson bracket or symplectic form as was the case in Lie-Poisson reduction.

Euler-Poincaré Reduction starts with a left (respectively right) invariant Lagrangian $L : TG \rightarrow \mathbb{R}$ on the tangent bundle of a Lie group G . Recall that this means that $L(T_g L_h(v)) = L(v)$, respectively $L(T_g R_h(v)) = L(v)$, for all $g, h \in G$ and all $v \in T_g G$.

Theorem 6.6 *Let G be a Lie group, $L : TG \rightarrow \mathbb{R}$ a left-invariant Lagrangian, and $l := L|_{\mathfrak{g}} : \mathfrak{g} \rightarrow \mathbb{R}$ be its restriction to $\mathfrak{g}$. For a curve $g(t) \in G$, let $\xi(t) = g(t)^{-1} \cdot \dot{g}(t) := T_{g(t)} L_{g(t)^{-1}} \dot{g}(t) \in \mathfrak{g}$. Then the following are equivalent:*

- (i) $g(t)$ satisfies the Euler-Lagrange equations for L on G .
- (ii) The variational principle

$$\delta \int_a^b L(g(t), \dot{g}(t)) dt = 0$$

holds, for variations with fixed endpoints.

- (iii) The **Euler-Poincaré equations** hold:

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi}.$$

- (iv) The variational principle

$$\delta \int_a^b l(\xi(t)) dt = 0$$

holds on $\mathfrak{g}$, using variations of the form $\delta \xi = \dot{\eta} + [\xi, \eta]$, where $\eta(t)$ is an arbitrary path in $\mathfrak{g}$ that vanishes at the endpoints, i.e $\eta(a) = \eta(b) = 0$.

Proof The equivalence (i) $\iff$ (ii) is the variational principle of Hamilton (see Theorem 3.5). To show that (ii) $\iff$ (iv) we need to compute variations $\delta \xi$ induced by δg . We will do it for matrix groups to simplify the exposition.

Let $\xi = g^{-1} \dot{g}$ and g_ϵ a family of curves in G such that $g_0(t) = g(t)$ and denote $\delta g := (dg_\epsilon(t)/d\epsilon)|_{\epsilon=0}$. Then we have

$$\delta \xi = \frac{d}{d\epsilon} \Big|_{\epsilon=0} (g_\epsilon^{-1} \dot{g}_\epsilon) = -g^{-1} (\delta g) g^{-1} \dot{g} + g^{-1} \frac{d^2 g}{dt d\epsilon} \Big|_{\epsilon=0}. \quad (6.15)$$

Let $\eta := g^{-1} \delta g$, that is, $\eta(t)$ is an arbitrary curve in $\mathfrak{g}$ with the only restriction that it vanishes at the endpoints. Then we get

$$\frac{d\eta}{dt} = \frac{d}{dt} \left(g^{-1} \frac{d}{d\epsilon} \Big|_{\epsilon=0} g_\epsilon \right) = -g^{-1} \dot{g} g^{-1} (\delta g) + g^{-1} \frac{d^2 g}{dt d\epsilon} \Big|_{\epsilon=0}. \quad (6.16)$$

Taking the difference of (6.15) and (6.16) we get

$$\delta \xi - \dot{\eta} = -g^{-1} (\delta g) g^{-1} \dot{g} + g^{-1} \dot{g} g^{-1} (\delta g) = \xi \eta - \eta \xi = [\xi, \eta],$$

that is, $\delta\xi = \dot{\eta} + [\xi, \eta]$.

Left invariance of L together with the formula just deduced prove the equivalence of (ii) and (iv).

To avoid the assumption that G is a matrix group and to do the general case for any Lie group, the same proof works using the following lemma.

Lemma 6.7 *Let $g : U \subset \mathbb{R}^2 \rightarrow G$ be a smooth map and denote its partial derivatives by*

$$\xi(t, \varepsilon) := T_{g(t, \varepsilon)} L_{g(t, \varepsilon)^{-1}} \frac{\partial g(t, \varepsilon)}{\partial t}, \quad \eta(t, \varepsilon) := T_{g(t, \varepsilon)} L_{g(t, \varepsilon)^{-1}} \frac{\partial g(t, \varepsilon)}{\partial \varepsilon}. \quad (6.17)$$

Then

$$\frac{\partial \xi}{\partial \varepsilon} - \frac{\partial \eta}{\partial t} = [\xi, \eta]. \quad (6.18)$$

Conversely, if $U \subset \mathbb{R}^2$ is simply connected and $\xi, \eta : U \rightarrow \mathfrak{g}$ are smooth functions satisfying (6.18), then there exists a smooth function $g : U \rightarrow G$ such that (6.17) holds.

Let us show that (iii) $\iff$ (iv). We have

$$\begin{aligned} \delta \int_a^b l(\xi(t)) dt &= \int_a^b \left\langle \frac{\delta l}{\delta \xi}, \delta \xi \right\rangle dt = \int_a^b \left\langle \frac{\delta l}{\delta \xi}, \dot{\eta} + \text{ad}_\xi \eta \right\rangle dt \\ &= \int_a^b \left\langle \frac{\delta l}{\delta \xi}, \dot{\eta} \right\rangle dt + \int_a^b \left\langle \frac{\delta l}{\delta \xi}, \text{ad}_\xi \eta \right\rangle dt \\ &= \int_a^b \left\langle -\frac{d}{dt} \frac{\delta l}{\delta \xi} + \text{ad}_\xi^* \frac{\delta l}{\delta \xi}, \eta \right\rangle dt, \end{aligned}$$

where the last equality follows from integration by parts and because the curve $\eta(t)$ vanishes at the endpoints. Thus, $\delta \int_a^b l(\xi(t)) dt = 0$ if and only if the right hand side of the previous equality vanishes for any $\eta(t)$ that vanishes at the endpoints, which proves that this is equivalent to

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi}.$$

which are the Euler-Poincaré equations. ■

In case of right invariant Lagrangians on TG the same theorem holds with the changes that the Euler-Poincaré equations are

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = -\text{ad}_\xi^* \frac{\delta l}{\delta \xi}$$

and the variations are taken of the form $\delta\xi = \dot{\eta} - [\xi, \eta]$.

As was the case for Lie-Poisson dynamics, there is a reconstruction procedure in this case. The goal is to find the solution $v(t) \in T_{g(t)}G$ of the Euler-Lagrange equations with initial conditions $g(0) = g_0$ and $\dot{g}(0) = v_0$ knowing the solution of the Euler-Poincaré equations. To do this, first solve the initial value problem for the Euler-Poincaré equations:

$$\begin{cases} \frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} \\ \xi(0) = \xi_0 := g_0^{-1} v_0 \end{cases}$$

and then solve the “linear differential equation with time-dependent coefficients”

$$\begin{cases} \dot{g}(t) = g(t)\xi(t) \\ g(0) = g_0. \end{cases}$$

The Euler-Poincaré reduction theorem guarantees then that $v(t) = \dot{g}(t) = g(t) \cdot \xi(t)$ is a solution of the Euler-Lagrange equations with initial condition $v_0 = g_0 \xi_0$.

A similar statement holds, with obvious changes for right invariant Lagrangian systems on TG .

The relationship between Lie-Poisson and Euler-Poincaré reduction, similar to the link between the Hamiltonian and Lagrangian formulations discussed in Section 3.3, is the following. Define the Legendre transformation $\mathbb{F}l : \mathfrak{g} \rightarrow \mathfrak{g}^*$ by

$$\mathbb{F}l(\xi) = \frac{\delta l}{\delta \xi} = \mu,$$

and let $h(\mu) := \langle \mu, \xi \rangle - l(\xi)$. Assuming that $\mathbb{F}l$ is a diffeomorphism, we get

$$\frac{\delta h}{\delta \mu} = \xi + \left\langle \mu, \frac{\delta \xi}{\delta \mu} \right\rangle - \left\langle \frac{\delta l}{\delta \xi}, \frac{\delta \xi}{\delta \mu} \right\rangle = \xi.$$

So the Euler-Poincaré equations for l are equivalent to the Lie-Poisson equations for h :

$$\frac{d}{dt} \left(\frac{\delta l}{\delta \xi} \right) = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} \iff \dot{\mu} = \text{ad}_{\frac{\delta h}{\delta \mu}}^* \mu.$$

There is one more element to be discussed: the reduction of Hamilton’s phase space principle due to [Cendra et al.], called *Hamilton-Poincaré reduction*.

Theorem 6.8 *Let G be a Lie group, $H : T^*G \rightarrow \mathbb{R}$ a left-invariant Hamiltonian, and $h := H|_{\mathfrak{g}^*} : \mathfrak{g}^* \rightarrow \mathbb{R}$ be its restriction to $\mathfrak{g}^*$. For a curve*

$\alpha(t) \in T^*G$, let $g(t) = \pi(\alpha(t)) \in G$, where $\pi : T^*G \rightarrow G$ is the cotangent bundle projection, and define $\mu(t) = g(t)^{-1} \cdot \alpha(t) := T_e L_{g(t)} \alpha(t) \in \mathfrak{g}^*$, and $\xi(t) = g(t)^{-1} \cdot \dot{g}(t) := T_{g(t)} L_{g(t)^{-1}} \dot{g}(t) \in \mathfrak{g}$. The following are equivalent:

- (i) **Hamilton's Phase Space Principle.** The curve $\alpha(t) \in T^*G$ is a critical point of the action

$$\int_{t_0}^{t_1} (\langle \Theta, \dot{\alpha}(t) \rangle - H(\alpha(t))) dt,$$

where the variations $\delta\alpha$ satisfy $T\pi_G(\delta\alpha(t_i)) = 0$, for $i = 0, 1$.

- (ii) Hamilton's equations hold on T^*G .

- (iii) **The Hamilton-Poincaré Variational Principle.** The curve $(\mu(t), \xi(t)) \in \mathfrak{g}^* \times \mathfrak{g}$ is a critical point of the action

$$\int_{t_0}^{t_1} (\langle \mu(t), \xi(t) \rangle - h(\mu(t))) dt,$$

with variations $\delta\xi(t) = \dot{\eta}(t) + [\xi(t), \eta(t)] \in \mathfrak{g}$, where $\eta(t)$ is an arbitrary curve satisfying $\eta(t_i) = 0$, for $i = 0, 1$, and $\delta\mu(t) \in \mathfrak{g}^*$ is arbitrary.

- (iv) The Lie-Poisson equations hold:

$$\dot{\mu} = -\text{ad}_{\frac{\delta h}{\delta \mu}}^* \mu.$$

There is, of course, a similar statement for right invariant systems where one has to change the sign in front of the Lie-Poisson equation and in front of the bracket defining $\delta\xi$.

Proof The equivalence (i) $\iff$ (ii) is Hamilton's phase space variational principle which holds on any cotangent bundle (see Theorem 3.7). The equivalence (ii) $\iff$ (iv) is the Lie-Poisson reduction and reconstruction of dynamics (see Theorems 6.1 and 6.3). We now show that (iii) $\iff$ (iv). The variation

$$\begin{aligned} & \delta \int_{t_0}^{t_1} (\langle \mu(t), \xi(t) \rangle - h(\mu(t))) dt \\ &= \int_{t_0}^{t_1} \left(\langle \delta\mu, \xi \rangle + \langle \mu, \delta\xi \rangle - \left\langle \delta\mu, \frac{\delta h}{\delta \mu} \right\rangle \right) dt \\ &= \int_{t_0}^{t_1} \left(\left\langle \delta\mu, \xi - \frac{\delta h}{\delta \mu} \right\rangle + \langle \mu, \dot{\eta} + \text{ad}_{\xi} \eta \rangle \right) dt \\ &= \int_{t_0}^{t_1} \left(\left\langle \delta\mu, \xi - \frac{\delta h}{\delta \mu} \right\rangle + \langle -\dot{\mu} + \text{ad}_{\xi}^* \mu, \eta \rangle \right) dt \end{aligned}$$

vanishes for any functions $\delta\mu(t) \in \mathfrak{g}^*$ and $\eta(t) \in \mathfrak{g}$ (vanishing at the endpoints

t_0 and t_1) if and only if $\xi = \delta h / \delta \mu$ and $\dot{\mu} = -\text{ad}_\xi^* \mu$, that is, when the Lie Poisson equations in (iv) hold. ■

If the Lagrangian L is hyperregular then any of the statements in Theorem 6.8 are equivalent to any of the statements in Theorem 6.6. The link between the reduced Lagrangian l and the reduced Hamiltonian h is given by $h(\mu) = \langle \mu, \xi \rangle - l(\xi)$, where $\mu = \delta l / \delta \xi$, as was already remarked earlier.

Let us work out a few examples in detail.

(1) Free rigid body. As we saw in Example 1 of Section 6.3, the restriction of the Lagrangian of the free rigid body to $\mathbb{R}^3 \cong \mathfrak{so}(3)$ is $L(\Omega) = \Omega \cdot \mathbb{I}\Omega/2$ (see (1.36) with $\Pi = \mathbb{I}\Omega$) and hence, using the identity $\text{ad}_u^* \tilde{\Pi} = (\Pi \times u)^\sim$ and Theorem 6.6, we get the Euler-Poincaré equations

$$\mathbb{I}\dot{\Omega} = \mathbb{I}\Omega \times \Omega$$

and we recognize the Euler equations, this time formulated in terms of the body angular velocity Ω as opposed to the body angular momentum Π . The Legendre transformation is $\Pi = \mathbb{I}\Omega$. Looking back at the computations involving the variational principle carried out at the end of Section 1.2 and setting in those computations $\ell = 0$, one recognizes the variational principle on $\mathfrak{so}(3)$ for the free rigid body Lagrangian.

(2) Lagrangian systems on the affine algebra. Consider the affine Lie algebra introduced in example 2 of Section 6.3 and let $L : TG \rightarrow \mathbb{R}$ be a left invariant Lagrangian. The Euler-Poincaré equations for $l : \mathbb{R}^2 \rightarrow \mathbb{R}$ are

$$\frac{d}{dt} \frac{\partial l}{\partial u} = -v \frac{\partial l}{\partial v}, \quad \frac{d}{dt} \frac{\partial l}{\partial v} = u \frac{\partial l}{\partial u}.$$

(3) Incompressible homogeneous fluids, vorticity representation. In Section 6.3, Example 3, we have studied the motion of an incompressible homogeneous fluid in vorticity representation as a Lie-Poisson system. Now we shall derive the same equations (6.11) as Euler-Poincaré equations. We recall that $\text{ad}_v^* \omega = \mathcal{L}_v \omega$ (see (6.8)) and

$$l(v) = \frac{1}{2} \int_D \|v\|^2 \mu = \frac{1}{2} \langle \omega, v \rangle,$$

where $\text{dv}^\flat = \omega$ (see (6.10)). Therefore, $\delta l / \delta v = \omega$ and hence the Euler-Poincaré equations (for a right invariant system on the tangent bundle of a Lie group) are given by

$$\frac{\partial \omega}{\partial t} + \mathcal{L}_v \omega = 0.$$

(4) Incompressible homogeneous fluids, velocity representation. In Section 6.3, Example 4, we have presented in Lie-Poisson setting the equations of motion of an ideal incompressible homogeneous fluid in velocity representation. These are the Euler equations (6.14) which were not explicitly derived there because this will be done now. To do this, one could proceed by using the general expression for the Euler-Poincaré equations and formula (6.13) for the coadjoint action. Even though (6.13) was not proved in these notes, let us do it first this way. Since

$$l(v) = \frac{1}{2} \int_D \|v\|^2 \mu$$

is the kinetic energy of the fluid and we use the L^2 weak pairing to identify $\mathfrak{X}_{\text{div}}(D)$ with itself, we have $\delta l / \delta v = v$ and so the Euler-Poincaré equations are

$$\frac{\partial v}{\partial t} = -\text{ad}_v^* v = -\mathbb{P}(\nabla_v v + (v \cdot \nabla v)^\sharp).$$

Write, according to the Helmholtz decomposition,

$$\nabla_v v + (v \cdot \nabla v)^\sharp = \mathbb{P}(\nabla_v v + (v \cdot \nabla v)^\sharp) - \nabla q$$

for some smooth function $q : D \rightarrow \mathbb{R}$ and note that $(v \cdot \nabla v)^\sharp = \nabla(\|v\|^2/2)$. The Euler-Poincaré equations become hence

$$\frac{\partial v}{\partial t} + \nabla_v v = -\nabla \left(\frac{1}{2} \|v\|^2 + q \right)$$

which are the Euler equations for the pressure $q + \|v\|^2/2$. Therefore, *the Euler equations for an ideal incompressible homogeneous fluid are the spatial representation of the geodesic spray on $\text{Diff}_{\text{vol}}(D)$ for the right invariant metric whose value on $\mathfrak{X}_{\text{div}}(D)$ is the L^2 inner product. This proves that the solutions of the Euler equations are geodesics on $\text{Diff}_{\text{vol}}(D)$.*

Even though the computation above proves the claim that the Euler-Poincaré equations are the Euler equations for an ideal incompressible homogeneous fluid, it is unsatisfactory since it relies on the unproven formula (6.13). Due to its importance we shall derive it once more directly.

We begin by recalling that $\delta l / \delta v = v$ and hence for any $w \in \mathfrak{X}_{\text{div}}(D)$ we have

$$\left\langle -\text{ad}_v^* \frac{\delta l}{\delta v}, w \right\rangle = \left\langle \frac{\delta l}{\delta v}, -\text{ad}_v w \right\rangle = \langle v, [v, w] \rangle. \quad (6.19)$$

In order to isolate w in $\langle v, [v, w] \rangle$, let us compute the Lie derivative of $v^\flat(w)\mu$

along the vector field v in two different ways. We have

$$\begin{aligned}\mathcal{L}_v \left(v^\flat(w)\mu \right) &= \mathbf{i}_v \mathbf{d} \left(v^\flat(w)\mu \right) + \mathbf{d} \mathbf{i}_v \left(v^\flat(w)\mu \right) \\ &= \mathbf{d} \left(v^\flat(w) \mathbf{i}_v \mu \right).\end{aligned}\tag{6.20}$$

On the other hand, since $\mathcal{L}_v \mu = (\operatorname{div} v)\mu = 0$, we get

$$\begin{aligned}\mathcal{L}_v \left(v^\flat(w)\mu \right) &= \left(\mathcal{L}_v v^\flat \right) (w)\mu + v^\flat(\mathcal{L}_v w)\mu + v^\flat(w)\mathcal{L}_v \mu \\ &= \left(\mathcal{L}_v v^\flat \right) (w)\mu + v^\flat([v, w])\mu.\end{aligned}\tag{6.21}$$

Then, from equations (6.20) and (6.21), we get

$$\mathbf{d} \left(v^\flat(w) \mathbf{i}_v \mu \right) = \left(\mathcal{L}_v v^\flat \right) (w)\mu + v^\flat([v, w])\mu$$

and hence by integration

$$\int_D \mathbf{d} \left(v^\flat(w) \mathbf{i}_v \mu \right) = \int_D \left(\mathcal{L}_v v^\flat \right) (w)\mu + \int_D v^\flat([v, w])\mu.$$

The left hand side is zero by Stokes' theorem and (6.5) and so

$$\begin{aligned}\langle v, [v, w] \rangle &= \int_D g(v, [v, w])\mu = \int_D v^\flat([v, w])\mu \\ &= - \int_D \left(\mathcal{L}_v v^\flat \right) (w)\mu = - \int_D g \left(\left(\mathcal{L}_v v^\flat \right)^\sharp, w \right) \mu \\ &= - \int_D g \left(\mathbb{P} \left(\left(\mathcal{L}_v v^\flat \right)^\sharp \right), w \right) \mu = - \left\langle \mathbb{P} \left(\left(\mathcal{L}_v v^\flat \right)^\sharp \right), w \right\rangle\end{aligned}$$

which, using (6.19), shows that

$$\operatorname{ad}_v^* \frac{\delta l}{\delta v} = \mathbb{P} \left(\left(\mathcal{L}_v v^\flat \right)^\sharp \right)$$

by weak non-degeneracy of the L^2 pairing on divergence free vector fields tangent to the boundary. Thus the Euler-Poincaré equations are

$$\frac{\partial v}{\partial t} + \mathbb{P} \left(\left(\mathcal{L}_v v^\flat \right)^\sharp \right) = 0.$$

Since $(\mathcal{L}_v v^\flat)^\sharp = \nabla_v v + \frac{1}{2} \nabla \|v\|^2$, writing $\nabla_v v = \mathbb{P}(\nabla_v v) - \nabla p$, for some smooth function $p : D \rightarrow \mathbb{R}$, the above equation becomes

$$\begin{aligned}0 &= \frac{\partial v}{\partial t} + \mathbb{P} \left(\nabla_v v + \frac{1}{2} \nabla \|v\|^2 \right) = \frac{\partial v}{\partial t} + \mathbb{P} \left(\mathbb{P}(\nabla_v v) - \nabla p + \frac{1}{2} \nabla \|v\|^2 \right) \\ &= \frac{\partial v}{\partial t} + \mathbb{P}(\nabla_v v) = \frac{\partial v}{\partial t} + \nabla_v v + \nabla p\end{aligned}$$

which are the Euler equations (6.14).

(5) KdV equation and the Virasoro algebra. Here we show, following [OKh87], that the periodic KdV equation is the Euler-Poincaré equation on the Virasoro algebra $\mathfrak{v}$ corresponding to the geodesic flow of the L^2 right invariant metric on the Virasoro group.

The Lie algebra $\mathfrak{X}(S^1)$ of vector fields on the circle, identified with the periodic functions of period 1, with Lie bracket given by $[u, v] = uv' - u'v$, has a unique central extension by $\mathbb{R}$. This unique central extension is the **Virasoro algebra**

$$\mathfrak{v} := \{(u, a) \in \mathfrak{X}(S^1) \times \mathbb{R}\},$$

with Lie bracket

$$[(u, a), (v, b)] := \left(-uv' + u'v, \gamma \int_0^1 u'(x)v''(x) dx \right),$$

where the first argument is the (left) Lie bracket on $\mathfrak{X}(S^1)$, the second is the Gelfand-Fuchs cocycle, and $\gamma \in \mathbb{R}$ is a constant.

We identify $\mathfrak{v}^*$ with $\mathfrak{v}$ using the weak L^2 pairing

$$\langle (u, a), (v, b) \rangle = ab + \int_0^1 u(x)v(x) dx.$$

We also consider the right invariant weak Riemannian metric whose value on $\mathfrak{v}$ is the expression above. We are interested in the geodesic equations for this metric. To see who they are, we compute the Euler-Poincaré equations for the kinetic energy of this metric.

We begin by computing the coadjoint action of $\mathfrak{v}$ on $\mathfrak{v}^*$:

$$\begin{aligned} \langle \text{ad}_{(u,a)}^*(v, b), (w, c) \rangle &= \langle (v, b), [(u, a), (w, c)] \rangle \\ &= \left\langle (v, b), \left(-uw' + u'w, \gamma \int_0^1 u'(x)w''(x) dx \right) \right\rangle \\ &= b\gamma \int_0^1 u'w'' dx - \int_0^1 v u w' dx + \int_0^1 v u' w dx \\ &= \int_0^1 (b\gamma u''' + 2u'v + uv') w dx \\ &= \langle (b\gamma u''' + 2u'v + uv', 0), (w, c) \rangle, \end{aligned}$$

where from the third to the fourth equality we have used integration by parts, twice for the first term and once for the second term, as well as the null conditions on the boundary. So,

$$\text{ad}_{(u,a)}^*(v, b) = (b\gamma u''' + 2u'v + uv', 0) \quad (6.22)$$

and the Euler-Poincaré equations (for a right invariant system) determined by $l : \mathfrak{v} \rightarrow \mathbb{R}$ are

$$\begin{aligned} \frac{d}{dt} \left(\frac{\delta l}{\delta u}, \frac{\delta l}{\delta a} \right) &= -\text{ad}_{(u,a)}^* \left(\frac{\delta l}{\delta u}, \frac{\delta l}{\delta a} \right) \\ &= - \left(\gamma \frac{\delta l}{\delta a} u''' + 2u' \frac{\delta l}{\delta u} + u \left(\frac{\delta l}{\delta u} \right)', 0 \right), \end{aligned}$$

that is,

$$\frac{d}{dt} \frac{\delta l}{\delta a} = 0, \quad \frac{d}{dt} \frac{\delta l}{\delta u} = - \left(\gamma \frac{\delta l}{\delta a} u''' + 2u' \frac{\delta l}{\delta u} + u \left(\frac{\delta l}{\delta u} \right)' \right). \quad (6.23)$$

If we are interested in the L^2 geodesic flow, then we take

$$l(u, a) = \frac{1}{2} \left(a^2 + \int_0^1 u^2(x) dx \right)$$

so that $\frac{\delta l}{\delta a} = a$ and $\frac{\delta l}{\delta u} = u$. Thus the corresponding Euler-Poincaré equations are:

$$\frac{da}{dt} = 0, \quad \frac{du}{dt} = -\gamma a u''' - 3u'u.$$

Hence a is constant and we get

$$u_t + 3u_x u + \gamma a u_{xxx} = 0,$$

which is one of the forms of the KdV equation (for example by choosing $a = \gamma = 1$). To get the expression (4.10), that is,

$$u_t + 6u_x u + u_{xxx} = 0$$

one needs to rescale time ($\tau(t) = t/2$) and make an appropriate choice of the constants ($a = 1/(2\gamma)$). This shows that *the solutions of the KdV equation are geodesics of the L^2 right invariant metric on the Virasoro group.*

(6) Camassa-Holm equation and the Virasoro algebra. Let us compute the Euler-Poincaré equations for the kinetic energy of the H^1 metric

$$\langle (u, a), (v, b) \rangle := ab + \int_0^1 (u(x)v(x) + u'(x)v'(x)) dx$$

on the Virasoro algebra. As before, we identify $\mathfrak{v}^*$ with $\mathfrak{v}$ using the L^2 inner product. Thus $\delta l / \delta a = a$ and $\delta l / \delta u = u - u''$. The Euler-Poincaré equations

(6.23) are given by $a = \text{constant}$ and

$$\begin{aligned}\frac{d}{dt}(u - u'') &= -\left(\gamma a u''' + 2u'(u - u'') + u(u - u'')'\right) \\ &= -3uu' + 2u'u'' + uu''' - \gamma a u'''.\end{aligned}$$

This is the Camassa-Holm equation:

$$u_t - u_{txx} = -3uu_x + 2u_x u_{xx} + uu_{xxx} - \gamma a u_{xxx}.$$

Shifting $u \mapsto u + \gamma a$ brings the Camassa-Holm equation into yet another form, often used in the literature

$$u_t - u_{txx} = -3uu_x + 2u_x u_{xx} + uu_{xxx} - 3\gamma a u_x.$$

Thus, *the solutions of the Camassa-Holm equation are geodesics of the H^1 right invariant metric on the Virasoro group*, a result due to [Mis02].

An identical computation shows that *the solutions of Hunter-Saxon equation*

$$u_{txx} = -2u_x u_{xx} - uu_{xxx}$$

are geodesics of the right invariant degenerate metric on the Virasoro group whose value on $\mathfrak{v}$ is

$$\langle (u, a), (v, b) \rangle = ab + \int_0^1 u'(x)v'(x)dx,$$

a result due to [KhMis03]. Eliminating the degeneracy of the metric means looking at the Hunter-Saxon equation on the homogeneous space which is the quotient of the Virasoro group by rotations of the circle.

To see this, note first that if l is the kinetic energy of this degenerate metric then $\delta l / \delta u = -u''$, so again by (6.23) we get

$$-u_{txx} = -2u_x u_{xx} - uu_{xxx} - \gamma a u_{xxx}.$$

Reversing time and shifting $u \mapsto u - \gamma a$ yields the Hunter-Saxon equation.

7 Symplectic Reduction

In previous lectures we focussed our attention to the particular cases of reduction of Hamiltonian and Lagrangian systems where the phase space is either the cotangent space or the tangent space of a given Lie group. In this lecture we shall present the general case of symplectic reduction as formulated by [MaWei74]. All the hypotheses will require regularity assumptions. The singular case is considerably more involved and its complete treatment can be found in [OR04]. We shall present with detailed proofs the Marsden-Weinstein

theorem, nowadays often called “point reduction”. This will be done at geometric and dynamic level. Then we shall present three important examples of reduced manifolds with all computations done in detail. The orbit reduction method and the so-called “shifting trick” will then be presented without proofs. This chapter ends with a discussion of the semidirect product reduction theorem and its application to the motion of the heavy top presented in the Introduction.

7.1 Point Reduction

In this section we shall review the symplectic point reduction theorem and give its formulation and classical proof due to [MaWei74]. This procedure is of paramount importance in symplectic geometry and geometric mechanics. It underlies all the other reduction methods that one can find now in the literature as well as all the various generalizations that have proved their usefulness in areas as varied as algebraic geometry and topology, differential and symplectic topology, classical, continuum, and quantum mechanics, field theory, dynamical systems, bifurcation theory, and control theory.

The setup of the problem is the following. Let (P, Ω) be a symplectic manifold on which a Lie group G with Lie algebra $\mathfrak{g}$ acts in a Hamiltonian fashion with associated equivariant momentum map $\mathbf{J} : P \rightarrow \mathfrak{g}^*$. If $\mu \in \mathbf{J}(P) \subset \mathfrak{g}^*$ denote by $G_\mu := \{g \in G \mid \text{Ad}_g^* \mu = \mu\}$ the coadjoint isotropy subgroup of μ . The Marsden-Weinstein reduction theorem is the following.

Theorem 7.1 (Symplectic point reduction) *Assume that μ is a regular value of $\mathbf{J}$ and that the coadjoint isotropy subgroup G_μ acts freely and properly on $\mathbf{J}^{-1}(\mu)$. Then the quotient manifold $P_\mu := \mathbf{J}^{-1}(\mu)/G_\mu$ has a unique symplectic form Ω_μ characterized by the identity $\iota_\mu^* \Omega = \pi_\mu^* \Omega_\mu$, where $\iota_\mu : \mathbf{J}^{-1}(\mu) \hookrightarrow P$ is the inclusion and $\pi_\mu : \mathbf{J}^{-1}(\mu) \rightarrow P_\mu$ the projection. The symplectic manifold (P_μ, Ω_μ) is called the **symplectic point reduced space** at μ .*

In what follows we shall need the following notations. If (V, Ω) is a finite dimensional symplectic vector space and $W \subset V$ a subspace, define its **symplectic orthogonal** by $W^\Omega = \{v \in V \mid \Omega(v, w) = 0 \text{ for all } w \in W\}$. An elementary linear algebra argument, using the identity $\dim W + \dim W^\Omega = \dim V$, shows that $(W^\Omega)^\Omega = W$. If $z \in P$, denote by $G \cdot z$ and $G_\mu \cdot z$ the G - and G_μ -orbits through z respectively. It is important to note that the set $\mathbf{J}^{-1}(\mu)$ is G -invariant if and only if $G_\mu = G$. In general, $\mathbf{J}^{-1}(\mu)$ is only G_μ -invariant. The key ingredient in the proof of the reduction theorem is the following result

that we shall state in the setting of Poisson manifolds because of its usefulness in that general situation (not covered in these lectures).

Lemma 7.2 (Reduction Lemma) *Let P be a Poisson manifold and let $\mathbf{J} : P \rightarrow \mathfrak{g}^*$ an equivariant momentum map of the canonical G -action on P . Let $G \cdot \mu$ denote the coadjoint orbit through a regular value $\mu \in \mathfrak{g}^*$ of $\mathbf{J}$. Then*

- (i) $\mathbf{J}^{-1}(G \cdot \mu) = G \cdot \mathbf{J}^{-1}(\mu) := \{g \cdot z \mid g \in G \text{ and } \mathbf{J}(z) = \mu\};$
- (ii) $G_\mu \cdot z = (G \cdot z) \cap \mathbf{J}^{-1}(\mu);$
- (iii) $\mathbf{J}^{-1}(\mu)$ and $G \cdot z$ **intersect cleanly**, i.e.,

$$T_z(G_\mu \cdot z) = T_z(G \cdot z) \cap T_z(\mathbf{J}^{-1}(\mu));$$

- (iv) if (P, Ω) is symplectic, then $T_z(\mathbf{J}^{-1}(\mu)) = (T_z(G \cdot z))^\Omega.$

Proof (i) Since $\mathbf{J}^{-1}(G \cdot \mu)$ is a G -invariant set by equivariance of $\mathbf{J}$ and $\mathbf{J}^{-1}(\mu) \subset \mathbf{J}^{-1}(G \cdot \mu)$, it follows that $G \cdot \mathbf{J}^{-1}(\mu) \subset \mathbf{J}^{-1}(G \cdot \mu)$. Conversely, $z \in \mathbf{J}^{-1}(G \cdot \mu)$ if and only if $\mathbf{J}(z) = \text{Ad}_{g^{-1}}^* \mu$ for some $g \in G$, which is equivalent to $\mu = \text{Ad}_g^* \mathbf{J}(z) = \mathbf{J}(g \cdot z)$, i.e., $g \cdot z \in \mathbf{J}^{-1}(\mu)$ and hence $z = g \cdot (g^{-1} \cdot z) \in G \cdot \mathbf{J}^{-1}(\mu)$.

(ii) $g \cdot z \in \mathbf{J}^{-1}(\mu) \Leftrightarrow \mu = \mathbf{J}(g \cdot z) = \text{Ad}_{g^{-1}}^* \mathbf{J}(z) = \text{Ad}_{g^{-1}}^* \mu \Leftrightarrow g \in G_\mu.$

(iii) First suppose that $v_z \in T_z(G \cdot z) \cap T_z(\mathbf{J}^{-1}(\mu))$. Then $v_z = \xi_P(z)$ for some $\xi \in \mathfrak{g}$ and $T_z \mathbf{J}(v_z) = 0$ which, by infinitesimal equivariance (written in the form $T_z \mathbf{J}(\xi_P(z)) = -\text{ad}_\xi^* \mathbf{J}(z)$) gives $\text{ad}_\xi^* \mu = 0$; i.e., $\xi \in \mathfrak{g}_\mu$. If $v_z = \xi_P(z)$ for $\xi \in \mathfrak{g}_\mu$ then $v_z \in T_z(G_\mu \cdot z)$. The reverse inclusion is immediate since, by (ii), $G_\mu \cdot z$ is included in both $G \cdot z$ and $\mathbf{J}^{-1}(\mu)$.

(iv) The condition $v_z \in (T_z(G \cdot z))^\Omega$ means that $\Omega(z)(\xi_P(z), v_z) = 0$ for all $\xi \in \mathfrak{g}$. This is equivalent to $\langle T_z \mathbf{J}(v_z), \xi \rangle = \mathbf{d}J^\xi(z)(v_z) = 0$ for all $\xi \in \mathfrak{g}$ by definition of the momentum map. Thus, $v_z \in (T_z(G \cdot z))^\Omega$ if and only if $v_z \in \ker T_z \mathbf{J} = T_z(\mathbf{J}^{-1}(\mu))$. ■

We are now ready to prove the Symplectic Point Reduction Theorem.

Proof Since π_μ is a surjective submersion, if Ω_μ exists, it is uniquely determined by the condition $\pi_\mu^* \Omega_\mu = \iota_\mu^* \Omega$. This relation also defines Ω_μ in the following way. For $v \in T_z \mathbf{J}^{-1}(\mu)$, let $[v] = T_z \pi_\mu(v) \in T_{[z]} P_\mu$, where $[z] = \pi_\mu(z)$. Then $\pi_\mu^* \Omega_\mu = \iota_\mu^* \Omega$ is equivalent to

$$\Omega_\mu([z])([v], [w]) = \Omega(z)(v, w)$$

for all $v, w \in T_z \mathbf{J}^{-1}(\mu)$. To see that this relation defines Ω_μ , that is, it is independent of the choices made to define it, let $y = \Phi_g(z)$, $v' = T_z \Phi_g(v)$,

and $w' = T_z \Phi_g(w)$, where $g \in G_\mu$. If, in addition $[v''] = [v']$ and $[w''] = [w']$, then $v'' - v', w'' - w' \in \ker T_{g \cdot z} \pi_\mu = T_{g \cdot z}(G_\mu \cdot z)$ and thus

$$\begin{aligned} \Omega(y)(v'', w'') &= \Omega(y)((v'' - v') + v', (w'' - w') + w') \\ &= \Omega(y)(v'' - v', w'' - w') + \Omega(y)(v'' - v', w') \\ &\quad + \Omega(y)(v', w'' - w') + \Omega(y)(v', w'). \end{aligned}$$

The second and third terms vanish by Lemma 7.2 (iv). The first term vanishes by Lemma 7.2 (iii) and (iv). Thus we have

$$\Omega(y)(v'', w'') = \Omega(y)(v', w')$$

and we conclude

$$\begin{aligned} \Omega(y)(v'', w'') &= \Omega(y)(v', w') = \Omega(\Phi_g(z))(T_z \Phi_g(v), T_z \Phi_g(w)) \\ &= (\Phi_g^* \Omega)(z)(v, w) = \Omega(z)(v, w) \end{aligned}$$

since the action is symplectic. This proves that $\Omega_\mu([z])([v], [w])$ is well defined and satisfies the relation in the statement of the theorem. It is smooth since $\pi_\mu^* \Omega_\mu = \iota_\mu^* \Omega$ is smooth. Thus we have a well defined smooth two-form Ω_μ on P_μ .

Since $d\Omega = 0$, we get

$$\pi_\mu^* d\Omega_\mu = d\pi_\mu^* \Omega_\mu = d\iota_\mu^* \Omega = \iota_\mu^* d\Omega = 0.$$

Since π_μ is a surjective submersion, the pull-back map π_μ^* on forms is injective, so we can conclude that $d\Omega_\mu = 0$.

Finally, we prove that Ω_μ is non-degenerate. Suppose that $\Omega_\mu([z])([v], [w]) = 0$ for all $w \in T_z(\mathbf{J}^{-1}(\mu))$. This means that $\Omega(z)(v, w) = 0$ for all $w \in T_z(\mathbf{J}^{-1}(\mu))$, which is equivalent to $v \in (T_z(\mathbf{J}^{-1}(\mu)))^\Omega = T_z(G \cdot z)$ by Lemma 7.2 (iv). Hence $v \in T_z(\mathbf{J}^{-1}(\mu)) \cap T_z(G \cdot z) = T_z(G_\mu \cdot z)$ by Lemma 7.2 (iii) so that $[v] = 0$, thus proving the weak non-degeneracy of Ω_μ . ■

There are several important comments related to the Symplectic Point Reduction Theorem which will be addressed below.

(1) Define the *symmetry algebra* at $z \in P$ by $\mathfrak{g}_z := \{\xi \in \mathfrak{g} \mid \xi_P(z) = 0\}$. An element $\mu \in \mathfrak{g}^*$ is a regular value of $\mathbf{J}$ if and only if $\mathfrak{g}_z = 0$ for all $z \in \mathbf{J}^{-1}(\mu)$.

To prove this, recall that z is a regular point if and only if $T_z \mathbf{J}$ is surjective which is equivalent to $\{0\} = \{\xi \in \mathfrak{g} \mid \langle \xi, T_z \mathbf{J}(v) \rangle = 0, \text{ for all } v \in T_z P\}$. Since $\langle \xi, T_z \mathbf{J}(v) \rangle = \Omega(z)(\xi_P(z), v)$ by the definition of the momentum map, it thus follows that z is a regular point of $\mathbf{J}$ if and only if $\{0\} = \{\xi \in \mathfrak{g} \mid \Omega(z)(\xi_P(z), v) = 0 \text{ for all } v \in T_z P\}$. As $\Omega(z)$ is nondegenerate, this is in turn equivalent to $\mathfrak{g}_z = \{0\}$.

(2) The previous statement affirms that only points with trivial symmetry algebra are regular points of $\mathbf{J}$. This is important in concrete examples because it isolates the singular points easily. Another way to look at this statement is to interpret it as saying that points with symmetry are bifurcation points of $\mathbf{J}$. This simple observation turns out to have many important consequences, for example in the convexity theorems for momentum maps. Another consequence of statement (1) is that if $\mu \in \mathbf{J}(P) \subset \mathfrak{g}^*$ is a regular value of $\mathbf{J}$ then the action is **locally free**, which means, by definition, that the symmetry algebras of all the points in $\mathbf{J}^{-1}(\mu)$ vanish. In this case the reduction construction can be carried out locally.

(3) Even if $\Omega = -d\Theta$ and the action of G leaves Θ invariant, Ω_μ need not be exact. We shall prove in the next section that coadjoint orbits with their orbit symplectic form are reduced spaces. This immediately gives an example of a situation where the original symplectic form is exact but the reduced one is not. The sphere is the reduced space of $T^*SO(3)$. The canonical symplectic form on $T^*SO(3)$ is clearly exact whereas the area form on S^2 is not.

(4) If one looks at the proof of the theorem carefully, one notices that the hypothesis that μ is a regular value of $\mathbf{J}$ was not really used. What was necessary is that μ is a **clean value** of $\mathbf{J}$ which means, by definition, that $\mathbf{J}^{-1}(\mu)$ is a manifold and $T_z(\mathbf{J}^{-1}(\mu)) = \ker T_z\mathbf{J}$.

(5) The freeness and properness of the G_μ action on $\mathbf{J}^{-1}(\mu)$ are used only to guarantee that P_μ is a manifold. So these hypotheses can be replaced by the requirement that P_μ is a manifold and that $\pi_\mu : \mathbf{J}^{-1}(\mu) \rightarrow P_\mu$ a submersion.

(6) A point in $\mathfrak{g}^*$ is said to be **generic** if its coadjoint orbit is of maximal dimension. Duflo and Vergne have shown that the set of generic points is Zariski open in $\mathfrak{g}^*$ and that the coadjoint isotropy algebra of a generic point is necessarily Abelian. Regarding the momentum map, one should be warned that if μ is a regular value of $\mathbf{J}$, it need not be a generic point in $\mathfrak{g}^*$. As we shall see in the next section, the cotangent lift of the left (or right) translation of a Lie group G has all its values regular. However, among those, there are points that are not generic, such as the origin, in $\mathfrak{g}^*$.

(7) The connected components of the point reduced spaces P_μ can be viewed in a natural way as symplectic leaves of the Poisson manifold $(P/G, \{\cdot, \cdot\}_{P/G})$, provided that G acts freely and properly on P . Indeed, the smooth map $\kappa_\mu : P_\mu \rightarrow P/G$ naturally defined by the commutative diagram

$$\begin{array}{ccc}
\mathbf{J}^{-1}(\mu) & \xrightarrow{\iota_\mu} & P \\
\pi_\mu \downarrow & & \downarrow \pi \\
P_\mu & \xrightarrow{\kappa_\mu} & P/G
\end{array}$$

is a Poisson injective immersion. Moreover, the κ_μ -images in P/G of the connected components of the symplectic manifolds (P_μ, Ω_μ) are its symplectic leaves (see e.g. [OR04]). We will return to this in §7.4, where we will summarize the orbit reduction process and link it with these observations.

Note that, in general, κ_μ is only an injective immersion. So the topology of the image of κ_μ , homeomorphic to the topology of P_μ , is stronger than the subspace topology induced by the ambient space P/G . For example, we can have a subset of $\kappa_\mu(P_\mu)$ which is compact in the induced topology from P/G and not compact in the intrinsic topology of $\kappa_\mu(P_\mu)$ (relative to which it is homeomorphic to P_μ endowed with the quotient topology).

(8) We describe how reduction can be carried out for a non-equivariant momentum map $\mathbf{J} : P \rightarrow \mathfrak{g}^*$. If P is connected, the expression $\mathbf{J}(g \cdot z) - \text{Ad}_{g^{-1}}^* \mathbf{J}(z)$ turns out to be independent of $z \in P$. Setting $\sigma(g) := \mathbf{J}(g \cdot z) - \text{Ad}_{g^{-1}}^* \mathbf{J}(z)$ one obtains a group one-cocycle with values in $\mathfrak{g}^*$, that is, σ satisfies the cocycle identity $\sigma(gh) = \sigma(g) + \text{Ad}_{g^{-1}}^* \sigma(h)$ (see, e.g. [MaRa94]). If this cocycle is a coboundary, that is, there is some $\lambda \in \mathfrak{g}^*$ such that $\sigma(g) = \lambda - \text{Ad}_{g^{-1}}^* \lambda$, then the momentum map can be modified by the addition of $-\lambda$ to become equivariant. If σ is not a coboundary then there is no way one can modify $\mathbf{J}$ to make it equivariant. This $\mathfrak{g}^*$ -valued group one-cocycle $\sigma : G \rightarrow \mathfrak{g}^*$ is called the **nonequivariance group one-cocycle** defined by $\mathbf{J}$.

To carry out reduction, modify the coadjoint action of G on $\mathfrak{g}^*$ in the following way: $g \cdot \mu := \text{Ad}_{g^{-1}}^* \mu + \sigma(g)$. Relative to this affine action the momentum map $\mathbf{J}$ is equivariant and the reduction procedure works by dividing the level set $\mathbf{J}^{-1}(\mu)$ by the μ -isotropy subgroup for this affine action.

(9) If the regularity assumptions in the Symplectic Point Reduction Theorem do not hold, then P_μ is a stratified space, all of whose strata are symplectic manifolds. This result is considerably more difficult to prove and we refer to [OR04] and references therein for an exposition of this theory. One could even further relax the requirements, namely, do not even assume that there

is a momentum map for a given canonical Lie group action on (P, Ω) . In this case there is a generalization of the momentum map due to Condevaux, Dazord, and Molino, the so-called cylinder valued momentum map, for which the reduction procedure can be implemented. We refer again to [OR04] and references therein for a treatment of this subject.

7.2 Reduction and Reconstruction of Dynamics

The geometric theorem presented in the previous section has a dynamic counterpart that will be discussed now. We keep the same conventions and notations as in §7.1.

Theorem 7.3 (Point reduction of dynamics) *Let $\Phi : G \times P \rightarrow P$ be a free proper canonical action of the Lie group G on the connected symplectic manifold (P, Ω) . Assume that this action has an associated momentum map $\mathbf{J} : P \rightarrow \mathfrak{g}^*$, with nonequivariance one-cocycle $\sigma : G \rightarrow \mathfrak{g}^*$. Let $\mu \in \mathfrak{g}^*$ be a value of $\mathbf{J}$ and denote by G_μ the isotropy subgroup of μ under the affine action of G on $\mathfrak{g}^*$.*

- (i) *Let $H : P \rightarrow \mathbb{R}$ be a smooth G -invariant function. The flow F_t of the Hamiltonian vector field X_H leaves the connected components of $\mathbf{J}^{-1}(\mu)$ invariant and commutes with the G -action, so it induces a flow F_t^μ on the reduced space $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$ defined by*

$$\pi_\mu \circ F_t \circ \iota_\mu = F_t^\mu \circ \pi_\mu.$$

*The vector field generated by the flow F_t^μ on (P_μ, Ω_μ) is Hamiltonian with associated **reduced Hamiltonian function** $H_\mu : P_\mu \rightarrow \mathbb{R}$ defined by*

$$H_\mu \circ \pi_\mu = H \circ \iota_\mu.$$

The vector fields X_H and X_{H_μ} are π_μ -related.

- (ii) *Let $F : P \rightarrow \mathbb{R}$ be another smooth G -invariant function. Then $\{F, H\}$ is also G -invariant and $\{F, H\}_\mu = \{F_\mu, H_\mu\}_{P_\mu}$, where $\{\cdot, \cdot\}_{P_\mu}$ denotes the Poisson bracket associated to the reduced symplectic form Ω_μ on P_μ .*

Proof (i) By Noether's Theorem 5.4, the flow F_t leaves the connected components of $\mathbf{J}^{-1}(\mu)$ invariant. Since H is G -invariant and the G -action is canonical, it follows by Proposition 2.7 that F_t commutes with the G -action. Thus

F_t induces a flow F_t^μ on P_μ that makes the following diagram commutative:

$$\begin{array}{ccc} \mathbf{J}^{-1}(\mu) & \xrightarrow{F_t \circ \iota_\mu} & \mathbf{J}^{-1}(\mu) \\ \pi_\mu \downarrow & & \downarrow \pi_\mu \\ P_\mu & \xrightarrow{F_t^\mu} & P_\mu. \end{array}$$

The G -invariance of H implies the existence of a smooth function $H_\mu : P_\mu \rightarrow \mathbb{R}$ uniquely determined by the identity $H_\mu \circ \pi_\mu = H \circ \iota_\mu$. Let $Y \in \mathfrak{X}(P_\mu)$ be the vector field on P_μ whose flow is F_t^μ . By construction, Y is π_μ -related to X_H . Indeed, differentiating the relation given by the diagram above relative to t at $t = 0$, we obtain

$$T\pi_\mu \circ X_H \circ \iota_\mu = Y \circ \pi_\mu.$$

Let us check that $Y = X_{H_\mu}$. For $z \in \mathbf{J}^{-1}(\mu)$ and $v \in T_z \mathbf{J}^{-1}(\mu)$ we have

$$\begin{aligned} \Omega_\mu(\pi_\mu(z)) (Y(\pi_\mu(z)), T_z \pi_\mu(v)) &= \Omega_\mu(\pi_\mu(z)) (T_z \pi_\mu(X_H(z)), T_z \pi_\mu(v)) \\ &= \Omega(z)(X_H(z), v) = \mathbf{d}H(z)(v) = \mathbf{d}(H_\mu \circ \pi_\mu)(z)(v) \\ &= \mathbf{d}H_\mu(\pi_\mu(z)) (T_z \pi_\mu(v)) = \Omega_\mu(\pi_\mu(z)) (X_{H_\mu}(\pi_\mu(z)), T_z \pi_\mu(v)), \end{aligned}$$

which, by nondegeneracy of Ω_μ , shows that $Y = X_{H_\mu}$.

(ii) The G -invariance of $\{F, H\}$ is a straightforward corollary of Proposition 2.8. Recall that the function $\{F, H\}_\mu$ is uniquely characterized by the identity $\{F, H\}_\mu \circ \pi_\mu = \{F, H\} \circ \iota_\mu$. By the definition of the Poisson bracket on (P_μ, Ω_μ) , π_μ -relatedness of the relevant Hamiltonian vector fields, and the identity $\iota_\mu^* \Omega = \pi_\mu^* \Omega_\mu$, we have for any $z \in \mathbf{J}^{-1}(\mu)$ denoting $[z]_\mu := \pi_\mu(z)$,

$$\begin{aligned} \{F_\mu, H_\mu\}_{P_\mu}([z]_\mu) &= \Omega_\mu([z]_\mu) (X_{F_\mu}([z]_\mu), X_{H_\mu}([z]_\mu)) \\ &= \Omega_\mu([z]_\mu) (T_z \pi_\mu(X_F(z)), T_z \pi_\mu(X_H(z))) \\ &= (\pi_\mu^* \Omega_\mu)(z) (X_F(z), X_H(z)) = (\iota_\mu^* \Omega)(z) (X_F(z), X_H(z)) \\ &= \Omega(z) (X_F(z), X_H(z)) = \{F, H\}(z), \end{aligned}$$

that is, the function $\{F_\mu, H_\mu\}_{P_\mu}$ also satisfies the relation $\{F_\mu, H_\mu\}_{P_\mu} \circ \pi_\mu = \{F, H\} \circ \iota_\mu$, which proves the desired equality $\{F_\mu, H_\mu\}_{P_\mu} = \{F, H\}_\mu$. ■

This theorem shows how dynamics on P descends to dynamics on all reduced manifolds P_μ for any $\mu \in \mathfrak{g}^*$, if the group action is free and proper. For the singular case see [OR04].

Let us now pose the converse question. Assume that an integral curve $c_\mu(t)$ of the reduced Hamiltonian system X_{H_μ} on (P_μ, Ω_μ) is known. Let $z_0 \in \mathbf{J}^{-1}(\mu)$ be given. Can one determine from this data the integral curve of the

Hamiltonian system X_H with initial condition z_0 ? The answer to this question is affirmative as we shall see below.

The general method of **reconstruction of dynamics** is the following ([AbMa78] §4.3, [MaMoRa90], [MaRa03]). Pick a smooth curve $d(t)$ in $\mathbf{J}^{-1}(\mu)$ such that $d(0) = z_0$ and $\pi_\mu(d(t)) = c_\mu(t)$. We shall give later concrete choices for such curves in terms of connections. Then, if $c(t)$ denotes the integral curve of X_H with $c(0) = z_0$, we can write $c(t) = g(t) \cdot d(t)$ for some smooth curve $g(t)$ in G_μ . We shall determine now $g(t)$ and therefore $c(t)$. Below, $\Phi : G \times P \rightarrow P$ denotes the left action of G on P and $\Phi_g : P \rightarrow P$ is the diffeomorphism of P given by the group element $g \in G$. We have

$$\begin{aligned} X_H(c(t)) &= \dot{c}(t) = T_{d(t)}\Phi_{g(t)}\dot{d}(t) + T_{d(t)}\Phi_{g(t)}\left(T_{g(t)}L_{g(t)^{-1}}\dot{g}(t)\right)_P(d(t)) \\ &= T_{d(t)}\Phi_{g(t)}\left(\dot{d}(t) + (T_{g(t)}L_{g(t)^{-1}}\dot{g}(t))_P(d(t))\right) \end{aligned}$$

which implies

$$\begin{aligned} \dot{d}(t) + (T_{g(t)}L_{g(t)^{-1}}\dot{g}(t))_P(d(t)) &= T_{g(t) \cdot d(t)}\Phi_{g(t)^{-1}}X_H(c(t)) \\ &= T_{g(t) \cdot d(t)}\Phi_{g(t)^{-1}}X_H(g(t) \cdot d(t)) \\ &= \left(\Phi_{g(t)}^*X_H\right)(d(t)) = X_H(d(t)) \end{aligned}$$

since, by hypothesis, $H = \Phi_g^*H$ and thus, by Proposition 2.7, $X_H = X_{\Phi_g^*H} = \Phi_g^*X_H$ for any $g \in G$. This equation is solved in two steps as follows:

- **Step 1:** Find a smooth curve $\xi(t)$ in $\mathfrak{g}_\mu$ such that

$$\xi(t)_P(d(t)) = X_H(d(t)) - \dot{d}(t). \quad (7.1)$$

- **Step 2:** With $\xi(t) \in \mathfrak{g}_\mu$ determined above, solve the nonautonomous differential equation on G_μ

$$\dot{g}(t) = T_e L_{g(t)}\xi(t), \quad \text{with} \quad g(0) = e. \quad (7.2)$$

Here are some useful remarks regarding the solution of each step.

(1) The first step is of algebraic nature. For example, if G is a matrix Lie group, (7.1) is just a matrix equation. If one is willing to work with more geometric structure, this equation can be solved explicitly. Typically, one endows the left principal G_μ -bundle $\pi_\mu : \mathbf{J}^{-1}(\mu) \rightarrow P_\mu$ with a connection. Recall that a (left) connection on this bundle is given by a $\mathfrak{g}_\mu$ -valued one-form $A \in \Omega^1(\mathbf{J}^{-1}(\mu); \mathfrak{g}_\mu)$ satisfying for all $z \in \mathbf{J}^{-1}(\mu)$ the relations

$$A(z) (\xi_{\mathbf{J}^{-1}(\mu)}(z)) = \xi, \quad \text{for all} \quad \xi \in \mathfrak{g}_\mu$$

and

$$A(g \cdot z)(T_z \Phi_g(v_z)) = \text{Ad}_g(A(z)(v_z)), \quad \text{for all } g \in G_\mu, v_z \in T_z \mathbf{J}^{-1}(\mu).$$

Let G_μ act on P by restriction so that $\xi_P = \xi_{\mathbf{J}^{-1}(\mu)}$ for any $\xi \in \mathfrak{g}_\mu$. Choose in Step 1 of the reconstruction method the curve $d(t)$ to be the horizontal lift of $c_\mu(t)$ through z_0 , that is, $d(t)$ is uniquely characterized by the conditions $A(d(t))(\dot{d}(t)) = 0$, $\pi_\mu(d(t)) = c_\mu(t)$, for all t , and $d(0) = z_0$. Then the solution of (7.1) is given by

$$\xi(t) = A(d(t))(X_H(d(t))).$$

(2) The second step is the main difficulty in finding a complete answer to the reconstruction problem; equation (7.2) cannot be solved explicitly, in general. For matrix groups this is a linear system with time dependent coefficients. However, if G is Abelian, this equation can be solved *by quadratures*. To see how this works, we need the formula of the derivative of the exponential map at any point in the Lie algebra.

If G is a Lie group and $\exp : \mathfrak{g} \rightarrow G$ the exponential map then

$$T_\xi \exp = T_e L_{\exp \xi} \circ \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} \text{ad}_\xi^n$$

for any $\xi \in \mathfrak{g}$. If G is Abelian, then $\text{ad}_\xi = 0$ and so $T_\xi \exp = T_e L_{\exp \xi}$, a formula that we shall use below.

So let us return to the second step in the reconstruction method for an Abelian group G . Since the connected component of the p -dimensional Lie group G is a cylinder $\mathbb{R}^k \times \mathbb{T}^{p-k}$, the exponential map $\exp(\xi_1, \dots, \xi_p) = (\xi_1, \dots, \xi_k, \xi_{k+1} \pmod{2\pi}, \dots, \xi_p \pmod{2\pi})$ is onto, so we can write $g(t) = \exp \eta(t)$ for some smooth curve $\eta(t) \in \mathfrak{g}$ satisfying $\eta(0) = 0$. Equation (7.2) gives then $\xi(t) = T_{g(t)} L_{g(t)^{-1}} \dot{g}(t) = \dot{\eta}(t)$ since $\dot{g}(t) = T_{\eta(t)} \exp \dot{\eta}(t) = T_e L_{\exp \eta(t)} \dot{\eta}(t)$ by the comments above. Therefore, in this case, the solution of (7.2) is given by

$$g(t) = \exp \left(\int_0^t \xi(s) ds \right).$$

This reconstruction method is crucial in the determination of various geometric phases in mechanical problems; see [MaMoRa90] for details.

7.3 Examples of Reduced Manifolds

The projective space. Consider $\mathbb{C}^{2n} = \mathbb{R}^{4n} = T^*\mathbb{R}^{2n}$ endowed with the canonical symplectic structure $d\mathbf{q} \wedge d\mathbf{p}$ for $(\mathbf{q}, \mathbf{p}) \in T^*\mathbb{R}^{2n}$. The flow of

the Hamiltonian vector field given by the harmonic oscillator Hamiltonian $H(\mathbf{q}, \mathbf{p}) := (\|\mathbf{q}\|^2 + \|\mathbf{p}\|^2)/2$ is 2π -periodic and hence defines the circle action $\theta \cdot (\mathbf{q}, \mathbf{p}) \mapsto (\mathbf{q} \cos \theta + \mathbf{p} \sin \theta, -\mathbf{q} \sin \theta + \mathbf{p} \cos \theta)$ on $\mathbb{R}^{2n}$ or $\theta \cdot (\mathbf{q} + i\mathbf{p}) := e^{-i\theta}(\mathbf{q} + i\mathbf{p})$ on $\mathbb{C}^n$. Since the circle is compact this action is necessarily proper and it is obvious that it is free away from the origin. The infinitesimal generator of this action is X_H , which shows that $H : T^*\mathbb{R}^n \rightarrow \mathbb{R}$ is an invariant momentum map for this circle action. All hypotheses of the Symplectic Point Reduction Theorem hold and therefore $H^{-1}(1/2)$ is a symplectic manifold. However, since $H^{-1}(1/2)$ is diffeomorphic to the unit sphere S^{2n-1} , we can immediately conclude that $H^{-1}(1/2)/S^1 \cong S^{2n-1}/S^1$. This suggests that this reduced manifold is in fact symplectically diffeomorphic to complex projective space $\mathbb{CP}^{n-1}$. We shall prove this below.

Recall that $\mathbb{CP}^{n-1}$ is the space of complex lines through the origin in $\mathbb{C}^n$. Let $\pi : \mathbb{C}^n \setminus \{0\} \rightarrow \mathbb{CP}^{n-1}$ be the tautological projection that sends the vector $\mathbf{z} \neq 0$ to the complex line it spans, denoted by $[\mathbf{z}]$ when thought of as an element of $\mathbb{CP}^{n-1}$. Consider the inclusion $\iota_{1/2} : H^{-1}(1/2) = S^{2n-1} \hookrightarrow \mathbb{C}^n \setminus \{0\}$ and note that it preserves the equivalence relations, that is, the span of $e^{-i\theta}\mathbf{z}$ equals the span of $\mathbf{z} \neq 0$. Therefore, $\iota_{1/2}$ induces a smooth map $\hat{\iota}_{1/2} : S^{2n-1}/S^1 \rightarrow \mathbb{CP}^{n-1}$, uniquely characterized by the relation $\pi \circ \iota_{1/2} = \hat{\iota}_{1/2} \circ \pi_{1/2}$, which is easily seen to be bijective. Since the inverse of this map is the quotient of the smooth map $\mathbf{z} \in \mathbb{C}^n \setminus \{0\} \mapsto S^1 \cdot (\mathbf{z}/\|\mathbf{z}\|) \in S^{2n-1}/S^1$ by the equivalence relation defining projective space, it follows that this inverse is also smooth and therefore $\hat{\iota}_{1/2} : S^{2n-1}/S^1 \rightarrow \mathbb{CP}^{n-1}$ is a diffeomorphism. In what follows it is convenient to define the map $\varphi : S^{2n-1} \rightarrow \mathbb{CP}^{n-1}$ by $\varphi := \pi \circ \iota_{1/2} = \hat{\iota}_{1/2} \circ \pi_{1/2}$. Let us record all of these maps in the following commutative diagram:

$$\begin{array}{ccc}
 S^{2n-1} & \xrightarrow{\iota_{1/2}} & \mathbb{C}^n \setminus \{0\} \\
 \pi_{1/2} \downarrow & \searrow \varphi & \downarrow \pi \\
 S^{2n-1}/S^1 & \xrightarrow{\hat{\iota}_{1/2}} & \mathbb{CP}^{n-1}
 \end{array}$$

Finally, we recall the symplectic form on projective space (see e.g. [MaRa94], §5.3). Let $[\mathbf{z}] \in \mathbb{CP}^{n-1}$ be such that $\|\mathbf{z}\| = 1$ and let $\mathbf{w}_1, \mathbf{w}_2 \in (\mathbb{C}\mathbf{z})^\perp$. Then

the symplectic form Ω_{FS} on $\mathbb{CP}^{n-1}$ is given by

$$\Omega_{FS}([z]) (T_z \pi(\mathbf{w}_1), T_z \pi(\mathbf{w}_2)) = -\operatorname{Im}(\mathbf{w}_1 \cdot \mathbf{w}_2).$$

This symplectic form is associated to the **Fubini-Study** Kähler metric on $\mathbb{CP}^{n-1}$, a subject not discussed in these lectures; it is the negative of its imaginary part.

We need to prove that $\hat{\iota}_{1/2}^* \Omega_{FS} = \Omega_{1/2}$. By the characterization of the reduced symplectic form $\Omega_{1/2}$, this is equivalent to $\iota_{1/2}^*(\mathbf{d}\mathbf{q} \wedge \mathbf{d}\mathbf{p}) = \pi_{1/2}^* \Omega_{1/2} = \pi_{1/2}^* \hat{\iota}_{1/2}^* \Omega_{FS} = (\hat{\iota}_{1/2} \circ \pi_{1/2})^* \Omega_{FS} = \varphi^* \Omega_{FS}$, which is the identity that shall be verified below. Since $\varphi = \pi \circ \iota_{1/2}$, we have for any $\mathbf{z} \in S^{2n-1}$ and any $\mathbf{w}_1, \mathbf{w}_2 \in T_{\mathbf{z}} S^{2n-1} = (\mathbb{C}\mathbf{z})^\perp$ the identity $T_{\mathbf{z}} \varphi(\mathbf{w}_j) = T_{\mathbf{z}} \pi(\mathbf{w}_j)$ for $j = 1, 2$, and hence

$$\begin{aligned} (\varphi^* \Omega_{FS})(\mathbf{z}) (\mathbf{w}_1, \mathbf{w}_2) &= \Omega_{FS}([z]) (T_{\mathbf{z}} \varphi(\mathbf{w}_1), T_{\mathbf{z}} \varphi(\mathbf{w}_2)) \\ &= \Omega_{FS}([z]) (T_{\mathbf{z}} \pi(\mathbf{w}_1), T_{\mathbf{z}} \pi(\mathbf{w}_2)) \\ &= -\operatorname{Im}(\mathbf{w}_1 \cdot \mathbf{w}_2) = \iota_{1/2}^*(\mathbf{d}\mathbf{q} \wedge \mathbf{d}\mathbf{p}) (\mathbf{w}_1, \mathbf{w}_2) \end{aligned}$$

as was remarked at the beginning of §2.1. This proves that *the symplectic reduced space $(H^{-1}(1/2)/S^1, \Omega_{1/2})$ is symplectically diffeomorphic to the complex projective space $(\mathbb{CP}^{n-1}, \Omega_{FS})$.*

Kaluza-Klein construction in electromagnetism. Let us revisit the motion of a particle with charge e and mass m moving in a given time independent divergence free magnetic field $\mathbf{B} := B_x \mathbf{i} + B_y \mathbf{j} + B_z \mathbf{k}$, where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ is the usual orthonormal basis of $\mathbb{R}^3$. In §2.5 we have shown that Newton's equations (2.13) for the Lorentz force law

$$m \frac{d\mathbf{v}}{dt} = \frac{e}{c} \mathbf{v} \times \mathbf{B},$$

where $\mathbf{v} := \dot{\mathbf{q}}$ is the velocity of the particle, are equivalent to Hamilton's equations in $T^*\mathbb{R}^3 := \{(\mathbf{q}, \mathbf{p}) \mid \mathbf{x} = (x, y, z), \mathbf{p} := m\mathbf{v} = (m\dot{x}, m\dot{y}, m\dot{z}) = (p_x, p_y, p_z) \in \mathbb{R}^3\}$ endowed with the magnetic symplectic form

$$\Omega_B = \mathbf{d}x \wedge \mathbf{d}p_x + \mathbf{d}y \wedge \mathbf{d}p_y + \mathbf{d}z \wedge \mathbf{d}p_z - \frac{e}{c} B$$

and the Hamiltonian given by the kinetic energy of the particle

$$H = \frac{1}{2m} \|\mathbf{p}\|^2 = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2);$$

B denotes the closed two-form on $\mathbb{R}^3$ associated to the divergence free vector field $\mathbf{B}$, that is,

$$B := \mathbf{i}_B(\mathbf{d}x \wedge \mathbf{d}y \wedge \mathbf{d}z) = B_x \mathbf{d}y \wedge \mathbf{d}z + B_y \mathbf{d}z \wedge \mathbf{d}x + B_z \mathbf{d}x \wedge \mathbf{d}y.$$

In addition, we have shown that writing $B = \mathbf{d}A$, or equivalently, $\mathbf{B} = \nabla \times \mathbf{A}$, where $\mathbf{A}^\flat = A$, that is, $\mathbf{A} = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}$ and $A = A_x dx + A_y dy + A_z dz$, the same Lorentz force law equations are Hamiltonian on $T^*\mathbb{R}^3$ endowed with the canonical symplectic structure but the momentum shifted Hamiltonian

$$H_A(\mathbf{q}, \mathbf{p}) = \frac{1}{2m} \left\| \mathbf{p} - \frac{e}{c} \mathbf{A} \right\|^2.$$

In §3.3, writing these equations in Lagrangian form, we remarked that they are not geodesic, essentially because of the magnetic symplectic form. Then we used the Kaluza-Klein construction to find a new Lagrangian L_{KK} (see (3.6)) on the enlarged configuration space $Q_{KK} := \mathbb{R}^3 \times S^1 = \{(\mathbf{q}, \theta) \mid \mathbf{q} \in \mathbb{R}^3, \theta \in S^1\}$ which turned out to be the quadratic form of an $\mathbf{A}$ -dependent Riemannian metric on Q_{KK} , called the Kaluza-Klein metric. Thus the Euler-Lagrange equations for L_{KK} are the geodesic equations of this metric on $\mathbb{R}^3 \times S^1$. Furthermore, we Legendre transformed L_{KK} to get a Hamiltonian (see (3.8)) on $T^*\mathbb{R}^3 = \{\mathbf{q}, \theta, \mathbf{p}, \pi \mid \mathbf{q}, \mathbf{p} \in \mathbb{R}^3, \theta \in S^1, \pi \in \mathbb{R}\}$ given by

$$H_{KK}(\mathbf{q}, \mathbf{p}, \theta, \pi) = \frac{1}{2m} \|\mathbf{p} - \pi \mathbf{A}\|^2 + \frac{1}{2} \pi^2.$$

Relative to the canonical symplectic form, Hamilton's equations for H_{KK} are the geodesic equations for the Kaluza-Klein metric (expressed in Hamiltonian form). Since H_{KK} does not depend on θ , π is conserved and setting $\pi = e/c$, we noted that H_{KK} , regarded as a function of only the variables $(\mathbf{q}, \mathbf{p})$, is up to a constant (namely $\pi^2/2$), equal to the momentum shifted Lorentz force Hamiltonian H_A (see (3.4)). These were just observations obtained by direct calculations.

Now we shall show how all of this is obtained from reduction theory. We start from the Hamiltonian system on $T^*Q_{KK} = T^*(\mathbb{R}^3 \times S^1)$ and note that S^1 acts on Q_{KK} by $\psi \cdot (\mathbf{q}, \theta) := (\mathbf{q}, \theta + \psi)$ where $\theta + \psi$ is taken modulo 1 (so we normalize the length of the circle to be 1). The infinitesimal generator defined by $\xi \in \mathbb{R}$ for this action is $\xi_{Q_{KK}}(\mathbf{q}, \theta) = (\mathbf{q}, \theta; \mathbf{0}, \xi)$. The Hamiltonian H_{KK} is obviously invariant under this action. The momentum map for this action, given by (5.11), is in this case

$$\langle \mathbf{J}(\mathbf{q}, \theta, \mathbf{p}, \pi), \xi \rangle = (\mathbf{p} d\mathbf{q} + \pi d\theta) \xi \frac{\partial}{\partial \theta} = \xi \pi,$$

that is, $\mathbf{J}(\mathbf{q}, \theta, \mathbf{p}, \pi) = \pi$, thus recovering the direct observation that π is a conserved quantity. Moreover, any value of $\mathbf{J}$ in $\mathbb{R}$ is a regular value and we have $\mathbf{J}^{-1}(e/c) = T^*\mathbb{R}^3 \times S^1 \times \{e/c\}$ on which S^1 acts on the S^1 -factor only. Thus the reduced space is $(T^*Q_{KK})_{e/c} = T^*\mathbb{R}^3$. The reduced Hamiltonian is obviously $H_A + e^2/2c^2$, so the reduced dynamics is given by H_A . Finally we

need to compute the reduced symplectic form Ω_{red} . We have

$$\begin{aligned} (\pi_{e/c}^* \Omega_{red})(\mathbf{q}, \theta, \mathbf{p}, \pi) & \left(\mathbf{X} \frac{\partial}{\partial \mathbf{q}} + a \frac{\partial}{\partial \theta} + \mathbf{Y} \frac{\partial}{\partial \mathbf{p}}, \mathbf{X}' \frac{\partial}{\partial \mathbf{q}} + a' \frac{\partial}{\partial \theta} + \mathbf{Y}' \frac{\partial}{\partial \mathbf{p}} \right) \\ & = \iota_{e/c}^* (\mathbf{d}\mathbf{q} \wedge \mathbf{d}\mathbf{p} + \mathbf{d}\theta \wedge \mathbf{d}\pi) \left(\mathbf{X} \frac{\partial}{\partial \mathbf{q}} + a \frac{\partial}{\partial \theta} + \mathbf{Y} \frac{\partial}{\partial \mathbf{p}}, \right. \\ & \quad \left. \mathbf{X}' \frac{\partial}{\partial \mathbf{q}} + a' \frac{\partial}{\partial \theta} + \mathbf{Y}' \frac{\partial}{\partial \mathbf{p}} \right) \\ & = \mathbf{X} \cdot \mathbf{Y}' - \mathbf{X}' \cdot \mathbf{Y}. \end{aligned}$$

Since

$$T_{(\mathbf{q}, \theta, \mathbf{p}, e/c)} \pi_{e/c} \left(\mathbf{X} \frac{\partial}{\partial \mathbf{q}} + a \frac{\partial}{\partial \theta} + \mathbf{Y} \frac{\partial}{\partial \mathbf{p}} \right) = \mathbf{X} \frac{\partial}{\partial \mathbf{q}} + \mathbf{Y} \frac{\partial}{\partial \mathbf{p}}$$

this identity shows that Ω_{red} is the canonical symplectic form on the reduced space $T^*\mathbb{R}^3$, thus recovering the second result in §2.5: *the Lorentz force law equations are Hamiltonian on $T^*\mathbb{R}^3$ relative to the canonical symplectic form and the Hamiltonian function H_A .*

Let us carry out the reduction in a different manner, by insisting that we get on the reduced space the kinetic energy Hamiltonian $\|\mathbf{p}\|^2/2m$. To do this, requires that we project $\mathbf{J}^{-1}(e/c) \rightarrow T^*\mathbb{R}^3$ by the S^1 -invariant smooth map

$$\varphi \left(\mathbf{q}, \theta, \mathbf{p}, \frac{e}{c} \right) := \left(\mathbf{q}, \mathbf{p} - \frac{e}{c} \mathbf{A} \right).$$

It is clear from the computation above that the reduced symplectic form will not be the canonical one anymore, since we are using a different map to identify $(T^*Q_{KK})_{e/c}$ with $T^*\mathbb{R}^3$. So what is the symplectic form now?

We compute it by using φ as the projection from the reduction theorem and get as before

$$\begin{aligned} (\varphi^* \Omega_{e/c})(\mathbf{q}, \theta, \mathbf{p}, \pi) & \left(\mathbf{X} \frac{\partial}{\partial \mathbf{q}} + a \frac{\partial}{\partial \theta} + \mathbf{Y} \frac{\partial}{\partial \mathbf{p}}, \mathbf{X}' \frac{\partial}{\partial \mathbf{q}} + a' \frac{\partial}{\partial \theta} + \mathbf{Y}' \frac{\partial}{\partial \mathbf{p}} \right) \\ & = \iota_{e/c}^* (\mathbf{d}\mathbf{q} \wedge \mathbf{d}\mathbf{p} + \mathbf{d}\theta \wedge \mathbf{d}\pi) \left(\mathbf{X} \frac{\partial}{\partial \mathbf{q}} + a \frac{\partial}{\partial \theta} + \mathbf{Y} \frac{\partial}{\partial \mathbf{p}}, \right. \\ & \quad \left. \mathbf{X}' \frac{\partial}{\partial \mathbf{q}} + a' \frac{\partial}{\partial \theta} + \mathbf{Y}' \frac{\partial}{\partial \mathbf{p}} \right) \\ & = \mathbf{X} \cdot \mathbf{Y}' - \mathbf{X}' \cdot \mathbf{Y}. \end{aligned}$$

However,

$$T_{(\mathbf{q}, \theta, \mathbf{p}, e/c)} \varphi \left(\mathbf{X} \frac{\partial}{\partial \mathbf{q}} + a \frac{\partial}{\partial \theta} + \mathbf{Y} \frac{\partial}{\partial \mathbf{p}} \right) = \mathbf{X} \frac{\partial}{\partial \mathbf{q}} + \left(\mathbf{Y} - \frac{e}{c} (\mathbf{X} \cdot \nabla) \mathbf{A} \right) \frac{\partial}{\partial \mathbf{p}},$$

so the previous equality yields

$$\begin{aligned} \Omega_{e/c} \left(\mathbf{q}, \mathbf{p} - \frac{e}{c} \mathbf{A} \right) & \left(\mathbf{X} \frac{\partial}{\partial \mathbf{q}} + \left(\mathbf{Y} - \frac{e}{c} (\mathbf{X} \cdot \nabla) \mathbf{A} \right) \frac{\partial}{\partial \mathbf{p}}, \right. \\ & \left. \mathbf{X}' \frac{\partial}{\partial \mathbf{q}} + \left(\mathbf{Y}' - \frac{e}{c} (\mathbf{X}' \cdot \nabla) \mathbf{A} \right) \frac{\partial}{\partial \mathbf{p}} \right) \\ & = \mathbf{X} \cdot \mathbf{Y}' - \mathbf{X}' \cdot \mathbf{Y}. \end{aligned}$$

Replacing now $\mathbf{p}$ by $\mathbf{p} + \frac{e}{c} \mathbf{A}$, noting that the right hand side of this equality does not depend on $\mathbf{p}$, replacing $\mathbf{Y}$ by $\mathbf{Y} + \frac{e}{c} (\mathbf{X} \cdot \nabla) \mathbf{A}$, and $\mathbf{Y}'$ by $\mathbf{Y}' + \frac{e}{c} (\mathbf{X}' \cdot \nabla) \mathbf{A}$, yields

$$\begin{aligned} \Omega_{e/c}(\mathbf{q}, \mathbf{p}) & \left(\mathbf{X} \frac{\partial}{\partial \mathbf{q}} + \mathbf{Y} \frac{\partial}{\partial \mathbf{p}}, \mathbf{X}' \frac{\partial}{\partial \mathbf{q}} + \mathbf{Y}' \frac{\partial}{\partial \mathbf{p}} \right) \\ & = \mathbf{X} \cdot \left(\mathbf{Y}' + \frac{e}{c} (\mathbf{X}' \cdot \nabla) \mathbf{A} \right) - \mathbf{X}' \cdot \left(\mathbf{Y} + \frac{e}{c} (\mathbf{X} \cdot \nabla) \mathbf{A} \right) \\ & = \mathbf{X} \cdot \mathbf{Y}' - \mathbf{X}' \cdot \mathbf{Y} - \frac{e}{c} (\mathbf{X}' \cdot ((\mathbf{X} \cdot \nabla) \mathbf{A}) - \mathbf{X} \cdot ((\mathbf{X}' \cdot \nabla) \mathbf{A})) \\ & = \left(d\mathbf{q} \wedge d\mathbf{p} - \frac{e}{c} B \right) \left(\mathbf{X} \frac{\partial}{\partial \mathbf{q}} + \mathbf{Y} \frac{\partial}{\partial \mathbf{p}}, \mathbf{X}' \frac{\partial}{\partial \mathbf{q}} + \mathbf{Y}' \frac{\partial}{\partial \mathbf{p}} \right) \end{aligned}$$

since

$$B \left(\mathbf{X} \frac{\partial}{\partial \mathbf{q}}, \mathbf{X}' \frac{\partial}{\partial \mathbf{q}} \right) = \mathbf{X}' \cdot ((\mathbf{X} \cdot \nabla) \mathbf{A}) - \mathbf{X} \cdot ((\mathbf{X}' \cdot \nabla) \mathbf{A})$$

as a straightforward computation in coordinates shows. Thus

$$\Omega_{e/c} = d\mathbf{q} \wedge d\mathbf{p} - \frac{e}{c} B = \Omega_B.$$

which proves the first assertion in the example of §2.5 and explains the appearance of the magnetic term B in the symplectic form: *the Lorentz force law equations are Hamiltonian on $T^*\mathbb{R}^3$ relative to the symplectic form Ω_B and the kinetic energy $\|\mathbf{p}\|^2/2m$ as Hamiltonian function.*

The phenomenon occurring here is very general and has to do with cotangent bundle reduction, a topic not covered in these lectures. What is happening here is the following: one has a principal bundle with a connection and searches for explicit realizations of the reduced spaces of the cotangent bundle of the total space by the structure group. It turns out that there are two natural ways to carry out this reduction, both at Poisson and at symplectic level. For details of this theory see [MaRa03] and [PR04]; for a quick summary without proofs see [OR04].

Coadjoint Orbits. For the Lie group G denote by $L_g, R_g : G \rightarrow G$ the left and right translations on G by $g \in G$. The lifts of these actions to T^*G , denoted by $\bar{L}, \bar{R} : G \times T^*G \rightarrow T^*G$ respectively, are proper, free, and have equivariant momentum maps $\mathbf{J}_L, \mathbf{J}_R : T^*G \rightarrow \mathfrak{g}^*$ given by (see (5.16))

$$\mathbf{J}_L(\alpha_g) = T_e^* R_g(\alpha_g), \quad \mathbf{J}_R(\alpha_g) = T_e^* L_g(\alpha_g).$$

Recall that

$$\bar{L}_h(\alpha_g) = T_{hg}^* L_{h^{-1}}(\alpha_g), \quad \bar{R}_h(\alpha_g) = T_{gh}^* R_{h^{-1}}(\alpha_g)$$

for any $h \in G$ and $\alpha_g \in T^*G$.

Let us compute the point reduced space $(T^*G)_\mu$ for any $\mu \in \mathfrak{g}^*$ relative to the left action $\bar{L}$. Notice that we do *not* require μ to be a generic point in $\mathfrak{g}^*$; that is, arbitrarily nearby coadjoint orbits may have a different dimensions. Since the action $\bar{L}$ is free, the symmetry algebra of every point in T^*G is zero and thus every $\mu \in \mathfrak{g}^*$ is a regular value of $\mathbf{J}_L$. Thus, the hypotheses of the Symplectic Point Reduction Theorem hold.

The submanifold $\mathbf{J}_L^{-1}(\mu) = \{\alpha_g \in T^*G \mid T_e^* R_g(\alpha_g) = \mu\} = \{T_g^* R_{g^{-1}} \mu \mid g \in G\}$ is the graph of the right invariant one-form α_μ on G whose value at the identity is μ . Thus $\alpha_\mu : G \rightarrow \mathbf{J}_L^{-1}(\mu)$ is a diffeomorphism. Let us show that it is G_μ -equivariant, that is,

$$\alpha_\mu \circ L_h = \bar{L}_h \circ \alpha_\mu \quad \text{for all } h \in G_\mu.$$

Indeed, for any $g \in G$ we have

$$\begin{aligned} (\bar{L}_h \circ \alpha_\mu)(g) &= \lambda_h(T_g^* R_{g^{-1}} \mu) = T_{hg}^* L_{h^{-1}} T_g^* R_{g^{-1}} \mu = T_{hg}^* R_{g^{-1}} T_h^* L_{h^{-1}} \mu \\ &= T_{hg}^* R_{g^{-1}} T_h^* R_{h^{-1}} T_e^* R_h T_h^* L_{h^{-1}} \mu = T_{hg}^* R_{(hg)^{-1}} \text{Ad}_{h^{-1}}^* \mu \\ &= T_{hg}^* R_{(hg)^{-1}} \mu = \alpha_\mu(hg). \end{aligned}$$

Therefore $\alpha_\mu : G \rightarrow \mathbf{J}_L^{-1}(\mu)$ induces a diffeomorphism $\bar{\alpha}_\mu : G/G_\mu \rightarrow (T^*G)_\mu$, where $G/G_\mu := \{G_\mu g \mid g \in G\}$. Recall now that the map $\varepsilon_\mu : G/G_\mu \rightarrow \mathcal{O}_\mu$ given by $\varepsilon_\mu(G_\mu g) = \text{Ad}_g^* \mu$ is the diffeomorphism that defines the smooth manifold structure of the orbit $\mathcal{O}_\mu$. Define the diffeomorphism $\bar{\varphi} : (T^*G)_\mu = \mathbf{J}_L^{-1}(\mu)/G_\mu \rightarrow \mathcal{O}_\mu$ by $\bar{\varphi} := \varepsilon_\mu \circ (\bar{\alpha}_\mu)^{-1}$ and note that $\varphi := \bar{\varphi} \circ \pi_\mu = \mathbf{J}_R|_{\mathbf{J}_L^{-1}(\mu)} : \mathbf{J}_L^{-1}(\mu) \rightarrow \mathcal{O}_\mu$ has the expression $\varphi(T_g^* R_{g^{-1}} \mu) = \text{Ad}_g^* \mu$.

$$\begin{array}{ccccc}
\mathbf{J}_L^{-1}(\mu) & \xleftarrow{\alpha_\mu} & G & & \\
\downarrow \pi_\mu & & \downarrow & & \\
(T^*G)_\mu & \xleftarrow{\bar{\alpha}_\mu} & G/G_\mu & \xrightarrow{\varepsilon_\mu} & \mathcal{O}_\mu
\end{array}$$

We claim that $\bar{\varphi}_* \Omega_\mu = \omega^-$, where Ω_μ is the reduced symplectic form and ω^- is the minus orbit symplectic form on $\mathcal{O}_\mu$. Since Ω_μ is characterized by the identity $\pi_\mu^* \Omega_\mu = \iota_\mu^* \Omega = -\iota_\mu^* \mathbf{d}\Theta$, where Θ is the canonical one-form on T^*G , this relation is equivalent to $\varphi^* \omega^- = -\iota_\mu^* \mathbf{d}\Theta$. We shall now prove this identity. In what follows we shall denote by $\text{Ad}_g : G \rightarrow G$ the conjugation automorphism $\text{Ad}_g(h) := ghg^{-1}$, for any $h \in G$. Thus $T_e \text{Ad}_g = \text{Ad}_g : \mathfrak{g} \rightarrow \mathfrak{g}$ is the adjoint representation of G on $\mathfrak{g}$.

Let

$$\alpha_\xi(t) = T_{g \exp(t\xi)}^* R_{\exp(-t\xi)g^{-1}} \mu$$

be an arbitrary smooth curve in $\mathbf{J}_L^{-1}(\mu)$ passing through $\alpha_\xi(0) = T_g^* R_{g^{-1}} \mu$, where $\xi \in \mathfrak{g}$. Since $R_{\exp(-t\xi)g^{-1}} = \text{Ad}_g \circ R_{\exp(-t\xi)} \circ L_{g^{-1}}$, we get

$$T_{g \exp(t\xi)} R_{\exp(-t\xi)g^{-1}} = \text{Ad}_g \circ T_{\exp(t\xi)} R_{\exp(-t\xi)} \circ T_{g \exp(t\xi)} L_{g^{-1}},$$

so letting $\nu = \text{Ad}_g^* \mu$, we have

$$\alpha_\xi(t) = T_{g \exp(t\xi)}^* L_{g^{-1}} T_{\exp(t\xi)}^* R_{\exp(-t\xi)} \nu = \bar{L}_g \bar{R}_{\exp(t\xi)} \nu.$$

Therefore, an arbitrary tangent vector at $T_g^* R_{g^{-1}} \mu \in \mathbf{J}_L^{-1}(\mu)$ to the submanifold $\mathbf{J}_L^{-1}(\mu)$ has the expression

$$\alpha'_\xi(0) = T_\nu \bar{L}_g (\xi_{T^*G}^R(\nu)), \quad (7.3)$$

where $\xi_{T^*G}^R$ is the infinitesimal generator of the right action $\bar{R}$ on T^*G .

Now note that $\varphi(\alpha_\xi(t)) = \text{Ad}_{g \exp(t\xi)}^* \mu = \text{Ad}_{\exp(t\xi)}^* \text{Ad}_g^* \mu = \text{Ad}_{\exp(t\xi)}^* \nu$ so that

$$T_{\alpha_\xi(0)} \varphi(\alpha'_\xi(0)) = \left. \frac{d}{dt} \right|_{t=0} \varphi(\alpha_\xi(t)) = \left. \frac{d}{dt} \right|_{t=0} \text{Ad}_{\exp(t\xi)}^* \nu = \text{ad}_\xi^* \nu. \quad (7.4)$$

By (7.4), we have for any $\xi, \eta \in \mathfrak{g}$ and $\nu = \text{Ad}_g^* \mu$,

$$\begin{aligned} & (\varphi^* \omega^-)(T_g^* R_{g^{-1}} \mu)(\alpha'_\xi(0), \alpha'_\eta(0)) \\ &= \omega^-(\nu)(T_{\alpha'_\xi(0)} \varphi(\alpha'_\xi(0)), T_{\alpha'_\eta(0)} \varphi(\alpha'_\eta(0))) \\ &= \omega^-(\nu)(\text{ad}_\xi^* \nu, \text{ad}_\eta^* \nu) = -\langle \nu, [\xi, \eta] \rangle \\ &= -J_R^{[\xi, \eta]}(\nu) = \{J_R^\xi, J_R^\eta\}(\nu), \end{aligned} \quad (7.5)$$

since for *right* actions we have $J_R^{[\xi, \eta]} = -\{J_R^\xi, J_R^\eta\}$ (see (5.12) for *left* actions).

On the other hand, since $T_g^* R_{g^{-1}} \mu = \bar{L}_g \nu$, the expression (7.3) and left invariance of Θ give

$$\begin{aligned} & (\iota^* \mathbf{d}\Theta)(T_g^* R_{g^{-1}} \mu)(\alpha'_\xi(0), \alpha'_\eta(0)) \\ &= \mathbf{d}\Theta(\bar{L}_g \nu)(T_\nu \bar{L}_g(\xi_{T^*G}^R(\nu)), T_\nu \bar{L}_g(\eta_{T^*G}^R(\nu))) \\ &= \mathbf{d}\Theta(\nu)(\xi_{T^*G}^R(\nu), \eta_{T^*G}^R(\nu)) \\ &= \xi_{T^*G}^R[\Theta(\eta_{T^*G}^R)](\nu) - \eta_{T^*G}^R[\Theta(\xi_{T^*G}^R)](\nu) - \Theta([\xi, \eta]_{T^*G}^R)(\nu). \end{aligned}$$

If $\pi : T^*G \rightarrow G$ is the cotangent bundle projection, the definition (2.8) of Θ gives for any $\alpha_g \in T^*G$,

$$\Theta(\eta_{T^*G}^R)(\alpha_g) = \langle \alpha_g, T_{\alpha_g} \pi(\eta_{T^*G}^R) \rangle = \langle \alpha_g, \eta_G^R(g) \rangle = J_R^\eta(\alpha_g)$$

since $\eta_{T^*G}^R$ and η_G^R are π -related. Thus, using the identity $[\xi_{T^*G}^R, \eta_{T^*G}^R] = [\xi, \eta]_{T^*G}^R$ valid for *right* actions (see (5.8) for *left* actions), as well as the definition of the momentum map, we can continue the computation above and write

$$\begin{aligned} & \xi_{T^*G}^R[\Theta(\eta_{T^*G}^R)](\nu) - \eta_{T^*G}^R[\Theta(\xi_{T^*G}^R)](\nu) - \Theta([\xi, \eta]_{T^*G}^R)(\nu) \\ &= X_{J_R^\xi}[J_R^\eta](\nu) - X_{J_R^\eta}[J_R^\xi](\nu) - J_R^{[\xi, \eta]}(\nu) \\ &= \{J_R^\eta, J_R^\xi\}(\nu) - \{J_R^\xi, J_R^\eta\}(\nu) + \{J_R^\xi, J_R^\eta\}(\nu) = -\{J_R^\xi, J_R^\eta\}(\nu). \end{aligned}$$

Thus we have proved the identity

$$-(\iota_\mu^* \mathbf{d}\Theta)(T_g^* R_{g^{-1}} \mu)(\alpha'_\xi(0), \alpha'_\eta(0)) = \{J_R^\xi, J_R^\eta\}(\nu). \quad (7.6)$$

Equations (7.5) and (7.6) prove that $\varphi^* \omega^- = -\iota_\mu^* \mathbf{d}\Theta$.

7.4 Orbit Reduction

Let us return to Remark (7) in §7.1 where we have commented on the fact that the inclusion $\iota_\mu : \mathbf{J}^{-1}(\mu) \hookrightarrow P$ induces a Poisson injective immersion $\kappa_\mu : P_\mu \rightarrow P/G$. So, the κ_μ -images in P/G of the connected components

of the point reduced symplectic manifolds (P_μ, Ω_μ) are the symplectic leaves of P/G . In this section we shall present, without proofs, how this is actually carried out concretely.

We begin with the easy observation that, as sets, $\kappa_\mu(P_\mu) = \mathbf{J}^{-1}(\mathcal{O}_\mu)/G$, where $\mathcal{O}_\mu$ is the coadjoint orbit through $\mu \in \mathfrak{g}^*$. If the momentum map $\mathbf{J} : P \rightarrow \mathfrak{g}^*$ is not equivariant then instead of the coadjoint orbit through μ one considers the orbit of the affine action $g \cdot \mu := \text{Ad}_{g^{-1}}^* \mu + \sigma(g)$, where σ is the $\mathfrak{g}^*$ -valued nonequivariance group one-cocycle defined by $\mathbf{J}$; we need to assume here that P is connected (see Remark (8) in §7.1). Recall that $\sigma(g) := \mathbf{J}(g \cdot z) - \text{Ad}_{g^{-1}}^* \mathbf{J}(z)$ and that if P is connected the right hand side of this equation is independent of $z \in P$. The group one-cocycle σ induces by derivation a real valued Lie algebra two-cocycle $\Sigma : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{R}$ which can be shown to equal $\Sigma(\xi, \eta) = J^{[\xi, \eta]}(z) - \{J^\xi, J^\eta\}(z)$ for every $z \in P$ and $\xi, \eta \in \mathfrak{g}$. Denote by $\xi_{\mathfrak{g}^*}(\nu) := -\text{ad}_\xi^* \nu + \Sigma(\xi, \cdot)$ the infinitesimal generator of the affine action of G on $\mathfrak{g}^*$, where $\nu \in \mathfrak{g}^*$. The affine action orbit $\mathcal{O}_\mu$ carries two symplectic forms given by

$$\omega_{\mathcal{O}_\mu}^\pm(\nu)(\xi_{\mathfrak{g}^*}(\nu), \eta_{\mathfrak{g}^*}(\nu)) = \pm \langle \nu, [\xi, \eta] \rangle \mp \Sigma(\xi, \eta), \quad (7.7)$$

for any $\xi, \eta \in \mathfrak{g}$. They are the natural modifications of the usual orbit symplectic forms on coadjoint orbits. For the proofs of the statements above see [AbMa78], [LiMa87], or [OR04]. From now on we shall not make any equivariance hypotheses on $\mathbf{J}$ and shall work with the affine orbit $\mathcal{O}_\mu \subset \mathfrak{g}^*$ through μ . The set $P_{\mathcal{O}_\mu} := \mathbf{J}^{-1}(\mathcal{O}_\mu)/G$ is called the **orbit reduced space** associated to the orbit $\mathcal{O}_\mu$. The smooth manifold structure (and hence the topology) on $P_{\mathcal{O}_\mu}$ is the one that makes the bijective map $\kappa_\mu : P_\mu \rightarrow P_{\mathcal{O}_\mu}$ into a diffeomorphism.

The next theorem characterizes the symplectic form and the Hamiltonian dynamics on $P_{\mathcal{O}_\mu}$.

Theorem 7.4 (Symplectic orbit reduction) *Assume that the free proper symplectic action of the Lie group G on the symplectic manifold (P, Ω) admits an associated momentum map $\mathbf{J} : P \rightarrow \mathfrak{g}^*$.*

- (i) *On $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ there is a unique immersed smooth manifold structure such that the projection $\pi_{\mathcal{O}_\mu} : \mathbf{J}^{-1}(\mathcal{O}_\mu) \rightarrow P_{\mathcal{O}_\mu}$ is a surjective submersion, where $P_{\mathcal{O}_\mu}$ is endowed with the manifold structure making κ_μ into a diffeomorphism. This smooth manifold structure does not depend on the choice of μ in the orbit $\mathcal{O}_\mu$. If $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ is a submanifold of P in its own right, then the immersed topology by κ_μ and the induced topology on $P_{\mathcal{O}_\mu}$ coincide.*

- (ii) $P_{\mathcal{O}_\mu}$ is a symplectic manifold with the symplectic form $\Omega_{\mathcal{O}_\mu}^\mathfrak{X}$ uniquely characterized by the relation

$$\iota_{\mathcal{O}_\mu}^* \Omega = \pi_{\mathcal{O}_\mu}^* \Omega_{\mathcal{O}_\mu}^\mathfrak{X} + \mathbf{J}_{\mathcal{O}_\mu}^* \omega_{\mathcal{O}_\mu}^+,$$

where $\mathbf{J}_{\mathcal{O}_\mu}$ is the restriction of $\mathbf{J}$ to $\mathbf{J}^{-1}(\mathcal{O}_\mu)$, $\iota_{\mathcal{O}_\mu} : \mathbf{J}^{-1}(\mathcal{O}_\mu) \hookrightarrow P$ is the inclusion, and $\omega_{\mathcal{O}_\mu}^+$ is the +orbit symplectic form on $\mathcal{O}_\mu$ given by (7.7). The symplectic manifolds P_μ and $P_{\mathcal{O}_\mu}$ are symplectically diffeomorphic by κ_μ .

- (iii) Let H be a G -invariant function on P and define $\tilde{H} : P/G \rightarrow \mathbb{R}$ by $H = \tilde{H} \circ \pi$. Then the Hamiltonian vector field X_H is also G -invariant and hence induces a vector field on P/G which coincides with the Hamiltonian vector field $X_{\tilde{H}}$. Moreover, the flow of $X_{\tilde{H}}$ leaves the symplectic leaves $P_{\mathcal{O}_\mu}$ of P/G invariant. This flow restricted to the symplectic leaves is again Hamiltonian relative to the symplectic form $\Omega_{\mathcal{O}_\mu}^\mathfrak{X}$ and the Hamiltonian function $H_{\mathcal{O}_\mu}$ given by

$$H_{\mathcal{O}_\mu} \circ \pi_{\mathcal{O}_\mu} = H \circ i_{\mathcal{O}_\mu} \quad \text{or} \quad H_{\mathcal{O}_\mu} = \tilde{H}|_{\mathcal{O}_\mu}.$$

Moreover, if $F : P \rightarrow \mathbb{R}$ is another smooth G -invariant function, then $\{F, H\}$ is also G -invariant and $\{F, H\}_{\mathcal{O}_\mu} = \{F_{\mathcal{O}_\mu}, H_{\mathcal{O}_\mu}\}_{P_{\mathcal{O}_\mu}}$, where $\{\cdot, \cdot\}_{\mathcal{O}_\mu}$ is the Poisson bracket on the symplectic manifold $P_{\mathcal{O}_\mu} = \mathbf{J}^{-1}(\mathcal{O}_\mu)/G$.

The proof of this theorem in the regular case and when $\mathcal{O}_\mu$ is an embedded submanifold of $\mathfrak{g}^*$ can be found in [Marle76], [KaKoSt78], and [Marsden81]. For the general case, when $\mathcal{O}_\mu$ is not a submanifold of $\mathfrak{g}^*$ see [OR04]. Here is the main idea of the proof. Consider for each value $\mu \in \mathfrak{g}^*$ of $\mathbf{J}$ the G -equivariant bijection

$$s : [g, z] \in G \times_{G_\mu} \mathbf{J}^{-1}(\mu) \mapsto g \cdot z \in \mathbf{J}^{-1}(\mathcal{O}_\mu),$$

where $G \times_{G_\mu} \mathbf{J}^{-1}(\mu) := (G \times \mathbf{J}^{-1}(\mu))/G_\mu$, the G_μ -action being the diagonal action. Endow $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ with the smooth manifold structure that makes the bijection s into a diffeomorphism. Then $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ with this smooth structure is an immersed submanifold of P . This is the manifold structure on $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ used in the statement of Theorem 7.4.

In the particular case when $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ is a smooth submanifold of P in its own right, this manifold structure coincides with the one induced by the mapping s described above since in this situation the bijection s becomes a diffeomorphism relative to the a priori given smooth manifold structure on $\mathbf{J}^{-1}(\mathcal{O}_\mu)$.

If μ is a regular value of $\mathbf{J}$ and $\mathcal{O}_\mu$ is an embedded submanifold of $\mathfrak{g}^*$,

then $\mathbf{J}$ is transverse to $\mathcal{O}_\mu$ and hence $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ is automatically an embedded submanifold of P .

The orbit $\mathcal{O}_\mu$ can be used to transform point reduction at an arbitrary $\mu \in \mathfrak{g}^*$ to point reduction at zero for a larger manifold. Suppose that we are in the hypotheses of the Symplectic Point Reduction Theorem 7.1. Form the point reduced space P_μ and consider the G -orbit through μ in $\mathfrak{g}^*$ (in general the orbit under the affine action) endowed with the $+$ orbit symplectic form. The group G acts canonically on the left on $\mathcal{O}_\mu$ with momentum map given by the inclusion $i : \mathcal{O}_\mu \hookrightarrow \mathfrak{g}^*$. Let $P \oplus \mathcal{O}_\mu$ denote the symplectic manifold $P \times \mathcal{O}_\mu$ endowed with the symplectic structure $\Omega - \omega_{\mathcal{O}_\mu}^+ := \pi_1^* \Omega - \pi_2^* \omega_{\mathcal{O}_\mu}^+$, where $\pi_1 : P \times \mathcal{O}_\mu \rightarrow P$ and $\pi_2 : P \times \mathcal{O}_\mu \rightarrow \mathcal{O}_\mu$ are the projections on the first and second factors respectively. The Lie group G acts canonically on $P \oplus \mathcal{O}_\mu$ by $g \cdot (z, \nu) := (g \cdot z, \text{Ad}_{g^{-1}}^* \nu)$. As discussed in §5.3, example (6), this action has the momentum map $\mathbf{J} - i : P \oplus \mathcal{O}_\mu \rightarrow \mathfrak{g}^*$. This momentum map is equivariant if $\mathbf{J}$ is, in which case, $\mathcal{O}_\mu$ is taken to be the coadjoint orbit. With these notations we have the following result.

Theorem 7.5 (Shifting theorem) *The reduced symplectic manifolds P_μ and $(P \oplus \mathcal{O}_\mu)_0$ are symplectically diffeomorphic.*

One should not read into this theorem more than it states. It is tempting to quote it in order to dismiss the reduction procedure at all points $\mu \neq 0$. This would be an error, for the price one pays to reduce only at zero is heavy: the original phase space is enlarged by multiplication with the orbit $\mathcal{O}_\mu$ whose topology can be quite involved and who is, in general, not an embedded submanifold of $\mathfrak{g}^*$. Specifically, when dealing with singular reduction it is important to study reduction at non-zero values of the momentum map carefully. The Shifting Theorem 7.5 only hides the differential topological difficulties by burying them into $\mathcal{O}_\mu$.

7.5 Semidirect Product Reduction

In this section we present the general Semidirect Product Reduction Theorem as found in [MaRaWe84a, MaRaWe84b]. We do not attempt to give a history of the subject here since it can be found in many other papers and books. To avoid any technical complications, all of this section deals only with finite dimensional objects, even though the range of applicability of the theorems presented here goes far beyond that to many continuum and quantum mechanical systems.

Let V be a vector space and assume that $\sigma : G \rightarrow \text{Aut}(V)$ is a representation of the Lie group G on V ; $\text{Aut}(V)$ denotes the Lie group of linear isomorphisms of V onto itself whose Lie algebra is $\text{End}(V)$, the space of all linear maps of V to itself. Denote by $\sigma' : \mathfrak{g} \rightarrow \text{End}(V)$ the induced Lie algebra representation, that is,

$$\xi \cdot v := \xi_V(v) := \sigma'(\xi)v := \left. \frac{d}{dt} \right|_{t=0} \sigma(\exp t\xi)v$$

Given G , V , and ρ define the semidirect product $S := G \ltimes V$ as the Lie group whose underlying manifold is $G \times V$ and multiplication

$$(g_1, v_1)(g_2, v_2) := (g_1 g_2, v_1 + \sigma(g_1)v_2)$$

for $g_1, g_2 \in G$ and $v_1, v_2 \in V$. The identity element is $(e, 0)$ and $(g, v)^{-1} = (g^{-1}, -\sigma(g^{-1})v)$. Note that V is a normal subgroup of S and that $S/V = G$.

Let $\mathfrak{g}$ be the Lie algebra of G and let $\mathfrak{s} := \mathfrak{g} \ltimes V$ be the Lie algebra of S ; it is the semidirect product of $\mathfrak{g}$ with V using the representation σ' and its underlying vector space is $\mathfrak{g} \times V$. The Lie bracket on $\mathfrak{s}$ is given by

$$[(\xi_1, v_1), (\xi_2, v_2)] = ([\xi_1, \xi_2], \sigma'(\xi_1)v_2 - \sigma'(\xi_2)v_1)$$

for $\xi_1, \xi_2 \in \mathfrak{g}$ and $v_1, v_2 \in V$. Identify $\mathfrak{s}^*$ with $\mathfrak{g}^* \times V^*$ by using the duality pairing on each factor. The following formulas are useful for our next considerations. They are obtained by straightforward (and sometimes lengthy) computations:

- the adjoint action of S on $\mathfrak{s}$:

$$\text{Ad}_{(g,u)}(\xi, v) = (\text{Ad}_g \xi, \sigma(g)v - \sigma'(\text{Ad}_g \xi)u), \text{ for } (g, u) \in S, (\xi, v) \in \mathfrak{s};$$

- the coadjoint action of S on $\mathfrak{s}^*$:

$$\text{Ad}_{(g,u)}^*(\nu, a) = (\text{Ad}_{g^{-1}}^* \nu + (\sigma'_u)^* \sigma_*(g)a, \sigma_*(g)a),$$

for $(g, u) \in S$, $(\nu, a) \in \mathfrak{s}^*$, where $\sigma_*(g) := \sigma(g^{-1})^* \in \text{Aut}(V^*)$, $\sigma'_u : \mathfrak{g} \rightarrow V$ is the linear map given by $\sigma'_u(\xi) := \sigma'(\xi)u$ and $(\sigma'_u)^* : V^* \rightarrow \mathfrak{g}^*$ is its dual;

- the lift λ of left translation of S on T^*S :

$$\lambda((g, u), (\alpha_h, v, a)) = (T_{gh}^* L_{g^{-1}} \alpha_h, u + \sigma(g)v, \sigma_*(g)a)$$

for $(g, u) \in S$, and $(\alpha_h, v, a) \in T_{(h,v)}^* S = T_h^* G \times \{v\} \times V^*$; λ induces a canonical S -action on the product Poisson manifold $T^*G \times V^*$ by ignoring the third factor (V^* has the trivial Poisson bracket);

- the lift ρ of right translation of S on T^*S :

$$\rho((g, u), (\alpha_h, v, a)) = \left(T_{gh}^* R_{g^{-1}} \alpha_h - \mathbf{d}f_{\sigma(g^{-1})u}^a(hg), v + \sigma(h)u, a \right)$$

for $(g, u) \in S$, and $(\alpha_h, v, a) \in T_{(h,v)}^* S = T_h^* G \times \{v\} \times V^*$, and where $f_u^a : G \rightarrow \mathbb{R}$ is the “matrix element function” $f_u^a(g) := \langle a, \sigma(g)u \rangle$; ρ induces a canonical S -action on $T^*G \times V^*$ by ignoring the third factor (V^* has the trivial Poisson bracket);

- the momentum map $\mathbf{J}_L : T^*S \rightarrow \mathfrak{s}_+^*$ for the left translation λ :

$$\mathbf{J}_L(\alpha_g, v, a) = T_{(e,0)}^* R_{(g,v)}(\alpha_g, v, a) = (T_e^* R_g \alpha_g + (\sigma'_v)^* a, a); \quad (7.8)$$

- the momentum map $\mathbf{J}_R : T^*S \rightarrow \mathfrak{s}_-^*$ for the right translation ρ :

$$\mathbf{J}_R(\alpha_g, v, a) = T_{(e,0)}^* R_{(g,v)}(\alpha_g, v, a) = (T_e^* L_g \alpha_g, \sigma(g)^* a); \quad (7.9)$$

- the $\pm$ Lie-Poisson bracket of $F, H : \mathfrak{s}^* \rightarrow \mathbb{R}$:

$$\begin{aligned} \{F, H\}_{\pm}(\mu, a) &= \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right] \right\rangle \pm \left\langle a, \sigma' \left(\frac{\delta F}{\delta \mu} \right) \frac{\delta H}{\delta a} \right\rangle \\ &\mp \left\langle a, \sigma' \left(\frac{\delta H}{\delta \mu} \right) \frac{\delta F}{\delta a} \right\rangle \text{ for } \mu \in \mathfrak{g}^*, a \in V^*; \end{aligned} \quad (7.10)$$

- the Hamiltonian vector field determined by $H : \mathfrak{s}^* \rightarrow \mathbb{R}$:

$$X_H(\mu, a) = \mp \left(\text{ad}_{\frac{\delta H}{\delta \mu}}^* \mu - \left(\sigma'_{\frac{\delta H}{\delta a}} \right)^* a, \sigma' \left(\frac{\delta H}{\delta \mu} \right)^* a \right). \quad (7.11)$$

Now we shall reduce in two steps. We start with the left action of S on T^*S . As we already know from general theory, the momentum map $\mathbf{J}_R$ is invariant under λ . The normal subgroup V of S acts on S by left translations and the lift of this action admits an equivariant momentum map (in this case invariant since V is Abelian), given by the second component $(\alpha_g, v, a) \mapsto a$ of $\mathbf{J}_L$. In addition, the projection $T^*S \rightarrow T^*G$ is clearly canonical so that the map

$$P_L : (\alpha_g, v, a) \in T^*S \mapsto (\alpha_g, a) \in T^*G \times V^*$$

is also canonical; $T^*G \times V^*$ has the product Poisson structure (see §5.3, example (6)). It is easy to see that $\mathbf{J}_R$ factors through P_L , that is, there is a smooth map

$$\tilde{\mathbf{J}}_R : (\alpha_g, a) \in T^*G \times V^* \mapsto (T_e^* L_g \alpha_g, \sigma(g)^* a) \in \mathfrak{s}_-^*$$

such that the following diagram is commutative:

$$\begin{array}{ccc}
& T^*S & \\
P_L \swarrow & & \searrow J_R \\
T^*G \times V^* & \xrightarrow{\tilde{J}_R} & \mathfrak{s}_-^*
\end{array}$$

Since $J_R = \tilde{J}_R \circ P_L$, all three maps are canonical, and P_L is onto, it follows that $\tilde{J}_R$ is also canonical.

The same phenomenon occurs when working with the right action ρ . Since there is a lot of asymmetry in the expression of all the maps involved, we shall repeat the argument. The momentum map J_L is right invariant. The normal subgroup V of S acts on the right on T^*S with momentum map $(\alpha_g, v, a) \mapsto \sigma(g)^*a$ given by the second component of J_R . This map is therefore canonical. Moreover, the map

$$(\alpha_g, u, a) \mapsto \alpha_g + \mathbf{d}f_{\sigma(g^{-1})u}^a(g) = \alpha_g + T_g^* R_{g^{-1}}(\sigma'_u)^* a$$

is a projection followed by a translation with an exact differential on the fibers and is hence a canonical map from T^*S to T^*G (see Proposition 2.15). Therefore

$$P_R : (\alpha_g, u, a) \in T^*S \mapsto (\alpha_g + T_g^* R_{g^{-1}}(\sigma'_u)^* a, \sigma(g)^* a) \in T^*G \times V^*$$

is a canonical map. Now notice that J_L factors through P_R , that is, there is a smooth map

$$\tilde{J}_L : (\alpha_g, a) \in T^*G \times V^* \mapsto (T_e^* R_g \alpha_g, \sigma(g^{-1})^* a) \in \mathfrak{s}_+^*$$

such that the following diagram is commutative:

$$\begin{array}{ccc}
& T^*S & \\
P_R \swarrow & & \searrow J_L \\
T^*G \times V^* & \xrightarrow{\tilde{J}_L} & \mathfrak{s}_+^*
\end{array}$$

As before, this implies that $\tilde{\mathbf{J}}_L$ is a canonical map.

The origin of the maps P_L and P_R is also transparent. The space $T^*G \times V^*$ is diffeomorphic to the orbit space of T^*S by the left or right V -action. The diffeomorphisms that implement this identification are easily seen to be

$$[\alpha_g, u, a] \mapsto (\alpha_g, a)$$

for the left V -action and

$$[\alpha_g, u, a] \mapsto (\alpha_g + \mathbf{d}f_{\sigma(g^{-1})u}^a(g), \sigma(g)^*a)$$

for the right V -action, where $[\alpha_g, u, a]$ denotes the left or right V -orbit through (α_g, u, a) . Using these diffeomorphisms, the projections onto the orbit spaces become P_L and P_R respectively.

We summarize these considerations in the following theorem.

Theorem 7.6 *The maps $\tilde{\mathbf{J}}_L, \tilde{\mathbf{J}}_R : T^*G \times V^* \rightarrow \mathfrak{s}_\pm^*$ given by*

$$\begin{aligned}\tilde{\mathbf{J}}_L(\alpha_g, a) &= (T_e^* R_g \alpha_g, \sigma(g^{-1})^* a) \\ \tilde{\mathbf{J}}_R(\alpha_g, a) &= ((T_e^* L_g \alpha_g, \sigma(g)^* a)\end{aligned}$$

*are canonical. These maps are reductions of momentum maps by the action of the normal subgroup V and are themselves momentum maps for the left, respectively right, actions of S on the product Poisson manifold $T^*G \times V^*$, where V^* carries the trivial Poisson bracket.*

The procedure used here to reduce in two steps is very general and can be applied to many other situations, such as central extensions of groups, for example. We refer to [Marsden et. al.] and references therein for the general theory of reduction by stages and many other examples.

Let us study the reduction of dynamics implied by this theorem. So, consider a Hamiltonian $H : T^*G \times V^* \rightarrow \mathbb{R}$ and assume that it is invariant under the left action of S on $T^*G \times V^*$. In particular, for each $a \in V^*$ the function $H_a : T^*G \rightarrow \mathbb{R}$ given by $H_a(\alpha_g) := H(\alpha_g, a)$ is invariant under the lift to T^*G of the left action of the stabilizer $G_a := \{g \in G \mid \sigma(g)^*a = a\}$ on G . Then it follows that H induces a smooth function $H_L : \mathfrak{s}_-^* \rightarrow \mathbb{R}$ defined by $H_L \circ \tilde{\mathbf{J}}_R = H$, that is, $H_L(T_e^* L_g \alpha_g, \sigma(g)^*a) = H(\alpha_g, a)$. For right invariant systems, one interchanges, as usual, “left” by “right” and “−” by “+”. However, in this case, because the maps involved are different we record H_R separately: $H_R \circ \tilde{\mathbf{J}}_L = H$, that is, $H_R(T_e^* R_g \alpha_g, \sigma(g^{-1})^*a) = H(\alpha_g, a)$.

It turns out that the evolution of $a \in V^*$ is particularly simple. We begin with the left action and work on $\mathfrak{s}_-^*$. Let $c_a(t) \in T^*G$ denote an integral

curve of the Hamiltonian system for H_a and let $g_a(t)$ be its projection on G . Then $t \mapsto (c_a(t), a)$ is an integral curve of H on $T^*G \times V^*$ so that the curve $t \mapsto \tilde{\mathbf{J}}_R(c_a(t), a)$ is an integral curve of H_L on $\mathfrak{s}_-^*$. Thus, $t \mapsto \sigma(g_a(t))^*a$ is the evolution of the initial condition $a \in V^*$ in $\mathfrak{s}_-^*$.

For right actions the situation is identical, but we shall find another formula. If $c_a(t)$ and $g_a(t)$ are as before, the curve $t \mapsto \tilde{\mathbf{J}}_L(c_a(t), a)$ is an integral curve of H_R on $\mathfrak{s}_+^*$. Hence $t \mapsto \sigma(g_a(t)^{-1})^*a$ is the evolution of $a \in V^*$ in $\mathfrak{s}_+^*$. This proves the following theorem.

Theorem 7.7 *Let $H : T^*G \times V^* \rightarrow \mathbb{R}$ be a left invariant function relative to the S -action on $T^*G \times V^*$. Then H induces a Hamiltonian $H_L : \mathfrak{s}_-^* \rightarrow \mathbb{R}$ defined by $H_L(T_e^*L_g\alpha_g, \sigma(g)^*a) = H(\alpha_g, a)$ which then yields Lie-Poisson equations on $\mathfrak{s}_-^*$. The curve $(c_a(t), a) \in T^*G \times V^*$ is a solution of Hamilton's equations defined by H on the product Poisson manifold $T^*G \times V^*$, where V^* is endowed with the trivial Poisson bracket, if and only if $\tilde{\mathbf{J}}_R(c_a(t), a)$ is a solution of the Lie-Poisson system on $\mathfrak{s}_-^*$ defined by H_L . In particular, the evolution of $a \in V^*$ is given by $\sigma(g_a(t))^*a$, where $g_a(t)$ is the projection of $c_a(t)$ on G . For right invariant systems one interchanges “left” by “right”, “−” by “+”, and defines $H_R : \mathfrak{s}_+^* \rightarrow \mathbb{R}$ by $H_R(T_e^*R_g\alpha_g, \sigma(g^{-1})^*a) = H(\alpha_g, a)$. In this case, the evolution of $a \in V^*$ is given by $\sigma(g_a(t)^{-1})^*a$.*

The combination of these two theorems is quite powerful in examples. Often, a physical system is given by a Hamiltonian on $T^*G \times V^*$, where V^* is usually a space of parameters of the system. This Hamiltonian is left or right invariant under the $G \otimes V$ -action on $T^*G \times V^*$. Then, the theorems just proved, guarantee that one can reduce the given system to Lie-Poisson equations on $(\mathfrak{g} \otimes V)^*$ and one knows already that the second equation has as solution the “dragging along by the action” of the initial condition. For systems in continuum mechanics, this appears usually as a “Lie transport” equation, such as the conservation of mass, of entropy, or the frozen magnetic lines in the fluid in the magneto-hydrodynamics approximation.

We close these considerations by presenting a symplectic counterpart of Theorem 7.6. We shall use the Symplectic Orbit Reduction Theorem 7.4 to determine, up to connected components, the symplectic leaves of $(T^*G)/G_a$ for any $a \in V^*$. Fix in all that follows an $a \in V^*$ and let $\mathfrak{g}_a := \{\xi \in \mathfrak{g} \mid \sigma'(\xi)^*a = 0\}$ be the Lie algebra of G_a . The lift to T^*G of left translation of G_a on G has the equivariant momentum map $\mathbf{J}_L^a : T^*G \rightarrow \mathfrak{g}_a^*$ given by restriction $\mathbf{J}_L^a(\alpha_g) = (T_e^*R_g\alpha_g)|_{\mathfrak{g}_a}$. The map $i_L^a : T^*G \rightarrow T^*S$ given by $i_L^a(\alpha_g) := (\alpha_g, 0, a)$ is a Poisson embedding which is equivariant relative to the left action of G_a on T^*G and the lifted left action λ of S on its cotan-

gent bundle T^*S . Therefore i_L^a induces a Poisson embedding on the quotients $\bar{i}_L^a : (T^*G)/G_a \rightarrow T^*S/S \cong \mathfrak{s}_-^*$. From the Symplectic Orbit Reduction Theorem we know that

$$\begin{aligned} (\mathbf{J}_L^a)^{-1}(\mu|_{\mathfrak{g}_a})/(G_a)_{\mu|_{\mathfrak{g}_a}} &\cong (\mathbf{J}_L^a)^{-1}(\mathcal{O}_{\mu|_{\mathfrak{g}_a}})/G_a \hookrightarrow (T^*G)/G_a \\ &\xrightarrow{\bar{i}_L^a} (T^*S)/S \xrightarrow{\bar{\mathbf{J}}_R} \mathfrak{s}_-^* \end{aligned}$$

where the first diffeomorphism is symplectic and given by the Orbit Symplectic Reduction Theorem and $\bar{\mathbf{J}}_R$ is the quotient of $\mathbf{J}_R : T^*S \rightarrow \mathfrak{s}_-^*$ implementing the Lie-Poisson reduction theorem (see §6.1). Where does the reduced space $(\mathbf{J}_L^a)^{-1}(\mathcal{O}_{\mu|_{\mathfrak{g}_a}})/G_a$ land by this sequence of symplectic and Poisson diffeomorphisms and embeddings? To see this, we compute

$$\begin{aligned} (\bar{\mathbf{J}}_R \circ \bar{i}_L^a) ((\mathbf{J}_L^a)^{-1}(\mathcal{O}_{\mu|_{\mathfrak{g}_a}})/G_a) &= (\mathbf{J}_R \circ i_L^a) ((\mathbf{J}_L^a)^{-1}(\mathcal{O}_{\mu|_{\mathfrak{g}_a}})) \\ &= \{(\nu, b) \in \mathfrak{s}^* \mid \text{there exists } g \in G \text{ such that } \sigma_*(g)a = b, \text{Ad}_g^* \nu \in \mathcal{O}_{\mu|_{\mathfrak{g}_a}}\} \\ &= \bigcup_{\chi|_{\mathfrak{g}_a} = \mu|_{\mathfrak{g}_a}} S \cdot (\chi, a), \end{aligned}$$

where $S \cdot (\chi, a)$ denotes the S -coadjoint orbit through (χ, a) in $\mathfrak{s}_-^*$. However, the identity

$$\{(\sigma'_u)^* a \mid u \in V\} = \{\nu \in \mathfrak{g}^* \mid \nu|_{\mathfrak{g}_a} = 0\}$$

shows that $S \cdot (\chi, a) = S \cdot (\mu, a)$ for all $\chi \in \mathfrak{g}^*$ satisfying $\chi|_{\mathfrak{g}_a} = \mu|_{\mathfrak{g}_a}$. Therefore the union above is actually one single orbit, namely $S \cdot (\mu, a)$, and we have shown that the reduced space $(\mathbf{J}_L^a)^{-1}(\mathcal{O}_{\mu|_{\mathfrak{g}_a}})/G_a$ lands in $S \cdot (\mu, a)$.

For right actions the same thing happens but we need the map $i_R^a : T^*G \rightarrow T^*S$ given by $i_R^a(\alpha_g) := (\alpha_g, 0, \sigma_*(g)a)$ to embed right (G_a, S) -equivariantly T^*G into T^*S . We shall also need another notation for the quotients relative to right actions and we shall adopt here $S \backslash T^*S$ and $G_a \backslash T^*G$. Similarly, a sign on a coadjoint orbit signifies the sign in front of the orbit symplectic structure. We have proved the following theorem.

Theorem 7.8 *The map $\bar{\mathbf{J}}_R \circ \bar{i}_L^a : (\mathbf{J}_L^a)^{-1}(\mathcal{O}_{\mu|_{\mathfrak{g}_a}})/G_a \rightarrow S \cdot (\mu, a)_-$ is a symplectic diffeomorphism thereby realizing this reduced space as a coadjoint orbit in $\mathfrak{s}_-^*$. The map $\bar{\mathbf{J}}_L \circ \bar{i}_R^a : G_a \backslash (\mathbf{J}_R^a)^{-1}(\mathcal{O}_{\mu|_{\mathfrak{g}_a}}) \rightarrow S \cdot (\mu, a)_+$ is a symplectic diffeomorphism thereby realizing this reduced space as a coadjoint orbit in $\mathfrak{s}_+^*$.*

In other words, forgetting about the precise maps involved and the orbit reduction formulation of this result, this theorem states that *there is a symplectic diffeomorphism between the coadjoint orbit $S \cdot (\mu, a) \subset (\mathfrak{g} \otimes V)^*$ and*

the reduced space obtained by reducing T^*G by the subgroup G_a at the point $\mu|_{\mathfrak{g}_a} \in \mathfrak{g}_a^*$.

There is a Lagrangian version of this theorem, that is, a formulation in terms of Euler-Poincaré type equations. It is *not true* that the Euler-Poincaré equations that we shall deduce below for $\mathfrak{g} \ltimes V$ are simply the general Euler-Poincaré equations explicitly written out for a semidirect product. The reduced Lagrangian formulation in the case of semidirect products is more subtle and was done in [HMR98]. We present only the situation of left representations and left invariant Lagrangians. There are clearly three other versions and, unfortunately, they are important because of various relative sign differences in the equations and the constrained variations. Since in these lectures we shall only deal with the heavy top, we refer to the above mentioned paper for additional details and examples.

The set-up of the problem is the following. Given are:

- a left representation $\sigma : G \rightarrow \text{Aut}(V)$ of a Lie group G on a vector space V which induces the left action of G on $TG \times V^*$ given by $h \cdot (v_g, a) := (T_g L_h(v_g), \sigma_*(g)a)$, for $v_g \in T_g G$ and $a \in V^*$;
- a smooth left invariant function $L : TG \times V^* \rightarrow \mathbb{R}$ relative to this action;
- in particular, if $a_0 \in V^*$ the function $L_{a_0} : TG \rightarrow \mathbb{R}$ given by $L_{a_0}(v_g) := L(v_g, a_0)$ is invariant under the lift to TG of left translation of G_{a_0} on G ;
- by left G -invariance of L the formula

$$l(T_g L_{g^{-1}} v_g, \sigma(g)^* a) = L(v_g, a)$$

defines a smooth function $l : \mathfrak{g} \times V^* \rightarrow \mathbb{R}$ and conversely any such function $l : \mathfrak{g} \times V^* \rightarrow \mathbb{R}$ determines a left invariant function $L : TG \times V^* \rightarrow \mathbb{R}$;

- for a curve $g(t) \in G$ with $g(0) = e$, let $\xi(t) := T_{g(t)} L_{g(t)^{-1}} \dot{g}(t) \in \mathfrak{g}$ and define the curve $a(t) \in V^*$ by

$$a(t) := \sigma(g(t))^* a_0 \tag{7.12}$$

for some given $a_0 \in V^*$; the unique solution of the linear differential equation with time dependent coefficients

$$\dot{a}(t) = \sigma'(\xi(t))^* a(t) \tag{7.13}$$

with initial condition a_0 is this curve $a(t)$.

With these notations we have the following.

Theorem 7.9 *The following statements are equivalent:*

(i) With $a_0 \in V^*$ fixed, Hamilton's variational principle

$$\delta \int_{t_1}^{t_2} L_{a_0}(g(t), \dot{g}(t)) dt = 0$$

holds, for variations $\delta g(t)$ of $g(t)$ vanishing at the endpoints.

- (ii) The curve $g(t)$ satisfies the Euler-Lagrange equations for L_{a_0} on G .
 (iii) The constrained variational principle

$$\delta \int_{t_1}^{t_2} l(\xi(t), a(t)) dt = 0$$

holds on $\mathfrak{g} \times V^*$, using variations of the form

$$\delta \xi = \dot{\eta} + [\xi, \eta], \quad \delta a = \sigma'(\eta)^* a,$$

where $\eta(t) \in \mathfrak{g}$ is any curve vanishing at the endpoints.

(iv) The **semidirect Euler-Poincaré equations**

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} + \left(\sigma'_{\frac{\delta l}{\delta a}} \right)^* a \quad (7.14)$$

hold on $\mathfrak{g} \times V^*$.

Proof The proof follows the same pattern as that of Theorem 6.6. The equivalence of (i) and (ii) is Hamilton's classical variational principle that holds for any manifold (see Theorem 3.5). To prove that (iii) and (iv) are equivalent, we compute the variation of l , integrate by parts, and use the conditions $\eta(t_1) = \eta(t_2) = 0$ to get

$$\begin{aligned} \delta \int_{t_1}^{t_2} l(\xi(t), a(t)) dt &= \int_{t_1}^{t_2} \left(\left\langle \frac{\delta l}{\delta \xi}, \delta \xi \right\rangle + \left\langle \delta a, \frac{\delta l}{\delta a} \right\rangle \right) dt \\ &= \int_{t_1}^{t_2} \left(\left\langle \frac{\delta l}{\delta \xi}, \dot{\eta} + [\xi, \eta] \right\rangle + \left\langle \sigma'(\eta)^* a, \frac{\delta l}{\delta a} \right\rangle \right) dt \\ &= \int_{t_1}^{t_2} \left(\left\langle -\frac{d}{dt} \frac{\delta l}{\delta \xi} + \text{ad}_\xi^* \frac{\delta l}{\delta \xi}, \eta \right\rangle + \left\langle \left(\sigma'_{\frac{\delta l}{\delta a}} \right)^* a, \eta \right\rangle \right) dt \\ &= \int_{t_1}^{t_2} \left\langle -\frac{d}{dt} \frac{\delta l}{\delta \xi} + \text{ad}_\xi^* \frac{\delta l}{\delta \xi} + \left(\sigma'_{\frac{\delta l}{\delta a}} \right)^* a, \eta \right\rangle dt. \end{aligned}$$

Since this is valid for any smooth path $\eta(t)$ vanishing at the endpoints, the variation of the integral of l vanishes subject to the constrained variations of ξ and a if and only if the semidirect Euler-Poincaré equations hold.

It remains to be shown that (i) and (iii) are equivalent. We begin by noticing that due to the G -invariance of L and the relation $a(t) = \sigma'(g(t))^* a_0$ the integrands in the two variational principles are equal. Now let, $\eta(t) :=$

$T_{g(t)}L_{g(t)^{-1}}\delta g(t) \in \mathfrak{g}$. At this point one could proceed with the proof exactly as was done in the one for Theorem 6.6 by assuming that we work only with matrix groups, or use Lemma 6.7 to do the general case. Let us work abstractly this time around. So, by Lemma 6.7, all variations $\delta g(t) \in TG$ of $g(t)$ with fixed endpoints induce and are induced by variations $\delta \xi(t) \in \mathfrak{g}$ of $\xi(t)$ of the form $\delta \xi = \dot{\eta} + [\xi, \eta]$ with $\eta(t)$ a smooth curve vanishing at the endpoints.

Thus if (i) holds, define $\eta(t) := T_{g(t)}L_{g(t)^{-1}}\delta g(t) \in \mathfrak{g}$ for a variation $\delta g(t)$ vanishing at the endpoints and set $\delta \xi(t) = T_{g(t)}\dot{g}(t)$. By Lemma 6.7 we have $\delta \xi = \dot{\eta} + [\xi, \eta]$ with $\eta(t)$ a smooth curve vanishing at the endpoints. In addition, the variation of $a(t) = \sigma'(g(t))^*a_0$ is $\delta a(t) = \sigma'(\eta(t))a(t)$. Thus (iii) holds.

Conversely, assume that (iii) holds. So if $\delta \xi = \dot{\eta} + [\xi, \eta]$ with $\eta(t)$ a smooth curve vanishing at the endpoints, define $\delta g(t) = T_e L_{g(t)}\eta(t) \in TG$. Lemma 6.7 guarantees then that this $\delta g(t)$ is the general variation of $g(t)$ vanishing at the endpoints. Finally, the relation $\delta a(t) = \sigma'(\eta(t))^*a(t)$ shows that the variation of $\sigma_*(g(t))a(t) = a_0$ vanishes, which is consistent with the fact that L_{a_0} depends only on $g(t)$ and $\dot{g}(t)$. Thus (i) holds. ■

We close this section by showing how the heavy top equations fit into the semidirect Lie-Poisson and Euler-Poincaré framework. In the process, many of the remarkable statements in §1.2 that appeared as computational coincidences will be explained through the theory that was just presented in this section. To do this, we shall use all the explicit formulas deduced in §1.2. The configuration space is $G = SO(3)$ and it represents the attitude of the heavy top. In coordinates it is given by Euler angles, as explained in §1.2. The parameter of the problem is $Mg\ell\chi$, where $M \in \mathbb{R}$ is the mass of the heavy top, $g \in \mathbb{R}$ is the value of the gravitational acceleration, $\ell \in \mathbb{R}$ is the distance from the fixed point (that is, the point of suspension of the rigid body) to the center of mass of the body, and $\chi \in \mathbb{R}^3$ is the unit vector pointing from the fixed point to the center of mass. Therefore, the parameter space of this problem is $V^* := \mathbb{R}^3$. Identifying $\mathbb{R}^3$ with itself using the usual inner product, gives $V := \mathbb{R}^3$. The representation $\sigma : SO(3) \rightarrow \text{Aut}(\mathbb{R}^3)$ is usual matrix multiplication on vectors, that is, $\sigma(A)\mathbf{v} := A\mathbf{v}$, for any $A \in SO(3)$ and $\mathbf{v} \in \mathbb{R}^3$. Dualizing we get $\sigma(A)^*\Gamma = A^*\Gamma = A^{-1}\Gamma$, for any $\Gamma \in V^* \cong \mathbb{R}^3$. The induced Lie algebra representation $\sigma' : \mathbb{R}^3 \cong \mathfrak{so}(3) \rightarrow \text{End}(\mathbb{R}^3)$ is given by $\sigma'(\Omega)\mathbf{v} = \sigma'_\mathbf{v}\Omega = \Omega \times \mathbf{v}$, for any $\Omega, \mathbf{v} \in \mathbb{R}^3$. Therefore, $(\sigma'_\mathbf{v})^*\Gamma = \mathbf{v} \times \Gamma$ and $\sigma'(\Omega)^*\Gamma = \Gamma \times \Omega$, for any $\mathbf{v} \in V \cong \mathbb{R}^3$, $\Omega \in \mathbb{R}^3 \cong \mathfrak{so}(3)$, and $\Gamma \in V^* \cong \mathbb{R}^3$. Recall also that $\text{ad}_\Omega^*\Pi = \Pi \times \Omega$ by using the isomorphism (1.14); see (1.21).

The expressions of the Hamiltonian and Lagrangian functions on $\mathfrak{se}(3)^* \cong$

$\mathbb{R}^3 \times \mathbb{R}^3$ and $\mathfrak{se}(3) \cong \mathbb{R}^3 \times \mathbb{R}^3$ respectively (see (1.42) and (1.46)),

$$\begin{aligned} H(\mathbf{\Pi}, \mathbf{\Gamma}) &= \frac{1}{2} \mathbf{\Pi} \cdot \mathbb{I}^{-1} \mathbf{\Pi} + Mg\ell \mathbf{\Gamma} \cdot \boldsymbol{\chi} \\ L(\mathbf{\Omega}, \mathbf{\Gamma}) &= \frac{1}{2} \mathbb{I} \mathbf{\Omega} \cdot \mathbf{\Omega} - Mg\ell \mathbf{\Gamma} \cdot \boldsymbol{\chi} \end{aligned}$$

yield $\delta H / \delta \mathbf{\Pi} = \mathbb{I}^{-1} \mathbf{\Pi} = \mathbf{\Omega}$, $\delta H / \delta \mathbf{\Gamma} = Mg\ell \boldsymbol{\chi}$, $\delta L / \delta \mathbf{\Omega} = \mathbb{I} \mathbf{\Omega} = \mathbf{\Pi}$, and $\delta L / \delta \mathbf{\Gamma} = -Mg\ell \boldsymbol{\chi}$, where, in this case, due to the dot product pairing, the partial functional derivatives are given by $\delta / \delta \mathbf{\Pi} = \nabla_{\mathbf{\Pi}}$ and $\delta / \delta \mathbf{\Gamma} = \nabla_{\mathbf{\Gamma}}$. So we can immediately write both the semidirect product Lie-Poisson (7.11) and Euler-Poincaré equations (7.14), (7.13) to get

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbf{\Omega} + Mg\ell \mathbf{\Gamma} \times \boldsymbol{\chi}, \quad \dot{\mathbf{\Gamma}} = \mathbf{\Gamma} \times \mathbf{\Omega}$$

which are the Euler-Poisson equations (1.44). The solution (7.12) of the second Euler-Poisson equation with initial condition $\mathbf{\Gamma}(0) = \mathbf{k}$ is $\mathbf{\Gamma} = A^{-1} \mathbf{k}$ which was the definition of $\mathbf{\Gamma}$ used in the expression (1.39) of the potential energy. The minus Lie-Poisson bracket (7.10) becomes in this case

$$\{F, H\}(\mathbf{\Pi}, \mathbf{\Gamma}) = -\mathbf{\Pi} \cdot (\nabla_{\mathbf{\Pi}} F \times \nabla_{\mathbf{\Pi}} H) - \mathbf{\Gamma} \cdot (\nabla_{\mathbf{\Pi}} F \times \nabla_{\mathbf{\Gamma}} H + \nabla_{\mathbf{\Gamma}} F \times \nabla_{\mathbf{\Pi}} H),$$

which is formula (1.45). One recognizes in the computations at the end of §1.2 part of the proof of Theorem 7.9. The remarkable map (1.43) that sends the Euler angles and their conjugate momenta to the variables $(\mathbf{\Pi}, \mathbf{\Gamma})$ is none other than the momentum map $\mathbf{J}_R$ (see (7.9)) expressed in the chart given by the Euler angles. So, of course, it will map Hamilton's equations on $T^*SO(3)$ to minus Lie-Poisson equations on $\mathfrak{se}(3)^*$, as the general theory presented in these lectures stipulates.

Since the Euler-Poisson equations are of Lie-Poisson type, their solutions must lie on coadjoint orbits. The generic ones are given by the surfaces defined by $\|\mathbf{\Gamma}\| = \text{constant}$ and $\mathbf{\Pi} \cdot \mathbf{\Gamma} = \text{constant}$. Any function of $\|\mathbf{\Gamma}\|$ and $\mathbf{\Pi} \cdot \mathbf{\Gamma}$ is a Casimir function for the Lie-Poisson bracket, as an easy verification shows. In particular, the Euler-Poisson equations always have these two functions as conserved quantities. Restricted to such a generic coadjoint orbit, the Euler-Poisson equations are Hamiltonian relative to the orbit symplectic form and have the total energy H conserved. To be completely integrable, one needs therefore one more conserved quantity, independent of H and commuting with it. It is known that this is possible only in three cases: the **Euler case** characterized by $\ell = 0$, that is, the center of mass coincides with the point of suspension of the rigid body, or equivalently, no forces act on the body and one has fixed its center of mass, the **Lagrange case** characterized by $\boldsymbol{\chi} = (0, 0, 1)$ and $I_1 = I_2$, that is, the body has an additional S^1 -symmetry around the line

connecting the point of suspension of the body with its center of mass, and the **Kowalewski case**, characterized by the conditions $I_1 = I_2 = 2I_3$ and the center of mass lies in the plane of the equal moments of inertia, so it can be assumed to be $\chi = (1, 0, 0)$ by simply adjusting the frame of reference. The last two cases have an additional integral. In the Lagrange top case, this is the momentum map of the S^1 -action. In the Kowalewski case, the origin of this additional integral remains to this day a mystery from the point of view of momentum maps, that is, it is not known a priori (that is, without solving explicitly the system) what one-dimensional Lie group action has as its momentum map this additional quartic integral.

To completely describe the kinematics of this system we shall record here the coadjoint orbits of the special Euclidean group $SE(3)$; for the proofs of all the formulas below see [MaRa94], §14.7. Let $\{e_1, e_2, e_3, f_1, f_2, f_3\}$ be an orthonormal basis of $\mathfrak{se}(3) \cong \mathbb{R}^3 \times \mathbb{R}^3$ such that $e_i = f_i, i = 1, 2, 3$. The dual basis of $\mathfrak{se}(3)^*$ via the dot product is again $\{e_1, e_2, e_3, f_1, f_2, f_3\}$. There is a single zero dimensional coadjoint orbit, namely the origin. The other orbits are two and four dimensional. There is no six dimensional coadjoint orbit since the Poisson bracket is degenerate having the two Casimir functions given above. There are three types of coadjoint orbits.

Type I: The orbit $\mathcal{O}$ through $(e, 0)$ equals

$$SE(3) \cdot (e, 0) = \{ (Ae, 0) \mid A \in SO(3) \} = S_{\|e\|}^2 \times \{0\}, \quad (7.15)$$

the two-sphere of radius $\|e\|$. The tangent space to $\mathcal{O}$ at $(e, 0)$ is the tangent space to the sphere of radius $\|e\|$ at the point e in the first factor. The minus orbit symplectic form is

$$\omega^-(e, 0)(\text{ad}_{(x,y)}^*(e, 0), \text{ad}_{(x',y')}^*(e, 0)) = -e \cdot (x \times x')$$

which equals $-1/\|e\|$ times the area element of the sphere of radius $\|e\|$ (see §6.3, example 1).

Type II: The orbit $\mathcal{O}$ through $(0, f)$ is given by

$$\begin{aligned} SE(3) \cdot (0, f) &= \{ (a \times Af, Af) \mid A \in SO(3), a \in \mathbb{R}^3 \} \\ &= \{ (u, Af) \mid A \in SO(3), u \perp Af \} = TS_{\|f\|}^2, \end{aligned} \quad (7.16)$$

the tangent bundle of the two-sphere of radius $\|f\|$; note that the vector part is the first component. The tangent space to $\mathcal{O}$ at $(0, f)$ equals $f^\perp \times f^\perp$, where $f^\perp$ denotes the plane perpendicular to f . Let $(u, v) \in \mathcal{O}$, that is, $\|v\| = \|f\|$

and $\mathbf{u} \perp \mathbf{v}$. The symplectic form in this case is

$$\begin{aligned} \omega^-(\mathbf{u}, \mathbf{v})(\text{ad}_{(\mathbf{x}, \mathbf{y})}^*(\mathbf{u}, \mathbf{v}), \text{ad}_{(\mathbf{x}', \mathbf{y}')}^*(\mathbf{u}, \mathbf{v})) \\ = -\mathbf{u} \cdot (\mathbf{x} \times \mathbf{x}') - \mathbf{v} \cdot (\mathbf{x} \times \mathbf{y}' - \mathbf{x}' \times \mathbf{y}). \end{aligned} \quad (7.17)$$

It can be shown that this form is exact, namely, $\omega^- = -d\theta$, where

$$\theta(\mathbf{u}, \mathbf{v})(\text{ad}_{(\mathbf{x}, \mathbf{y})}^*(\mathbf{u}, \mathbf{v})) = \mathbf{u} \cdot \mathbf{x}.$$

Thus $\mathcal{O}$ is symplectically diffeomorphic to T^*S^2 endowed with the canonical cotangent bundle symplectic structure (we identify T^*S^2 with TS^2 using the natural Riemannian metric on the sphere S^2).

Type III: The orbit $\mathcal{O}$ through $(\mathbf{e}, \mathbf{f})$, where $\mathbf{e} \neq \mathbf{0}, \mathbf{f} \neq \mathbf{0}$, equals

$$\text{SE}(3) \cdot (\mathbf{e}, \mathbf{f}) = \{ (\mathbf{A}\mathbf{e} + \mathbf{a} \times \mathbf{A}\mathbf{f}, \mathbf{A}\mathbf{f}) \mid \mathbf{A} \in \text{SO}(3), \mathbf{a} \in \mathbb{R}^3 \}. \quad (7.18)$$

To get a better description of this orbit, consider the smooth map

$$\varphi : (\mathbf{A}, \mathbf{a}) \in \text{SE}(3) \mapsto \left(\mathbf{A}\mathbf{e} + \mathbf{a} \times \mathbf{A}\mathbf{f} - \frac{\mathbf{e} \cdot \mathbf{f}}{\|\mathbf{f}\|^2} \mathbf{A}\mathbf{f}, \mathbf{A}\mathbf{f} \right) \in TS_{\|\mathbf{f}\|}^2,$$

which is right invariant under the isotropy group

$$\text{SE}(3)_{(\mathbf{e}, \mathbf{f})} = \{ (\mathbf{B}, \mathbf{b}) \mid \mathbf{B}\mathbf{e} + \mathbf{b} \times \mathbf{f} = \mathbf{e}, \mathbf{B}\mathbf{f} = \mathbf{f} \}$$

and induces hence a diffeomorphism $\bar{\varphi} : \text{SE}(3)/\text{SE}(3)_{(\mathbf{e}, \mathbf{f})} \rightarrow TS_{\|\mathbf{f}\|}^2$. The orbit $\mathcal{O}$ through $(\mathbf{e}, \mathbf{f})$ is diffeomorphic to $\text{SE}(3)/\text{SE}(3)_{(\mathbf{e}, \mathbf{f})}$ by the diffeomorphism

$$(\mathbf{A}, \mathbf{a}) \mapsto \text{Ad}_{(\mathbf{A}, \mathbf{a})}^*(\mathbf{e}, \mathbf{f}).$$

Composing these two maps and identifying TS^2 and T^*S^2 by the natural Riemannian metric on S^2 , we get the diffeomorphism $\Phi : \mathcal{O} \rightarrow T^*S_{\|\mathbf{f}\|}^2$ given by

$$\Phi(\text{Ad}_{(\mathbf{A}, \mathbf{a})}^*(\mathbf{e}, \mathbf{f})) = \left(\mathbf{A}\mathbf{e} + \mathbf{a} \times \mathbf{A}\mathbf{f} - \frac{\mathbf{e} \cdot \mathbf{f}}{\|\mathbf{f}\|^2} \mathbf{A}\mathbf{f}, \mathbf{A}\mathbf{f} \right).$$

Thus this orbit is also diffeomorphic to $T^*S_{\|\mathbf{f}\|}^2$. The tangent space at $(\mathbf{e}, \mathbf{f})$ to $\mathcal{O}$ is $\{(\mathbf{u}, \mathbf{v}) \mid \mathbf{u} \cdot \mathbf{f} + \mathbf{v} \cdot \mathbf{e} = 0 \text{ and } \mathbf{v} \cdot \mathbf{f} = 0\}$. If $(\bar{\mathbf{u}}, \bar{\mathbf{v}}) \in \mathcal{O}$, the orbit symplectic structure is given by formula (7.17), where $\bar{\mathbf{u}} = \mathbf{A}\mathbf{e} + \mathbf{a} \times \mathbf{A}\mathbf{f}$ and $\bar{\mathbf{v}} = \mathbf{A}\mathbf{f}$, for some $\mathbf{A} \in \text{SO}(3), \mathbf{a} \in \mathbb{R}^3$. Let

$$\mathbf{u} = \mathbf{A}\mathbf{e} + \mathbf{a} \times \mathbf{A}\mathbf{f} - \frac{\mathbf{e} \cdot \mathbf{f}}{\|\mathbf{f}\|^2} \mathbf{A}\mathbf{f} = \bar{\mathbf{u}} - \frac{\mathbf{e} \cdot \mathbf{f}}{\|\mathbf{f}\|^2} \bar{\mathbf{v}}, \quad \mathbf{v} = \mathbf{A}\mathbf{f} = \bar{\mathbf{v}},$$

be a pair of vectors $(\mathbf{u}, \mathbf{v})$ representing an element of $TS_{\|\mathbf{f}\|}^2$. Note that $\|\mathbf{v}\| = \|\mathbf{f}\|$ and $\mathbf{u} \cdot \mathbf{v} = 0$. Then a tangent vector to $TS_{\|\mathbf{f}\|}^2$ at $(\mathbf{u}, \mathbf{v})$ can be represented

as $\text{ad}_{(\mathbf{x}, \mathbf{y})}^*(\mathbf{u}, \mathbf{v}) = (\mathbf{u} \times \mathbf{x} + \mathbf{v} \times \mathbf{y}, \mathbf{v} \times \mathbf{x})$. The push-forward of the orbit symplectic form ω^- to $TS_{\|\mathbf{f}\|}^2$ is computed then to be

$$\begin{aligned} & (\Phi_*\omega^-)(\mathbf{u}, \mathbf{v})(\text{ad}_{(\mathbf{x}, \mathbf{y})}^*(\mathbf{u}, \mathbf{v}), \text{ad}_{(\mathbf{x}', \mathbf{y}')}^*(\mathbf{u}, \mathbf{v})) \\ &= -\mathbf{u} \cdot (\mathbf{x} \times \mathbf{x}') - \mathbf{v} \cdot (\mathbf{x} \times \mathbf{y}' - \mathbf{x}' \times \mathbf{y}) - \frac{\mathbf{e} \cdot \mathbf{f}}{\|\mathbf{f}\|^2} \mathbf{v} \cdot (\mathbf{x} \times \mathbf{x}'). \end{aligned}$$

A comparison with (7.17) shows that the first two terms represent the canonical cotangent bundle symplectic form on $T^*S_{\|\mathbf{f}\|}^2$. The last term is the following closed two-form on $TS_{\|\mathbf{f}\|}^2$:

$$\beta(\mathbf{u}, \mathbf{v}) \left(\text{ad}_{(\mathbf{x}, \mathbf{y})}^*(\mathbf{u}, \mathbf{v}), \text{ad}_{(\mathbf{x}', \mathbf{y}')}^*(\mathbf{u}, \mathbf{v}) \right) = -\frac{\mathbf{e} \cdot \mathbf{f}}{\|\mathbf{f}\|^2} \mathbf{v} \cdot (\mathbf{x} \times \mathbf{x}').$$

This two-form β is a magnetic term as in §2.5. Therefore $\mathcal{O}$ is the cotangent bundle of the two-sphere of radius $\|\mathbf{f}\|$ endowed with a magnetic symplectic form. The type II and III coadjoint orbits are diffeomorphic but not symplectomorphic.

The motion of the heavy top always lies on these coadjoint orbits and is a Hamiltonian system relative to the total energy H and the orbit symplectic structures presented here.

Bibliography

- [AbMa78] Abraham, R. and Marsden, J.E. [1978], *Foundations of Mechanics*. Addison-Wesley, second edition.
- [AMR88] Abraham, R., Marsden, J.E. and Ratiu, T.S. [1988], *Manifolds, Tensor Analysis, and Applications*. Volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, second edition.
- [AmCZ96] Ambrosetti, A. and Coti Zelati, V. [1996], *Periodic Solutions of Singular Lagrangian Systems*. Birkhäuser
- [Arnold79] Arnold V.I. [1979], *Mathematical Methods of Classical Mechanics*. Volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag.
- [BelRa04] Beltiță, D. and Ratiu, T.S. [2004], Symplectic leaves in real Banach Lie-Poisson spaces, *Geom and Funct. Anal.*, to appear.
- [Ben86] Benci, V. [1986], Periodic solutions of Lagrangian Systems on a compact manifold, *Journal of Differential Equations* **63**, 135–161.
- [BlBr92] Blanchard, P. and Bruning, E. [1992], *Variational Methods in Mathematical Physics*. Springer Verlag.
- [Bo00] Bona, P. [2000], Extended quantum mechanics. *Acta Physica Slovaca*, **50** (1), 1–198.
- [Bou71] Bourbaki, N. [1971], *Variétés différentielles et analytiques. Fascicule de résultats*. Hermann.
- [Bou89] Bourbaki, N. [1989], *Lie Groups and Lie Algebras. Chapters 1–3*. Springer Verlag.
- [Cendra et al.] Cendra, H, Marsden, J.E., Pekarsky, S., and Ratiu, T.S. [2003], Variational principles for Lie-Poisson and Hamilton-Poincaré equations, *Moscow Math. Journ.*, **3**(3), 833–867.
- [ChMa74] Chernoff, P. R. and Marsden, J.E. [1974], *Properties of Infinite Dimensional Hamiltonian Systems*, Lecture Notes in Mathematics, **425**, Springer-Verlag.
- [CB97] Cushman, R.H. and Bates, L.M. [1997], *Global Aspects of Integrable Systems*. Birkhäuser.
- [DeMe93] Dellnitz M. and Melbourne I. [1993], The equivariant Darboux Theorem, *Lectures in Appl. Math.*, **29**, 163–169.

- [DrJo89] Drazin, P.G. and Johnson, R.S. [1989], *Solitons: An Introduction*. Cambridge University Press.
- [DFN95] Dubrovin, B., Novikov, S.P., and Fomenko, A.T. [1995], *Modern Geometry I, II, III*, Volumes 93, 104, 124 of *Graduate Texts in Mathematics*. Springer-Verlag.
- [EbMa70] Ebin, D.G. and Marsden, J.E. [1970], Groups of diffeomorphisms and the motion of an incompressible fluid, *Ann. of Math.* **92**, 102–163.
- [GuSt84] Guillemin, V. and Sternberg, S. [1984], *Symplectic Techniques in Physics*. Cambridge University Press.
- [HMR98] Holm, D.D., Marsden, J.E., and Ratiu, T.S. [1998], The Euler-Poincaré equations and semidirect products with applications to continuum theories, *Advances in Math.*, **137**, 1–81.
- [JoSa98] José, J.V. and Saletan, E.J. [1998], *Classical Dynamics : A Contemporary Approach*. Cambridge University Press.
- [Jost] Jost, J. [1998], *Riemannian Geometry and Geometric Analysis*. University Text. Springer-Verlag, second edition.
- [KaKoSt78] Kazhdan, D. Kostant, B., and Sternberg, S. [1978], Hamiltonian group actions and dynamical systems of Calogero type, *Comm. Pure Appl. Math.*, **31**, 481–508.
- [KhMis03] Khesin, B. and Misiolek, G. [2003], Euler equations on homogeneous spaces and Virasoro orbits, *Adv. in Math.*, **176**, 116–144.
- [LiMa87] Libermann, P. and Marle, C.-M. [1987], *Symplectic Geometry and Analytical Mechanics*. Reidel.
- [Lang] Lang, S. [1999], *Fundamentals of Differential Geometry*. Volume 191 of *Graduate Texts in Mathematics*. Springer-Verlag, New York.
- [Marle76] Marle, C.-M. [1976], Symplectic manifolds, dynamical groups, and Hamiltonian mechanics, in *Differential Geometry and Relativity*, Cahen, M. and Flato, M. eds., D. Reidel, Boston, 249–269.
- [Marsden81] Marsden, J.E. [1981], *Lectures on Geometric Methods in Mathematical Physics*. Volume 37, SIAM, Philadelphia, 1981.
- [Marsden92] Marsden, J.E. [1992], *Lectures on Mechanics*. Volume 174 of *London Mathematical Society Lecture Note Series*. Cambridge University Press.
- [Marsden et. al.] Marsden, J.E., Misiolek, G., Ortega, J.-P., Perlmutter, M., and Ratiu, T.S. [2004], *Hamiltonian Reduction by Stages*. Lecture Notes in Mathematics, to appear. Springer-Verlag.
- [MaMoRa90] Marsden, J.E., Montgomery, R., and Ratiu, T.S. [1984], Reduction, symmetry, and phases in mechanics, *Memoirs Amer. Math. Soc.*, **88**(436), 1–110.
- [MaRa94] Marsden, J.E. and Ratiu, T.S. [1994], *Introduction to Mechanics and Symmetry*. Volume 75 of *Texts in Applied Mathematics*, second printing of second edition 2003. Springer-Verlag.
- [MaRa95] Marsden, J.E. and Ratiu, T.S. [2003], *Geometric Fluid Dynamics*. Unpublished notes.
- [MaRa03] Marsden, J.E. and Ratiu, T.S. [2003], *Mechanics and Symmetry. Reduction Theory*. In preparation.
- [MaRaWe84a] Marsden, J.E., Ratiu, T.S., and Weinstein, A. [1984a], Semidirect products and reduction in mechanics, *Trans. Amer. Math. Soc.*, **281**(1), 147–177.
- [MaRaWe84b] Marsden, J.E., Ratiu, T.S., and Weinstein, A. [1984b], Reduction and

- Hamiltonian structures on duals of semidirect product Lie algebras, *Contemporary Math.*, **28**, 55–100.
- [MaWei74] Marsden, J.E. and Weinstein, A. [1983], Reduction of symplectic manifolds with symmetry, *Rep. Math. Phys.*, **5**, 121–130.
- [MaWei83] Marsden, J.E. and Weinstein, A. [1983], Coadjoint orbits, vortices and Clebsch variables for incompressible fluids, *Physica D*, **7**, 305–323.
- [MawWil989] Mawhin, J. and Willem, M. [1989], *Critical Point Theory and Hamiltonian Systems*. Volume 74 of *Applied Mathematical Sciences*. Springer-Verlag, second edition.
- [McDSal95] McDuff, D. and Salamon, D. [1995], *Introduction to Symplectic Topology*. Clarendon Press.
- [MD93] Melbourne, I. and Dellnitz, M. [1993], Normal forms for linear Hamiltonian vector fields commuting with the action of a compact Lie group, *Proc. Camb. Phil. Soc.*, **114**, 235–268.
- [Milnor] Milnor, J. [1963], *Morse Theory*. Princeton University Press.
- [Mis02] Misiulek, G. [2002], Classical solutions of the periodic Camassa-Holm equation, *Geom. Funct. Anal.*, **12**, 1080–1104.
- [OdzRa03] Odziejewicz A. and Ratiu T.S. [2003], Banach Lie-Poisson spaces and reduction, *Comm. Math. Phys.*, **243**, 1–54.
- [OdzRa004] Odziejewicz, A. and Ratiu, T.S. [2004], Extensions of Banach Lie-Poisson spaces, *J. Func. Anal.*, **217**(1), 103–125.
- [OR04] Ortega, J.-P. and Ratiu, T.S. [2004], *Momentum Maps and Hamiltonian Reduction*. Volume 222 of *Progress in Mathematics*. Birkhäuser.
- [OKh87] Ovsienko, V.Y. and Khesin, B.A. [1987], Korteweg-de Vries superequations as an Euler equation, *Funct. Anal. Appl.*, **21**, 329–331.
- [Palais68] Palais, R. [1968], *Foundations of Global Non-Linear Analysis*, W. A. Benjamin, Inc., New York-Amsterdam.
- [PR04] Perlmutter, M. and Ratiu, T.S. [2004], Gauged Poisson structures, preprint.
- [Schm87] Schmid, R. [1987], *Infinite Dimensional Hamiltonian Systems*. Bibliopolis.
- [Serre] Serre, J.-P. [1992], *Lie Algebras and Lie Groups*, Volume 1500 of *Lecture Notes in Mathematics*. Springer-Verlag.
- [Sp79] Spivak, M. [1979], *Differential Geometry*, Volume I. New printing with corrections. Publish or Perish, Inc. Houston, Texas.
- [V96] Vaisman, I. [1996], *Lectures on the Geometry of Poisson Manifolds*. Volume 118 of *Progress in Mathematics*, Birkhäuser.
- [W83] Warner, F.W. [1983], *Foundation of Differentiable Manifolds and Lie Groups*. Volume 94 of *Graduate Texts in Mathematics*, Springer-Verlag.
- [W] Weinstein, A. [1983], The local structure of Poisson manifolds, *Journ. Diff. Geom* **18**, 523–557.