Alma Mater Studiorum · Università di Bologna

**Dipartimento di Fisica e Astronomia “Augusto Righi”**
**Corso di Laurea in Fisica**

# Symplectic Geometry and Applications in Hamiltonian Mechanics

**Relatore:**
**Prof. Rita Fioresi**

**Presentata da:**
**Andrea Barbieri**

Anno Accademico 2025/2026

# Contents

**Abstract**

The goal of this thesis is to present the Hamiltonian formulation of classical mechanics in the generalized context of differentiable manifolds. We show that this formulation of mechanics not only is best expressed using symplectic geometry, but it actually emerges naturally from the basic construction of the theory. The work is structured as follows. In the first two chapters, we focus on the mathematical background needed for describing this generalization of mechanics, namely the theory of differentiable manifolds and of Lie groups and Lie algebras. The third chapter is entirely dedicated to presenting the main ideas of symplectic geometry with a particular emphasys on those relevant to physical applications. Finally, in the fourth chapter, we show how Hamiltonian mechanics naturally emerges from symplectic geometry and demonstrate how some ideas of classical mechanics can be elegantly reformulated using this formalism.

**Abstract**

L'obiettivo di questa tesi è presentare la formulazione hamiltoniana della meccanica classica nel contesto generalizzato delle varietà differenziabili. Mostriamo non solo che questa formulazione della meccanica è meglio espressa utilizzando la geometria simplettica, ma che emerge naturalmente dai fondamenti basilari della teoria. Il lavoro è suddiviso come segue. Nei primi due capitoli ci concentriamo sui concetti matematici necessari per descrivere questa generalizzazione della meccanica, ovvero le varietà differenziabili e i gruppi e le algebre di Lie. Il terzo capitolo è interamente dedicato all'esposizione delle principali nozioni di geometria simplettica, con una particolare attenzione rivolta a quelle con applicazioni fisiche. Infine, nel quarto capitolo mostriamo come la meccanica hamiltoniana emerge in maniera naturale dalla geometria simplettica e presentiamo come alcune nozioni di meccanica classica possono essere elegantemente riformulate mediante questo formalismo.

# Introduction

Classical mechanics, as presented in the works of Newton (1642-1726) and Huygens (1629-1695), heavily relied on geometry. What followed, however, was the period of "mécanique analytique". Lagrange's (1736-1813) treatise of mechanics contained no pictures, as his methods did not require geometric or mechanical reasonings but just algebraic operations (see [1] and [2]). Later, at the beginning of the 20th century, the works of Poincaré (1854-1912) and Birkhoff (1884-1944) brought geometry back as a tool to make qualitative statements for problems which were not quantitatively tractable. More generally, geometry played a crucial role in the development of physics in the 20th century, as both the theory of General Relativity and Yang-Mills field theories, which together describe all four fundamental forces, are deeply rooted in geometry.
The geometry of both General Relativity and gauge theories is Riemannian geometry; however, there is another type of geometry which is even more deeply rooted in physics: *symplectic geometry*. The first symplectic manifold was introduced by Lagrange in 1808 as a tool in the study of celestial mechanics (see [2]). For many years symplectic ideas were implicitily present in the development of mechanics; most importantly, the ideas of Lagrange influenced the works of Hamilton (1805-1865), who first applied them to optics and later used them to arrive at his formulation of mechanics. Nevertheless, the geometrization of classical mechanics proceeded slowly and symplectic ideas were sparsely used by physicists up until the 1940s (see [1]). However, in these same years, symplectic geometry came into its own as a distinct mathematical discipline and rapidly grew. By the 1960s, symplectic geometry was cast in its modern form and the geometric formulation of classical mechanics was finally complete.

In this thesis, we present a fully geometric treatment of the Hamiltonian formulation of classical mechanics using the tools of symplectic geometry. The work is structured as follows.
The first chapter introduces calculus on differentiable manifolds, which is an indispensable tool for the geometric constructions that follow.
Meanwhile, in the second chapter, we present the main elements of the theory of Lie groups and Lie algebras, which are the main tools used to describe symmetries in modern physics.

In the third chapter, we introduce symplectic geometry starting with the algebraic properties of a symplectic form on a vector space and later defining symplectic manifolds. We discuss Hamiltonian vector fields and Poisson brackets as purely geometric constructions and talk about the cotangent bundle as a prominent example of a symplectic manifold. We end the chapter by introducing the momentum map and show how conserved quantities can be exploited to reduce the degrees of freedom of a system via the Regular Symplectic Reduction Theorem.

Finally, in the fourth chapter, we show how Hamiltonian mechanics can be generalized using the tools of symplectic geometry developed previously and discuss the Hamiltonian version of Noether's Theorem. Later, we show how the symplectic formulation of Hamiltonian mechanics can be obtained from the Lagrangian formulation over a manifold via the Legendre transform. We conclude the chapter by showing some notable examples of momentum maps in physics and use them to apply regular symplectic reduction to Kepler's problem.

# Chapter 1

# Mathematical background

In this chapter we introduce the basics of calculus on differentiable manifolds. In particular, we define differentiable manifolds, smooth maps as well as vector fields, flows and differential $k$-forms. For further reading see [3], [4] and [5].

## 1.1 Differentiable manifolds

In this section we introduce the notion of a differentiable manifold and provide some examples. We assume that the reader is familiar with point-set topology. The main reference for this section is Section 5 of [4].

**Definition 1.1.1.** A topological space $M$ is *locally Euclidean of dimension* $n$ if every point $p$ in $M$ has a neighborhood $U$ such that there is a homeomorphism $\phi$ from $U$ onto an open subset of $\mathbb{R}^n$. We call the pair $(U, \phi : U \to \mathbb{R}^n)$ a *chart*, $U$ a *coordinate neighborhood* and $\phi$ a *coordinate map* or *coordinate system* of $U$. A chart $(U, \phi)$ is said to be *centered* at $p \in U$ if $\phi(p) = 0$.

Before proceeding, we recall that a topological space is said to be *second countable* if it has a countable basis while it is said to be *Hausdorff* if for every two distinct points $x, y$ there are two disjoint open sets $U, V$ such that $x \in U, y \in V$.

**Definition 1.1.2.** A *topological manifold* is a Hausdorff, second countable, locally Euclidean space. It is said to be of *dimension* $n$ if it is locally Euclidean of dimension n.

It is worth noting that for the notion of dimension of a topological manifold to be well-defined we need that, for $n \neq m$, an open subset of $\mathbb{R}^n$ cannot be homeomorphic to an open subset of $\mathbb{R}^m$. This fact is known as *invariance of dimension.* If a manifold has more than one connected component, it is possible for each connected component to have

a different dimension.
For physical applications additional structure is required.

**Definition 1.1.3.** Two charts $(U, \phi : U \to \mathbb{R}^n)$, $(V, \psi : V \to \mathbb{R}^n)$ of a topological manifold are $C^\infty$-*compatible* if the two maps

$$\phi \circ \psi^{-1} : \psi(U \cap V) \to \phi(U \cap V) \quad , \quad \psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$$

are both $C^\infty$. These maps are called *transition functions*. If $U \cap V$ is empty the maps are automatically $C^\infty$-compatible.

We simplify the notation by writing $U_{\alpha\beta}$ instead of $U_\alpha \cap U_\beta$ and so on for multiple intersections.
Since we are interested only in $C^\infty$-compatibility we will simply speak of compatible charts and omit the $C^\infty$ prefix.

**Definition 1.1.4.** A $C^\infty$ *atlas*, or simply an *atlas*, on a locally Euclidean space $M$ is a collection $\mathfrak{U} = \{(U_\alpha, \phi_\alpha)\}$ of pairwise compatible charts that cover $M$.
An atlas $\mathfrak{M}$ on a locally Euclidean space is said to be *maximal* if it is not contained in a larger atlas, meaning that if any other atlas $\mathfrak{U}$ contains $\mathfrak{M}$ then $\mathfrak{U} = \mathfrak{M}$.

It is worth noting that, while the compatibility of charts is both reflexive and symmetric, it is not transitive. To show this it's enough to consider three charts $(U_i, \phi_i)$, with $i = 1, 2, 3$, such that $(U_1, \phi_1)$ is compatible with $(U_2, \phi_2)$ and $(U_2, \phi_2)$ is compatible with $(U_3, \phi_3)$. Now, since the functions are simultaneously defined only on $U_{123}$, the composition

$$\phi_3 \circ \phi_1^{-1} = (\phi_3 \circ \phi_2^{-1}) \circ (\phi_2 \circ \phi_1^{-1})$$

is $C^\infty$ only on $\phi_1(U_{123})$, and not on $\phi_1(U_{13})$, as we would need for the charts $\phi_1$ and $\phi_3$ to be compatible.
We say that a chart $(V, \psi)$ is *compatible with an atlas* $\{(U_\alpha, \phi_\alpha)\}$ if it is compatible with all the charts $(U_\alpha, \phi_\alpha)$.

**Lemma 1.1.5.** *Let $(U_\alpha, \phi_\alpha)$ be an atlas on a locally Euclidean space. If two charts $(V, \psi)$ and $(W, \sigma)$ are both compatible with the atlas $(U_\alpha, \phi_\alpha)$, then they are compatible with each other.*

*Proof.* For the proof see [4], Section 5.2, Lemma 5.8. □

Finally, we can define the main object of this section.

**Definition 1.1.6.** A $C^\infty$ or *smooth manifold* is a topological manifold $M$ together with a maximal atlas. The maximal atlas is also called a *differentiable structure*. A smooth manifold is said to have dimension n if all its connected components have dimension n.

As it was for topological manifolds, the notion of dimension is well-defined as we shall prove later on (Corollary 1.3.7). The following proposition shows that it is not necessary to exhibit a maximal atlas to prove that a topological manifold is also a smooth manifold, as the existence of *any* atlas suffices.

**Proposition 1.1.7.** *Any atlas $\mathfrak{U} = (U_\alpha, \phi_\alpha)$ on a locally Euclidean space is contained in a unique maximal atlas $\mathfrak{M}$.*

*Proof.* Adjoin to the atlas $\mathfrak{U}$ all the charts $(V_i, \psi_i)$ that are compatible with $\mathfrak{U}$ and call this new collection of charts $\mathfrak{M}$. By Lemma 1.1.5 all the charts $(V_i, \psi_i)$ are compatible with one another. Therefore, $\mathfrak{M}$ is an atlas as well. Any chart compatible with the new atlas has to be compatible with the original atlas $\mathfrak{U}$, so it must belong to the new atlas. This proves $\mathfrak{M}$ is maximal.
Let $\mathfrak{M}'$ be another maximal atlas containing $\mathfrak{U}$, then all the charts in $\mathfrak{M}'$ are compatible with $\mathfrak{U}$ so, by construction, they must belong to $\mathfrak{M}$. This shows that $\mathfrak{M}' \subset \mathfrak{M}$, since $\mathfrak{M}'$ is maximal we get that $\mathfrak{M}' = \mathfrak{M}$. This shows that the maximal atlas is unique. □

From this proposition we conclude that to prove that a given topological space is a smooth manifold it's sufficient to show that it is Hausdorff, second countable and that it has *an* atlas, not necessarily a maximal one. It is worth noting that, while any given atlas is contained in unique maximal atlas, a given topological manifold could admit distinct differentiable structures.
From now on, since we are only interested in smooth manifolds, we will simply speak of manifolds. We'll also use the terms "$C^\infty$"and "smooth" interchangeably.

**Example 1.1.8.** Euclidean space, $\mathbb{R}^n$, is a manifold with a single chart where the coordinate map is the identity map i.e., $(\mathbb{R}^n, \mathbb{1}_{\mathbb{R}^n})$.

**Example 1.1.9.** Consider a manifold $M$ with atlas $(U_\alpha, \phi_\alpha)$ and $V$, an open subset of $M$. Now, if we consider $V$ as a topological space with the subset topology, we get that the collection of charts $(U_\alpha \cap V, \phi_\alpha|_{U_\alpha \cap V})$ is an atlas for $V$, where $\phi_\alpha|_{U_\alpha \cap V}$ denotes the restriction of $\phi_\alpha$ to $U_\alpha \cap V$. As an open subset of a manifold, $V$ is automatically second countable and Hausdorff. Therefore, $V$ is a manifold.

**Example 1.1.10.** We view the unit circle, $S^1$, as the set of points $(x, y) \in \mathbb{R}^2$ satisfying $x^2 + y^2 = 1$ and we give an atlas made of four charts. We can cover $S^1$ with four open sets: the upper and lower semicircles $U_1, U_2$ and the right and left semicircles $U_3, U_4$.

$$U_1 = \left\{(x, y) \in \mathbb{R}^2 |\ y > 0\right\} \quad , \quad U_2 = \left\{(x, y) \in \mathbb{R}^2 |\ y < 0\right\} ,$$
$$U_3 = \left\{(x, y) \in \mathbb{R}^2 |\ x > 0\right\} \quad , \quad U_4 = \left\{(x, y) \in \mathbb{R}^2 |\ x < 0\right\} .$$

Now we can project the upper and lower semicircles onto the $x$-axis to get two homeomorphisms, namely $\phi_i : U_i \to (-1, 1)$ , $\phi_i(x, y) = x$ for $i = 1, 2$. Similarly, we can project the right and left semicircles onto the $y$-axis and get the homeomorphisms $\phi_i : U_i \to (-1, 1)$ , $\phi_i(x, y) = y$ for $i = 3, 4$. It is now easy to check that, for all the pairwise non-empty intersections $U_i \cap U_j$, the transition function $\phi_i \circ \phi^{-1_j}$ is smooth. For example, on $U_1 \cap U_4$ we have

$$(\phi_3 \circ \phi_1^{-1})(x) = \phi_3(x, \sqrt{1 - x^2}) = \sqrt{1 - x^2}$$

$$(\phi_1 \circ \phi_3^{-1})(y) = \phi_1(\sqrt{1 - y^2}, y) = \sqrt{1 - y^2}$$

which are both smooth. Similarly, this holds for every pair. Therefore, $S^1$ is a smooth manifold with atlas $\{(U_i, \phi_i)\}$.
An alternative atlas for $S^1$ is given by the *stereographic projections.* These maps are given by $\Phi_{\pm 1} : S^1 \setminus \{(0, \pm 1)\} \to \mathbb{R}$, where

$$\phi_{\pm 1}(x, y) = \frac{x}{1 \mp y} ,$$

with inverses given by

$$(\phi_{\pm 1})^{-1}(t) = \left( \frac{2t}{t^2 + 1}, \pm \frac{t^2 - 1}{t^2 + 1} \right) .$$

Since both maps are continous bijections with a continous inverse they're homeomorphisms. Thus they're charts $(U_{\pm 1}, \phi_{\pm 1})$. We now show these charts are compatible. We have that $\Phi_{\pm 1}(U_{+1} \cap U_{-1}) = \mathbb{R} \setminus \{0\}$ and the transition functions are

$$(\phi_{\pm 1} \circ (\phi_{\mp 1})^{-1})(t) = \frac{2t}{(t^2 + 1) \mp (\mp(t^2 - 1))} = \frac{1}{t} .$$

Both of these smooth on $\mathbb{R} \setminus \{0\}$, therefore $\{(U_{+1}, \phi_{+1}), (U_{-1}, \phi_{-1})\}$ is an atlas for $S^1$.

## 1.2 Smooth maps on manifolds

In this section define the notion of smooth maps between manifolds and provide some criteria to determine the smoothness of a map. The main reference for this section is Section 6 of [4].

**Definition 1.2.1.** Let $N$ and $M$ be manifolds of dimension $n$ and $m$ respectively and let $F : N \to M$ be a continuous map. $F : N \to M$ is said to be $C^\infty$, or *smooth*, at a point $p$ if there are charts $(U, \phi)$ about $p$ and $(V, \psi)$ about $F(p)$ such that the composition $\psi \circ F \circ \phi^{-1} : \phi(U \cap F^{-1}(V)) \to \mathbb{R}^m$

$$\begin{tikzcd} U \subset N \arrow[r, "F"] & V \subset M \arrow[d, "\psi"] \\ \phi(U) \subset \mathbb{R}^n \arrow[u, "\phi^{-1}"] \arrow[r, "\psi \circ F \circ \phi^{-1}"] & \mathbb{R}^m \end{tikzcd} \tag{1.1}$$

is smooth at $\phi(p)$. The continuous map $F : N \to M$ is said to be $C^\infty$, or *smooth*, if it is smooth at every point of $N$.

**Remark 1.2.2.**

(1) In Definition 1.2.1 we assume that $F : N \to M$ is continuous to guarantee that the set $F^{-1}(V)$ is open in $N$. This ensures that the set $\phi(U \cap F^{-1}(V))$ is open in $\mathbb{R}^n$ so that it makes sense to talk about differentiability.

(2) We usually reserve the term function for maps that have $\mathbb{R}$, or one of its subsets, as the target space, i.e. maps such as $f : M \to D \subset \mathbb{R}$.

A general principle of the theory of differentiable manifolds is that constructions should be intrinsically defined, meaning they should not depend on the coordinate system we choose to describe them. In particular, for the notion of a smooth map to be well defined it needs to be independent of the choice of charts, as we show in the next proposition.

**Proposition 1.2.3.** *Let $N$ and $M$ be manifolds. Suppose $F : N \to M$ is a smooth at $p \in N$. If $(U, \phi)$ and $(V, \psi)$ are arbitrary charts about $p$ and $F(p)$ respectively, then $\psi \circ F \circ \phi^{-1}$ is smooth at $\phi(p)$.*

*Proof.* Since $F$ is $C^\infty$ at $p$ there exist charts $(U_0, \phi_0)$ about $p$ and $(V_0, \psi_0)$ about $F(p)$ such that the composition $\psi_0 \circ F \circ \phi_0^{-1}$ is smooth at $\phi_0(p)$. Since both $N$ and $M$ are manifolds, the compatibility of charts implies that both $\phi_0 \circ \phi$ and $\psi \circ \psi_0^{-1}$ are $C^\infty$ maps between subsets of Euclidean spaces. Hence the composite map

$$\psi \circ F \circ \phi^{-1} = (\psi \circ \psi_0^{-1}) \circ (\psi_0 \circ F \circ \phi_0^{-1}) \circ (\phi_0 \circ \phi^{-1})$$

is a smooth map. In particular, it is smooth at $\phi(p)$. □

**Definition 1.2.4.** Let $F : N \to M$ be a map and $h$ a function on $M$. The *pullback* of $h$ by $F$, denoted $F^*h$, is the composite function $h \circ F$.

In this terminology, a function $f$ is smooth on a chart $(U, \phi)$ if and only if its pullback $(\phi^{-1})^* f$ by $(\phi^{-1})$ is $C^\infty$ on $\phi(U)$.

**Proposition 1.2.5.** *Let $N$ and $M$ be smooth manifolds of dimension $n$ and $m$ and let $F : N \to M$ be a continuous map. The following are equivalent:*

(1) *The map $F : N \to M$ is $C^\infty$*

(2) *There are atlases $\mathfrak{U}$ of $N$ and $\mathfrak{V}$ of $M$ such that, for every chart $(U, \phi)$ in $\mathfrak{U}$ and $(V, \psi)$ in $\mathfrak{V}$, the map*

$$\psi \circ F \circ \phi^{-1} : \phi(U \cap F^{-1}(V)) \to \mathbb{R}^m$$

*is* $C^\infty$.

(3) *For every chart* $(U, \phi)$ *on* $N$ *and* $(V, \psi)$ *on* $M$ *the map*

$$\psi \circ F \circ \phi^{-1} : \phi(U \cap F^{-1}(V)) \to \mathbb{R}^m$$

*is* $C^\infty$.

*Proof.* For the proof see Proposition 6.8 in [4]. □

As an immediate consequence, we get that the composition of smooth maps is itself a smooth map.

**Corollary 1.2.6.** *Let* $F : N \to M$ *and* $G : M \to P$ *be smooth maps between manifolds, then the composite* $G \circ F : N \to P$ *is smooth.*

*Proof.* See Proposition 6.9 in [4]. □

**Example 1.2.7.** Let $n$ and $m$ be two positive integers so that $\mathbb{R}^{n\times m}$ denotes the space of all $n \times m$ matrices. Since this is isomorphic to $\mathbb{R}^{nm}$ we can give it it the topology of $\mathbb{R}^{nm}$. We define the *general linear group* $\mathrm{GL}(n, \mathbb{R})$ to be the subset of $n \times n$ invertible matrices. This is an open subset of $\mathbb{R}^{n\times n}$ since it is the preimage of an open set under the continuous map $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ ,

$$\mathrm{GL}(n, \mathbb{R}) = \left\{ A \in \mathbb{R}^{n\times n} \mid \det(A) \neq 0 \right\} = \det{}^{-1}(\mathbb{R} \setminus \{0\}) \ .$$

Therefore, by Example 1.1.9, $\mathrm{GL}(n, \mathbb{R})$ is a manifold.
Consider the matrix multiplication map $\mu : \mathrm{GL}(n, \mathbb{R}) \times \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$. Since the $(i, j)$-entry of the product matrix of $A, B \in \mathrm{GL}(n, \mathbb{R})$ is a polynomial in the coordinates of the matrices, namely

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \ ,$$

the multiplication map $\mu$ is smooth. Analogously, by the inverse matrix formula we have that

$$(A^{-1})_{ij} = \frac{1}{\det A} (-1)^{i+j} ((i, j)\text{-minor of } A) \ ,$$

which is a smooth function of the entries of the matrix $A$ as long as $\det(A) \neq 0$. This shows that the inverse map $\iota : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$ is $C^\infty$ as well.

We now focus on the class of smooth bijective maps, whose inverse is also smooth.

**Definition 1.2.8.** A *diffeomorphism* is a smooth bijective map $F : N \to M$ whose inverse $F^{-1}$ is also $C^\infty$.

It turns out that every chart of an $n$-dimensional manifold is a diffeomorphism onto a subset of $\mathbb{R}^n$ and that every such diffeomorphism naturally induces a chart.

**Proposition 1.2.9.** *Let $M$ be a $n$-dimensional manifold. If $(U, \phi)$ is a chart on $M$, then the coordinate map $\phi : U \to \phi(U) \subset \mathbb{R}^n$ is a diffeomorphism. Conversely, if $U$ is an open subset of $M$ and $\phi : U \to \phi(U) \subset \mathbb{R}^n$ is a diffeomorphism, then $(U, \phi)$ is a chart in the differentiable structure of $M$.*

*Proof.* For the proof see Propositions 6.10 and 6.11 of [4]. □

Finally, we discuss how the notion of partial derivative generalizes to the context of manifolds. Let $M$ be a $n$-dimensional manifold, $(U, \phi)$ be a chart and $f$ a smooth function. We denote the standard coordinates on $\mathbb{R}^n$ with $r^1, \ldots, r^n$ so that the components of $\phi$ are given by $x^i = r^i \circ \phi$. For $p \in U$ we define the *partial derivative $\partial f/\partial x^i$ of $f$ with respect to $x^i$ at $p$* to be

$$\frac{\partial f}{\partial x^i}(p) := \frac{\partial}{\partial x^i}\Big|_p f := \frac{\partial}{\partial r^i}\Big|_{\phi(p)} (f \circ \phi^{-1}) := \frac{\partial (f \circ \phi^{-1})}{\partial r^i}(\phi(p)) \tag{1.2}$$

Since $p = \phi^{-1}(\phi(p))$ we get the equation

$$\frac{\partial f}{\partial x^i} \circ \phi^{-1} = \frac{\partial (f \circ \phi^{-1})}{\partial r^i}$$

on $\phi(U)$. Therefore, the partial derivative $\partial f/\partial x^i$ is $C^\infty$ on $U$ because its pullback by $\phi^{-1}$ is smooth on $\phi(U)$.

**Definition 1.2.10.** Let $F : N \to M$ be a smooth map, let $(U, \phi) = (U, x^1, \ldots, x^n)$ and $(V, \psi) = (V, y^1, \ldots, y^n)$ be charts on N and M respectively such that $F(U) \subset V$ and denote

$$F^i = y^i \circ F = r^i \circ \psi \circ F : U \to \mathbb{R} .$$

The matrix $[\partial F^i/\partial x^j]$ is called the *Jacobian matrix* of $F$ with respect to $(U, \phi)$ and $(V, \psi)$.

## 1.3 Tangent vectors and tangent spaces

In this section we introduce the concept of tangent vectors and tangent spaces. Later, we define the differential as a linear map between tangent spaces. The references for this section are Sections 8 and 11.2 of [4].

Let $p$ be a point on a manifold $M$ and consider all the set of pairs $(f, U)$ where $U$ is a neighborhood of $p$ and $f : U \to \mathbb{R}$ is $C^\infty$. We say two pairs $(f, U)$ and $(g, V)$ are equivalent if $f$ and $g$ agree on a neighborhood of $p$. This is an equivalence relation, so we call an equivalence class $[(f, U)]$ the *germ* of $f$ at $p$. We denote the set of all germs of functions on $M$ at $p$ as $C_p^\infty(M)$. Addition and multiplication of functions makes $C_p^\infty(M)$ into an algebra over $\mathbb{R}$.

**Definition 1.3.1.** Let $D : C_p^\infty(M) \to \mathbb{R}$ be an $\mathbb{R}$-linear map. We say $D$ is a *derivation at a point* $p$ if it satisfies the *Leibniz rule*

$$D(fg) = (Df)g(p) + f(p)(Dg)$$

The key concept in generalizing the notion of a tangent vector to abstract spaces such as manifolds is thinking about tangent vectors as derivation. As it turns out, in $\mathbb{R}^n$ the tangent space at any point (which is identified with $\mathbb{R}^n$ itself) is isomorphic to the vector space of all derivations at that point (see Theorem 2.2 in [4]). This motivates the following definition.

**Definition 1.3.2.** A *tangent vector* at a point $p$ in a manifold $M$ is a derivation at $p$. The tangent vectors at $p$ form a vector space called the *tangent space of* $M$ *at* $p$, denoted $T_pM$.

**Remark 1.3.3.** Since the notion of germs is local, it is clear that, for any neighborhood of a point $p \in M$, $C_p^\infty(U)$ is the same as $C_p^\infty(M)$, hence $T_pU = T_pM$.

**Example 1.3.4.** Let $(U, \phi) = (U, x^1, \dots, x^n)$ be a chart about a point $p$ is a manifold $M$ and let $r^1, \dots, r^n$ be the standard coordinates in $\mathbb{R}^n$ such that $x^i = r^i \circ \phi$. Using the fact that the partial derivatives on $\mathbb{R}^n$ are derivation we can show that the partial derivative $\partial/\partial x^i$ is tangent vector. If $f$ and $g$ are smooth functions on a neighborhood of $p$, by Equation (1.2) we get

$$\begin{aligned}
\frac{\partial\,(fg)}{\partial x^i}(p) &= \frac{\partial\,((fg)\circ\phi^{-1})}{\partial r^i}(\phi(p)) \\
&= \frac{\partial\,(f\circ\phi^{-1})}{\partial r^i}(\phi(p))(g\circ\phi^{-1})(\phi(p)) + (f\circ\phi^{-1})(\phi(p))\frac{\partial\,(g\circ\phi^{-1})}{\partial r^i}(\phi(p)) \\
&= \frac{\partial f}{\partial x^i}(p)g(p) + f(p)\frac{\partial g}{\partial x^i}(p)\ ,
\end{aligned}$$

which shows that the partial derivatives with respect to the coordinates are tangent vectors at every point in their respective coordinate neighborhood.

For simplicity's sake, when dealing with one dimensional manifolds, we will use $t$ as the coordinate and use the notation $d/dt$ instead of $\partial/\partial t$.

**Definition 1.3.5.** Let $F : N \to M$ be a $C^\infty$ map between manifolds. We call *differential at $p$ of $F$* the induced map $F_* : T_pN \to T_{F(p)}M$ defined by

$$(F_*(X_p))f = X_p(f \circ F)$$

for $f \in C^\infty_{F(p)}(M)$ and $X_p \in T_pN$. To make the dependence on $p$ explicit we sometimes write $F_{*,p}$. Occasionally, we will use the alternative notation $(dF)_p$.

**Proposition 1.3.6.** *If $F : N \to M$ and $G : M \to P$ are smooth maps between manifolds and $p \in N$, then*

$$(G \circ F)_{*,p} = G_{*,F(p)} \circ F_{*,p}$$

*Proof.* Let $X_p \in T_pN$ and $f$ be a smooth function at $G(F(p)) \in P$, then

$$\begin{aligned}((G \circ F)_*(X_p))f &= X_p(f \circ G \circ F) = (F_*(X_p))(f \circ G) \\ &= (G_*(F_*(X_p)))f = ((G_* \circ F_*)(X_p))f\end{aligned}$$

□

**Corollary 1.3.7.** *If $F : N \to M$ is a diffeomorphism and $p \in N$, then $F_* : T_pN \to T_{F(p)}M$ is an isomorphism. In particular if two open sets $U \subset \mathbb{R}^n$ and $V \in \mathbb{R}^m$ are diffeomorphic, then $n = m$*

*Proof.* For the proof see Corollaries 8.6 and 8.7 in [4]. □

The second claim, which takes the name of *(smooth) invariance of dimension*, finally shows that the dimension of a manifold is well-defined.

As we already remarked earlier in this section, the tangent space at any point in $\mathbb{R}^n$ is identified with both the Euclidean space itself, making it finite dimensional, and the derivations at that point (See Section 2 in [4]). This induces a basis on the tangent space given by $\{\partial/\partial r^i\}_{i=1}^n$, where $r^1, \ldots, r^n$ are the standard coordinates on $\mathbb{R}^n$. Since the charts of a manifold are local diffeomorphisms between the manifold and Euclidean space, their differential provides a way to induce a basis on the tangent space.

**Proposition 1.3.8.** *Let $(U, \phi) = (U, x^1, \ldots, x^n)$ and $(V, \psi) = (V, y^1, \ldots, y^n)$ be charts about a point $p$ in a manifold $M$, then:*

(1)

$$\phi_* \left( \frac{\partial}{\partial x^i} \right) = \left. \frac{\partial}{\partial r^i} \right|_{\phi(p)}$$

(2) *$T_pM$ has a basis given by*

$$\left.\frac{\partial}{\partial x^1}\right|_p, \dots, \left.\frac{\partial}{\partial x^n}\right|_p$$

*Proof.* See Propositions 8.8, 8.9, 8.10 in [4]. □

Since the differential is a linear map between vector spaces, once we fix a basis on each vector space, it can be represented by a matrix.

**Proposition 1.3.9.** *Let $F : N \to M$ be a smooth map between manifolds and let $(U, x^1, \dots, x^n)$ and $(V, y^1, \dots, y^n)$ charts about $p \in N$ and $F(p)$ respectively. Relative to the bases $\{\partial/\partial x^i\}$ for $T_pN$ and $\{\partial/\partial y^i\}$ for $T_{F(p)}M$ the differential $F_{*,p} : T_pN \to T_{F(p)}M$ is represented by the Jacobian matrix $[\partial F^i/\partial x^j]$.*

*Proof.* As a linear map, $F_{*,p}$ is fully determined by how it acts on the basis, meaning it's determined the numbers $a^i_j$ given by

$$F_*\left(\left.\frac{\partial}{\partial x^j}\right|_p\right) = \sum_{k=1}^m a^k_j \left.\frac{\partial}{\partial y^k}\right|_{F(p)} \quad , \quad j = 1, \dots, n\,.$$

Applying $y^i$ to both sides yields

$$a^i_j = \left(\sum_{k=1}^m a^k_j \left.\frac{\partial}{\partial y^k}\right|_{F(p)}\right) y_i = F_*\left(\left.\frac{\partial}{\partial x^j}\right|_p\right) y_i = \left.\frac{\partial}{\partial x^j}\right|_p (y^i \circ F) = \frac{\partial F^i}{\partial x^j}\ ;$$

where we used the fact that $(\partial/\partial y^k)(y^i) = \delta^i_k$. □

**Definition 1.3.10.** Let $F : N \to M$ be a $C^\infty$ map.

1. The map is said to be an *immersion* at $p$ if $F_{*,p}$ is injective. $F$ is called an *immersion* if it is an immersion at every point in $N$.
2. The map is said to be a *submersion* at $p$ if $F_{*,p}$ is surjective. $F$ is called a *submersion* if it is a submersion at every point in $N$.

Immersions (submersions) are charaterized by the property that, for an apporpriate choice of charts, they locally look like inclusions (projections). These facts are known as the immersion and submersion theorem respectively.

**Theorem 1.3.11.** *Let $N$ and $M$ be manifolds of dimension $n$ and $m$ respectively.*

(1) *Suppose $f : N \to M$ is an immersion at $p \in N$. Then, there are charts $(U, \phi)$ centered at $p \in N$ and $(V, \psi)$ centered at $f(p) \in M$ such that*

$$(\psi \circ f \circ \phi^{-1})(r^1, \ldots, r^n) = (r^1, \ldots, r^n, 0, \ldots 0)$$

(2) *Suppose $f : N \to M$ is a submersion at $p \in N$. Then, there are charts $(U, \phi)$ centered at $p \in N$ and $(V, \psi)$ centered at $f(p) \in M$ such that*

$$(\psi \circ f \circ \phi^{-1})(r^1, \ldots r^m, r^{m+1}, \ldots, r^n) = (r^1, \ldots, r^m)$$

*Proof.* See Section 11.2 in [4]. □

## 1.4 Vector fields

In this section we define the notion tangent bundle for a smooth manifold and define smooth vector fields on manifolds. The main reference for this section are sections 12 and 14 of [4].

We now want to focus on the additional structure that is provided to a manifold $M$ by the union of all its tangent spaces. Consider their (disjoint) union which we call the *tangent bundle*

$$TM = \bigsqcup_{p \in M} T_pM \ .$$

This construction comes equipped with a natural projection map $\pi : TM \to M$ given by $\pi(v) = p$ if $v \in T_pM$. To give a topology to the tangent bundle we can proceed as follows. Consider a chart $(U, \phi) = (U, x^1, \ldots, x^n)$ on $M$, by Remark 1.3.3 we have that

$$TU = \bigsqcup_{p \in U} T_pU = \bigsqcup_{p \in U} T_pM \ .$$

Since at any point $p \in U$ the tangent vectors $\{\partial/\partial x^i\}_i$ are a basis for $T_pM$, any $v \in T_pM$ can be expressed by a unique linear combination

$$v = \sum_i c^i \frac{\partial}{\partial x^i} \ .$$

Since the coefficients $c^i = c^i(v)$ depend in $v$ they are functions on $TU$. If we let $\overline{x}^i = x^i \circ \pi$ we can define $\tilde{\phi} : TU \to \phi(U) \times \mathbb{R}^n$ by

$$v \to (\overline{x}^1, \ldots, \overline{x}^n, c^1, \ldots, c^n)(v) \ ,$$

which has an inverse

$$(\phi(p), c^1, \ldots, c^n) \to \sum_i c^i \frac{\partial}{\partial x^i}\bigg|_p .$$

Since $\tilde{\phi}$ is a bijection we can transfer the topology from $\phi(U) \times \mathbb{R}^n \in \mathbb{R}^{2n}$ to $TU$. With this in mind, one can show (Lemma 12.1 in [4]) that the collection

$$\mathfrak{B} = \bigcup_\alpha \{A \mid A \text{ open in } TU_\alpha\,,\, U_\alpha \text{ a coordinate open set in } M\}$$

is a suitable basis for a topology; therefore we give $TM$ the topology generated by $\mathfrak{B}$. It follows that $TM$ is both second countable and Hausdorff and that any $C^\infty$ atlas $\{(U_\alpha, \phi_\alpha)\}$ for $M$ induces an atlas for $TM$ given by $\{(TU_\alpha, \tilde{\phi}_\alpha)\}$ (as shown in section 12 of [4]). This makes $TM$ a differentiable manifold. For further details see Section 12 of [4].

**Definition 1.4.1.** A *vector field* $X$ on a manifold $M$ is a map that assigns a tangent vector $X_p \in T_pM$ to every point $p \in M$. We say a vector field $X$ is *smooth* if the map $X : M \to TM$ is a smooth map between manifolds. We denote $\mathfrak{X}(M)$ the set of all smooth vector fields on $M$.

Vector fields can still act on functions just like tangent vectors, where, for any function $f$ on $M$, the operation is now defined pointwise

$$(Xf)(p) = X(p)f = X_p(f)$$

and it induces a function $Xf : M \to \mathbb{R}$. Additionally, if we consider the projection map $\pi : TM \to M$, every vector field satisfies $\pi \circ X = \mathrm{id}_M$.

**Definition 1.4.2.** Let $M$ be a manifold. A *frame* for the tangent bundle $TM$ over an open set $U \subset M$ is a collection of vector fields $X_1, \ldots, X_n$ on $M$ such that, for every $p \in U$, the tangent vectors $X_1(p), \ldots, X_n(p)$ form a basis for $T_pM$.

The following proposition characterizes smooth vector fields.

**Proposition 1.4.3.** *Let $X$ be a vector field on a manifold $M$, then the following are equivalent:*

(1) *The vector field $X$ is smooth.*

(2) *The manifold $M$ has an atlas such that, for any chart $(U, x^1, \ldots, x^n)$ of the atlas, the coefficients $a^i$ of $X = \sum a^i\, \partial/\partial x^i$ relative to the frame $\{\partial/\partial x^i\}$, are all smooth functions on $M$.*

(3) *On any chart $(U, x^1, \ldots, x^n)$ on $M$ the coefficients $a^i$ of $X = \sum a^i\, \partial/\partial x^i$ relative to the frame $\{\partial/\partial x^i\}$, are all smooth functions on $M$.*

(4) *For every smooth function $f$ on $M$, the function $Xf$ is smooth on $M$*

*Proof.* For the proof see Propositions 14.2 and 14.3 of [4]. □

If we denote $C^\infty(M)$ the ring of all smooth functions on $M$, since vector fields act pointwise on functions, we can think of them as derivation on $C^\infty(M)$. We define the *Lie bracket* $[X, Y]$ of two vector fields $X$ and $Y$ as follows

$$[X, Y]_p \, f = (X_p Y - Y_p X) f$$

for any germ $f$ of a smooth function at $p$. It can be shown that the Lie bracket satisfies the Leibniz rule so that $[X, Y]$ is vector field, which is also $C^\infty$ if $X$ and $Y$ are (see Section 14.4 of [4]). Therefore, the Lie bracket makes $\mathfrak{X}(M)$ into a Lie algebra.

**Definition 1.4.4.** Let $K$ be a field. A *Lie algebra* over $K$ is a vector space $V$ endowed with a product $[\cdot, \cdot] : V \times V \to V$, called the bracket, satisfying the following properties for all $a, b \in K$ and $X, Y, Z \in V$

(1) (bilinearity)
$$[aX + bY, Z] = a\,[X, Z] + b\,[Y, Z]$$
$$[Z, aX + bY] = a\,[Z, X] + b\,[Z, Y]$$

(2) (anticommutativity) $[Y, X] = -[X, Y]$

(3) (Jacobi identity)
$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

A map between Lie algebras $\phi : V \to W$ is called a Lie algebra homomorphism if it is linear and
$$\phi([v, w]) = [\phi(v), \phi(w)] \quad \forall v, w \in V \ .$$
A Lie algebra isomorphism is a bijective Lie algebra homomorphism.

**Definition 1.4.5.** Let $F : N \to M$ be a smooth map between manifolds. A vector field $X$ on $N$ is *F-related* to a vector field $\hat{X}$ on $M$ if for all $p \in N$

$$F_{*,p}(X_p) = \hat{X}_{F_p} \ . \tag{1.3}$$

**Proposition 1.4.6.** *Let $F : N \to M$ be a smooth map. If two smooth vector fields $X$ and $Y$ on $N$ are F-related to two smooth vector fields $\hat{X}$ and $\hat{Y}$, respectively, on $M$, then the Lie bracket $[X, Y]$ on $N$ is F-related to the Lie bracket $[\hat{X}, \hat{Y}]$ on M.*

*Proof.* For the proof see Proposition 14.17 in [4]. □

## 1.5 Submanifolds

In this section we define both immersed and embedded submanifolds, state the implicit function theorem on manifolds and discuss the relation between vector fields on submanifold and on the ambient space. The references for this section are Section 1.6 and 2.7 of [3] and Section 11.3 of [4].

**Definition 1.5.1.** Let $N$ be a smooth manifold. An *immersed submanifold* of $N$ is a pair $(M, \varphi)$ where $M$ is a smooth manifold and $\varphi : M \to N$ is an injective immersion. Two submanifolds $(M_1, \varphi_1)$ and $(M_2, \varphi_2)$ are said to be equivalent if there exists a diffeomorphism $\psi : M_1 \to M_2$ such that $\varphi_1 = \varphi_2 \circ \psi$.

Given an immersed submanifold $(M, \varphi)$ of a manifold $N$, consider the induced mapping $\tilde{\varphi} : M \to \varphi(M)$. Since $\varphi$ is a bijection, $\varphi(M)$ can inherit both a topology and a differentiable structure from this mapping, thus making $\tilde{\varphi}$ into a diffeomorphism. This implies that $(\varphi(M), i)$, where $i : \varphi(M) \to N$ is the natural inclusion, is an immersed submanifold of $N$ equivalent to $(M, \varphi)$. Therefore, up to equivalence, every immersed submanifold of $N$ can be assumed to be given by a subset paired with the corresponding natural inclusion mapping.
We note that, the topology induced on $\varphi(M)$ by the mapping $\tilde{\varphi}$ might not coincide with the subset topology induced by $N$. In fact, when $\varphi(M)$ is given the subset topology, it needs not to be a manifold at all (see Section 11.3 of [4]). To solve this issue, it is sufficient to require that the mapping $\varphi$ be open onto its image. This makes it so that the image $\varphi(M)$ remains homeomorphic to $M$ under $\varphi$ even when given the subset topology.

**Definition 1.5.2.** A smooth map $\varphi : M \to N$ is called an *embedding* if it is a one-to-one immersion which is open onto its image. A submanifold $(M, \varphi)$ of a manifold $N$ is called an *embedded submanifold* if $\varphi : M \to N$ is an embedding.

**Definition 1.5.3.** Let $F : N \to P$ be a smooth map between manifolds. We say that a value $c \in P$ is a *critical value* of $F$ if $F$ is not submersion for some $p \in F^{-1}(c)$. We say $c \in N$ is *regular value* of $F$ if it is not a critical value.

We remark that, in general, a regular value does not need to be the in the image of $F$. The following is the manifold version of the implicit function theorem, also known as the regular level set theorem.

**Theorem 1.5.4.** *Let $F : N \to P$ be a smooth map between manifolds and let $c \in P$ be a regular value of $F$ such that $M := F^{-1}(c)$ is non-empty. Then $(M, i)$ is an embedded submanifold of $N$ of dimension* $\dim(N) - \dim(P)$, *where $i : M \hookrightarrow N$ is the natural inclusion map. Furthermore, $i_{*,m}(T_m M) = \ker(F_{*,m})$ for every $m \in M$.*

*Proof.* See Corollary 1.8.3 in [3]. □

To end this section we discuss how vector fields on submanifolds are related to vector fields on the ambient space.

**Definition 1.5.5.** Let $N$ be a manifold and $(M, \varphi)$ be an immersed submanifold of $N$. We say that a vector field $X \in \mathfrak{X}(N)$ is *tangent* to $(M, \varphi)$ if $X_{\varphi(m)} \in \varphi_*(T_m M)$ for all $m \in M$

**Proposition 1.5.6.** *Let $N$ be a manifold and $(M, \varphi)$ be an immersed submanifold of $N$. For every vector field $X \in \mathfrak{X}(N)$ which tangent to $(M, \varphi)$, there is a unique vector field $\tilde{X}$ on $M$ which is $\varphi$ related to $X$, i.e., $\varphi_* \circ \tilde{X} = X \circ \varphi$.*

*Proof.* See Proposition 2.7.16 in [3]. □

## 1.6 Smooth curves

In this section we define smooth curves and show how they can be used to compute the differential of a map. The main reference for this section is Section 8 in [4].

We say a map $c : (a, b) \to M$ is a *smooth curve* in a manifold $M$ if it is smooth map from an open interval to M. We usually assume $0 \in (a, b)$ and say that the curve *starts at $p$* if $c(0) = p$. We define the curve's *velocity vector* $c'(t_0)$ at time $t_0 \in (a, b)$ to be the tangent vector given by

$$c'(t_0) \coloneqq c_* \left( \left. \frac{d}{dt} \right|_{t_0} \right) \in T_{c(t_0)} M.$$

Alternative notations for the velocity vector are

$$\frac{dc}{dt}(t_0) \quad \text{and} \quad \left. \frac{d}{dt} \right|_{t_0} c\,.$$

To avoid confusions, in the case where $c : (a, b) \to \mathbb{R}$, we denote the regular derivative with respect to $t$ with $\dot{c}(t)$.

**Proposition 1.6.1.** *Let $c : (a, b) \to M$ be a smooth curve and let $(U, x^1, \ldots, x^n)$ be a chart about $c(t)$. If we write $c^i = x^i \circ c : (a, b) \to \mathbb{R}$, then $c'(t)$ is given by*

$$c'(t) = \sum_i \dot{c}^i(t) \left. \frac{\partial}{\partial x^i} \right|_{c(t)} .$$

*Proof.* See Proposition 8.15 in [4]. □

**Proposition 1.6.2.** *For any point $p$ in a manifold $M$ and any tangent vector $X_p \in T_p M$ there are $\epsilon > 0$ and a smooth curve $c : (-\epsilon, \epsilon) \to M$ such that $c(0) = p$ and $c'(0) = X_p$.*

*Proof.* See Proposition 8.16 in [4]. □

**Proposition 1.6.3.** *Let $F : N \to M$ be a smooth map between manifolds, $p \in N$ and $X_p \in T_pN$. If $c$ is a smooth curve starting at $p$ with velocity $X_p$ at $p$, then $F_*(X_p)$ is the velocity vector of $F \circ c$ at $F(p)$, meaning that*

$$F_{*,p}(X_p) = \left.\frac{d}{dt}\right|_0 (F \circ c)(t) \; .$$

*Proof.* By hypothesis, $c(0) = p$ and $c'(0) = X_p$. Therefore,

$$F_{*,p}(X_p) = F_{*,p}(c'(0)) = (F_{*,p} \circ c*, 0) \left( \left.\frac{d}{dt}\right|_0 \right) \; .$$

Now, by the chain rule (Proposition 1.3.6), we have

$$F_{*,p}(X_p) = (F \circ c)_{*,0} \left( \left.\frac{d}{dt}\right|_0 \right) = \left.\frac{d}{dt}\right|_0 (F \circ c)\,(t) \; .$$

□

## 1.7 Integral curves and flows

In this section we define integral curves of vector fields, discuss their existence and uniqueness and show how they induce a map known as flow. The main reference for this section is Section 3.2 of [3].

**Definition 1.7.1.** Let $M$ be a manifold and let $X \in \mathfrak{X}(M)$. We say a smooth curve $c : I \to M$ is an *integral curve of* $X$ if

$$c'(t) = X_{c(t)} \;\text{ for all } t \in I \, . \tag{1.4}$$

We say an integral curve is *maximal* if it doesn't admit an extension and we say it is *complete* if $I = \mathbb{R}$. A vector field is called complete if its maximal integral curve is complete.

In local coordinates $(U, \phi) = (U, x^1, \dots, x^n)$, where the vector field is given by

$$X_p = \sum_{i=1}^{n} a^i(p) \left.\frac{\partial}{\partial x^i}\right|_p \, ,$$

if we call $\mathbf{q}(t) = \phi \circ c(t)$ and $q^i(t) = x^i \circ c(t)$, Equation (1.4) becomes

$$\dot{q}^i(t) = (a^i \circ \phi^{-1})(\mathbf{q}(t)) \qquad i = 1, \dots, n \, .$$

Since this is system of first order differential equations with smooth coefficients a unique solution is guaranteed. This leads to the following result.

**Proposition 1.7.2.** *Let $M$ be a manifold and let $X \in \mathfrak{X}(M)$. For every $m \in M$ there exists a unique maximal integral curve $\gamma_m : I_m \to M$ of $X$ starting at $p$, i.e. $\gamma_p(0) = p$.*

*Proof.* For the proof see Theorem 3.2.3 in [3]. □

**Definition 1.7.3.** Let $\mathcal{D}$ be an open neighborhood of $\{0\} \times M$ in $\mathbb{R} \times M$ and let $\Phi : \mathcal{D} \to M$ be a smooth map. Denote

$$\mathcal{D}_m = \{t \in \mathbb{R} \mid (t, m) \in \mathcal{D}\} \quad , \quad \mathcal{D}_t = \{m \in M \mid (t, m) \in \mathcal{D}\}$$

and let $\Phi_m : \mathcal{D}_m \to M$ and $\Phi_t : \mathcal{D}_t \to M$ denote the induced mappings

$$\Phi(t, m) = \Phi_t(m) = \Phi_m(t) \ .$$

We say $\Phi$ is a (smooth) flow on $M$ is the following hold:

(1) $\Phi_{t=0} = \mathrm{id}_M$

(2) if $(s, m) \in \mathcal{D}$ and $(t, \Phi_s(m)) \in \mathcal{D}$, then $(s + t, m)$ and $\Phi_t(\Phi_s(m)) = \Phi_{s+t}(m)$.

(3) $\mathcal{D}_m$ is an open interval for every $m \in M$.

(4) $\Phi_t(\mathcal{D}_t) \subset \mathcal{D}_{-t}$ for all $t \in \mathbb{R}$.

A flow is said to be *maximal* if it doesn't admit an extension which is itself a flow. It is called complete if $\mathcal{D} = \mathbb{R} \times M$.

For a given vector field $X$, if we consider all the maximal integral curves $\{\gamma_m : I_m \to M\}_{m \in M}$ of $X$ through each $m \in M$, we can form the subset

$$\mathcal{D}^X = \bigcup_{m \in M} I_m \times \{m\} \tag{1.5}$$

and define the mapping

$$\Phi^X : \mathcal{D}^X \to M \ , \quad \Phi^X(t, m) := \gamma_m(t) \ . \tag{1.6}$$

**Proposition 1.7.4.** *The assignment defined by Equations* (1.5) *and* (1.6) *defines a bijection between vector fields and maximal flows on $M$. Complete vector fields correspond to complete flows.*

*Proof.* See Proposition 3.2.6 in [3]. □

We refer to the map defined by Equation (1.6) as the *flow of the vector field $X$*. If the corresponding vector field is unambiguous we will simply write $\Phi : \mathcal{D} \to M$.

**Proposition 1.7.5.** *Let $M$ be a manifold, $X \in \mathfrak{X}(M)$ and $\Phi^X : \mathcal{D}^X \to M$ be its flow, then:*

(1) *For every $t \in \mathbb{R}$ such that $\mathcal{D}_t^X \neq \varnothing$, $\Phi_t$ is a diffeomorphism from $\mathcal{D}_t^X$ onto $\mathcal{D}_{-t}^X$ with inverse $\Phi_{-t}$.*

(2) *For every $t \in \mathbb{R}$ and every $s \in (0, t)$ we have $\mathcal{D}_t^X \subset \mathcal{D}_s^X$ and $\Phi_s(\mathcal{D}_t^X) \subset \mathcal{D}_{t-s}^X$.*

(3) *If $\Phi$ is complete, then $\Phi_s \circ \Phi_t = \Phi_{s+t}$ for all $s, t \in \mathbb{R}$.*

(4) *For $s \in \mathbb{R}$ let $\Phi^{sX} : \mathcal{D}^{sX} \to M$ be the flow of $sX$, then, for all $t \in \mathbb{R}$,*

$$\mathcal{D}_t^{sX} = \mathcal{D}_{st}^X \ , \ \ \Phi_t^{sX} = \Phi_{st}^X \, .$$

*Additionally, if $N$ is another manifold, $\varphi : M \to N$ is a smooth map, $Y \in \mathfrak{X}(N)$ and $t \in \mathbb{R}$, we have*

(5) *If $X$ and $Y$ are $\varphi$-related, then $\varphi(\mathcal{D}_t^X) \subset \mathcal{D}_t^Y$ and $\Phi_t^Y \circ \varphi|_{\mathcal{D}_t^X} = \varphi \circ \Phi_t^X$.*

(6) *If $\varphi$ is a diffeomorphism, then $\varphi(\mathcal{D}_t^X) = \mathcal{D}_t^{\varphi_* X}$ and $\Phi_t^{\varphi_* X} = \varphi \circ \Phi_t^X \circ (\varphi^{-1})|_{\mathcal{D}_t^{\varphi_* X}}$.*

*Proof.* For the proof see Propositions 3.2.10 and 3.2.13 in [3]. □

## 1.8 k-covectors and wedge product

In this section we briefly go over the properties of $k$-covectors and their wedge product over arbitrary vector spaces. The main reference for this section is Section 3 of [4].

We denote $V^k = V \times \cdots \times V$ the Cartesian product of $k$-copies of a real vector space $V$. We say a function $f : V^k \to \mathbb{R}$ is *$k$-linear* if it is linear is each of its $k$ arguments:

$$f(\dots, av + bw, \dots) = af(\dots, v, \dots) + bf(\dots, w, \dots)$$

for all $a, b \in \mathbb{R}$ and $v, w \in V$.

**Definition 1.8.1.** Let $V$ be a real vector space and let $S_k$ denote the group of permutations of the set $\{1, \dots, k\}$. We say a $k$-linear function $f : V^k \to \mathbb{R}$ is *alternating* if

$$f(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \operatorname{sgn}(\sigma) f(v_1, \dots, v_n)$$

for all permutations $\sigma \in S_k$. Such a function is also called a *$k$-covector*. The vector space of $k$-covectors is denoted $A_k(V)$. If $k = 1$ we simply talk about a covector.

In this context, the space $A_0(V)$ will be the space of constant functions, therefore $A_0(V) = \mathbb{R}$, meanwhile the space $A_1(V)$ is the space of linear functions over $V$ and coincides with

the dual space of $V$, which we denote $V^*$.
We now want to define a product of alternating functions, such that the resulting function is alternating as well. This motivates the following definition.

**Definition 1.8.2.** Let $f \in A_k(V)$ and $g \in A_l(V)$, we define their *wedge product*, also called *exterior product*, to be the $(k+l)$-covector $f \wedge g$ such that

$$(f \wedge g)(v_1, \dots, v_{k+l}) = \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} (\mathrm{sgn}(\sigma)) f(v_{\sigma(1)}, \dots, v_{\sigma(k)}) g(v_{\sigma(k+1)}, \dots, v_{\sigma(k+l)}) .$$

For the proof that $f \wedge g$ is actually alternating see Section 3.7 of [4].

**Proposition 1.8.3.** *Let $V$ be a real vector space and $f \in A_k(V)$, $g \in A_l(V)$ and $h \in A_m(V)$, then:*

(1) *The wedge product is anticommutative*
$$f \wedge g = (-1)^{kl} g \wedge f .$$
(2) *If $k$ is odd $f \wedge f = 0$.*
(3) *The wedge product is associative*
$$(f \wedge g) \wedge h = f \wedge (g \wedge h) .$$

*Proof.* For the proofs see Proposition 3.21, Corollary 3.23 and Proposition 3.25 of [4]. □

We will now construct a basis for $A_k(V)$. Let $e_1, \dots, e_n$ be a basis for $V$ and let $\alpha^1, \dots, \alpha^n$ be a basis for $V^*$. We introduce the *multi-index* notation

$$I = (i_1, \dots, i_k)$$

and write $e_I = (e_{i_1}, \dots, e_{i_k})$ and $\alpha^I = \alpha^{i_1} \wedge \dots \wedge \alpha^{i_k}$. It is clear that an alternating $k$-linear function is fully determined by its values on $(e_{i_1}, \dots, e_{i_k})$, where the $i_j$ are in strictly ascending order, meaning $1 \leq i_1 < \dots < i_k \leq n$. We denote the set of all such strictly ascending multi-indices as $\mathfrak{I}_{k,n}$.

**Proposition 1.8.4.** *Let $V$ be a real vector space and let $\alpha^1, \dots, \alpha^n$ be a basis for $V^*$. The alternating $k$-covectors $\alpha^I$, with $I$ strictly ascending multi-index, form a basis for $A_k(V)$.*

*Proof.* See Proposition 3.29 in [4]. □

It follows that $A_k(V)$ has dimension $\binom{n}{k}$, as this are is number of possible distinct strictly ascending multi-indices of length $k$ for $i_j \in \{1, \dots, n\}$. If $k \geq n$ then $A_k(V)$ is the trivial vector space, as every element of the basis $\alpha^{i_1} \wedge \dots \wedge \alpha^{i_k}$ will have a repeating covector in the wedge product, making each of the basis vector null via the associativity of the wedge product and Proposition 1.8.3/2.

## 1.9 Cotangent space and $k$-forms

In this section we restrict ourselves to the studying $k$-covectors over the tangent spaces of a manifold. We define differential $k$-forms, the differential of functions on a manifold and and the cotangent bundle. Finally, we give some criteria to determine the smoothness of a $k$-form. The main reference for this section is Chapter 5 of [4].

Let $M$ be a smooth manifold and $p \in M$. The *cotangent space* of $M$ at $p$, denoted $T_p^*M$, is defined to be the dual of tangent space $T_pM$, i.e. $T_p^*M = (T_pM)^*$.
An element of the cotangent space $T_p^*M$ is called a *covector at* $p$, which is a linear function $\omega_p : T_pM \to \mathbb{R}$. A *covector field*, or *differential* 1-*form*, or simply a 1-*form*, is a function $\omega$ that assigns a covector $\omega_p$ at $p$ to each point $p \in M$. A particular important case of 1-form is the differential of a function.

**Definition 1.9.1.** If $f$ is a smooth function on a manifold $M$, its *differential* is defined to be the 1-form $df$ on $M$ such that for any $p \in M$ and $X_p \in T_pM$

$$(df)_p(X_p) = X_p f \ .$$

Using the differential we can induce a basis on $T_p^*M$.

**Proposition 1.9.2.** *Let $(U, \phi) = (U, x^1, \dots, x^n)$. At each $p \in U$, the covectors $\{(dx^i)_p\}_{i=1}^n$ form a basis for $T_p^*M$, dual to the basis $\partial/\partial x^1, \dots, \partial/\partial x^n$ for $T_pM$.*

*Proof.* Since

$$(dx^i)_p \left( \left.\frac{\partial}{\partial x^j}\right|_p \right) = \frac{\partial x^i}{\partial x^j}(p) = \delta^i_j \ ,$$

the covectors $\{(dx^i)_p\}_{i=1}^n$ form a basis for $T_p^*M$. □

Given the result above, for every 1-form $\omega$ on $U$ there exists $n$ functions $a^i : U \to \mathbb{R}$ such that $\omega$ can be written as a linear combination

$$\omega = \sum_{i=1}^n a_i \, dx^i \ .$$

Given the collection of all cotangent spaces of $M$, we can construct a set called the *cotangent bundle* $T^*M$ given by

$$T^*M = \bigsqcup_{p \in M} T_p^*M \ .$$

This set turns out to be a manifold whose construction is analogous to the tangent bundle's. Without going into details (for a more in depth analysis see Section 17.3 of [4]) we remark the cotangent bundle admits a natural projection map $\pi : T^*M \to M$ and that any chart $(U, \phi) = (U, x^1, \dots, x^n)$ for $M$ induces a chart for $TM$ given by $\tilde{\phi} : T^*U \to \phi(U) \times \mathbb{R}^n$

$$\tilde{\phi}(\alpha) = (\phi \circ \pi, c_1, \dots, c_n)(\alpha)$$

where $c_1, \dots, c_n$ are the coefficient functions of the expansion of $\alpha$ in the basis $(dx^1)_p, \dots,$ $(dx^n)_p$ for $p \in U$. These maps are bijections and $T^*M$ has the topology induced by them from $\phi(U) \times \mathbb{R}^n$.
Since for any vector space $A_1(V) = V^*$ we actually have $T_p^*M = A_1(T_pM)$. We can expand this construction and consider the space of all $k$-covectors on a tangent space $A_k(T_pM)$. The union of this vector spaces as $p \in M$ varies forms a manifold $\bigwedge^k T^*M$ (for further details see Subsection 18.3 in [4]). We talk about *$k$-covector field*, or *$k$-form*, for a map that assigns a $k$-covector $\omega_p$ to each point $p \in M$.

**Definition 1.9.3.** Let $M$ be a manifold. A *frame* for $\bigwedge^k T^*M$ over an open set $U \subset M$ is a collection of $k$-forms $\omega_1, \dots, \omega_n$ on $M$ such that, for every $p \in U$, the $k$-covectors $\omega_1(p), \dots, \omega_n(p)$ form a basis for $A_k(T_pM)$.

**Remark 1.9.4.** Consider a chart $(U.x^1, \dots, x^n)$ for a manifold $M$. Since by Proposition 1.9.2 the covectors $(dx^1)_p, \dots, (dx^n)_p$ for a basis for $T_p^*M$ for all $p \in U$, by Proposition 1.8.4 every $k$-form on $U$ can be written as a linear combination

$$\omega = \sum_I a_I dx^I \ ,$$

where the $I$ are strictly ascending multi-indices and $a_I : U \to \mathbb{R}$. Therefore, the $k$-forms $\left\{dx^I\right\}_{I \in \mathfrak{I}_{k,n}}$ form a frame for $\bigwedge^k T^*M$ over $U$.

If $\omega$ is a $k$-form on a manifold $M$ and $X_1, \dots, X_n$ are vector fields on $M$, we can define a function $\omega(X_1, \dots, X_k)$ by

$$\omega(X_1, \dots, X_k)(p) = \omega_p((X_1)_p, \dots, (X_k)_p) \ . \tag{1.7}$$

**Proposition 1.9.5.** *Let $\omega$ be a $k$-form on a manifold. For any vector fields $X_1, \dots, X_k$ and any function $h : M \to \mathbb{R}$,*

$$\omega(X_1, \dots, hX_i, \dots, X_k) = h\omega(X_1, \dots, X_k) \ .$$

*Proof.* By definition of the function $\omega(X_1, \dots, X_k)$ and the fact that $\omega_p$ is multilinear, the following holds for all $p \in M$

$$\begin{aligned} \omega(X_1, \dots, hX_i, \dots, X_k)(p) &= \omega_p((X_1)_, \dots, h(p)(X_i)_p, \dots, (X_k)_p) \\ &= h(p)\omega_p((X_1)_, \dots, (X_i)_p, \dots, (X_k)_p) \\ &= (h\omega(X_1, \dots, X_i, \dots, X_k))(p) \ . \end{aligned}$$

□

We say that a $k$-form is *smooth* if it is $C^\infty$ as a map between manifolds $\omega : M \to \bigwedge^k T^*M$. The set of all smooth $k$-forms on a manifold $M$ is denoted $\Omega^k(M)$.

**Proposition 1.9.6.** *Let $\omega$ be a $k$-form on a manifold $M$, then the following are equivalent:*

(1) *The $k$-form $\omega$ is smooth on $M$*

(2) *The manifold has an atlas such that on every chart $(U, \phi) = (U, x^1, \ldots, x^n)$ in the atlas, the coefficients $a_I$ of the $\omega = \sum_I a_I dx^I$, relative to the frame $\{dx^I\}_{I \in \mathfrak{J}_{k,n}}$, are all $C^\infty$.*

(3) *On every chart $(U, \phi) = (U, x^1, \ldots, x^n)$ on $M$, the coefficients $a_I$ of the expansion $\omega = \sum_I a_I dx^I$, relative to the frame $\{dx^I\}_{I \in \mathfrak{J}_{k,n}}$, are all $C^\infty$.*

(4) *For any $k$ vector fields $X_1, \ldots, X_k \in \mathfrak{X}(M)$, the function $\omega(X_1, \ldots, X_k)$ is $C^\infty$ on $M$.*

*Proof.* See Proposition 18.7 in [4]. □

We can extend the wedge product between a $k$-covector and an $l$-covector to forms by defining it pointwise. If $\omega$ is a $k$-form and $\tau$ is an $l$-form, we define the $(k + l)$-form $\omega \wedge \tau$ by

$$(\omega \wedge \tau)(p) = \omega_p \wedge \tau_p \ .$$

**Proposition 1.9.7.** *If $\omega$ and $\tau$ are $C^\infty$ forms on $M$, then $\omega \wedge \tau$ is also $C^\infty$.*

*Proof.* See Proposition 18.10 in [4]. □

## 1.10 Exterior and Lie derivatives

In this section we define and briefly characterize some of the main operations involving $k$-forms: the exterior derivative, the interior product and the Lie derivative. The main reference for this section are Sections 19 and 20 of [4].

**Definition 1.10.1.** Let $\omega$ be a smooth $k$-form on a manifold $M$ and let $(U, x^1, \ldots, x^n)$ be a chart about $p \in M$. Suppose $\omega = \sum_I a_I dx^I$ on $U$. We say that the *exterior derivative* $d_U\omega$ of $\omega$ on $U$ is the $(k + 1)$-form

$$d\omega = d_U\omega = \sum_I da^I \wedge dx^I \ .$$

This construction is such that, for every point $p \in M$, $(d\omega)_p$ turns out to be independent of the choice of chart (see Section 19.3 in [4]). Thus, as the point varies the exterior derivative defines a map $d : \Omega^k(M) \to \Omega^{k+1}(M)$.

**Proposition 1.10.2.** *The exterior derivative satisfying the following:*

(1) *It is an antiderivation, meaning that if $\omega \in \Omega^k(M)$ and $\tau \in \Omega^l(M)$, then*

$$d(\omega \wedge \tau) = (d\omega) \wedge \tau + (-1)^{kl} \omega \wedge (d\tau) .$$

(2) $d \circ d = 0$.

(3) *If $f \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$, then $(df)(X) = Xf$.*

*Proof.* See Sections 19.1-19.4 of [4]. □

Consider a smooth map between manifolds $F : N \to M$ and a $k$-form $\omega$ on $M$. Using the differential $F_{*,p} : T_pN \to T_{F(p)}M$ we can naturally induce a $k$-covector at $p$ in N given by

$$(F^*\omega)_p(v_1, \dots, v_k) = \omega_{F(p)}(F_{*,p}(v_1), \dots, F_{*,p}(v_n)) \quad , \quad v_i \in T_pN .$$

As $p \in N$ varies this gives a $k$-form $F^*\omega$ on $N$ called the *pullback* of $\omega$ by $F$.

**Proposition 1.10.3.** *Let $F : N \to M$ be a smooth map of manifolds. The pullback by F satisfies the following properties:*

(1) *If $\omega$ and $\tau$ are $k$-forms and $a, b \in \mathbb{R}$, then*

$$F^*(a\omega + b\tau) = aF^*\omega + bF^*\tau .$$

(2) *If $\omega$ is a $k$-form and $\tau$ is a $l$-form, then*

$$F^*(\omega \wedge \tau) = (F^*\omega) \wedge (F^*\tau) .$$

(3) *If $\omega$ is a smooth $k$-form on $M$, then $F^*\omega$ is a smooth $k$-form on $N$.*

(4) *If $\omega \in \Omega^k(M)$, then $F^*(d\omega) = d(F^*\omega)$.*

*Proof.* See Propositions 18.9, 18.11, 19.5, 19.7 in [4]. □

**Definition 1.10.4.** Let $V$ be a real vector space, $\beta \in A_k(V)$ with $k \geq 2$ and $v \in V$. The *interior multiplication* or *contraction* of $\beta$ with $v$ is the $(k-1)$-covector $\iota_v\beta$, or $v \lrcorner \beta$, defined by

$$(v \lrcorner \beta)(v_2, \dots, v_k) = (\iota_v\beta)(v_2, \dots, v_k) = \beta(v, v_2, \dots, v_k) .$$

We define $\iota_v\beta = \beta(v) \in \mathbb{R}$ if $k = 1$ and $\iota_v\beta = 0$ if k=0.

Interior multiplication on a manifold is defined pointwise, where for any $k$-form $\omega$ and vector field on $M$, the $(k-1)$-form $\iota_X\omega$ is defined by $(\iota_X\omega)_p = \iota_{X_p}\omega_p$. If $X$ and $\omega$ are both $C^\infty$, then by Proposition 1.9.6/4 so is $\iota_X\omega$, since for any smooth vector fields $X_2, \dots, X_n$ the function $\iota_X\omega(X_2, \dots, X_k) = \omega(X_1, X_2, \dots, X_k)$ is smooth. Due to the alternating nature of $k$-forms, it is obvious that $\iota_X \circ \iota_X = 0$ for every vector field $X$.

A collection of $k$-forms $\{\omega_t\}$ is said to be a *one-parameter family* if $t$ runs over a subset of $\mathbb{R}$. Let $t \in I$, where $I$ is an open interval in $\mathbb{R}$, $M$ be a manifold and $\{\omega_t\}$ be a 1-parameter family of smooth $k$-forms on $M$. We say $\{\omega_t\}$ *depends smoothly on* $t$ if every point in $M$ has a coordinate neighborhood $(U, x^1, \dots, x^n)$ where the expansion coefficients $a^I$ of $\omega_t$, relative to frame $\left\{dx^I\right\}_{I\in\mathfrak{I}_{k,n}}$, are smooth functions on $I \times U$. Such a family is also called a *smooth family of k-forms* on $M$. If we consider a smooth family $\{\omega_t\}_{t\in I}$ and chart $(U, x^1, \dots, x^n)$ such that $(\omega_t)_p = \sum_{I\in\mathfrak{I}_{k,n}} b_I(t,p)(dx^I)_p$ for $(t,p) \in I \times U$, we can define the derivative with respect to $t$ to be

$$\left( \left.\frac{d}{dt}\right|_{t_0} \omega_t \right)_p = \sum_I \frac{\partial b_I}{\partial t}(t,p) dx^I|_p \ .$$

The construction is independent of the choice of chart (See Section 20.1 of [4]).

**Definition 1.10.5.** Let $M$ be a manifold, $X \in \mathfrak{X}(M)$, and $\omega \in \Omega^k(M)$. If we let $\Phi : \mathcal{D} \to M$, with $\mathcal{D} \subset \mathbb{R} \times M$, denote the flow of $X$ then we can define the *Lie derivative* $\mathcal{L}_X\omega$ at $p$ to be

$$\mathcal{L}_X\omega = \left.\frac{d}{dt}\right|_0 (\Phi_t^*\omega)_p \ .$$

The definition is well-posed since the smoothness of the flow makes it so that $\Phi_t^*\omega$ depends smoothly on t.

**Theorem 1.10.6.** *Let $X$ be a smooth vector field on a manifold $M$.*

(1) *The Lie derivative $\mathcal{L}_X\omega$ of a smooth $k$-form $\omega$ is a smooth $k$-form.*

(2) *The Lie derivative $\mathcal{L}_X$ is $\mathbb{R}$-linear and if $\omega \in \Omega^k(M)$ and $\tau \in \Omega^l(M)$, then*

$$\mathcal{L}_X(\omega \wedge \tau) = (\mathcal{L}_X\omega) \wedge \tau + \omega \wedge (\mathcal{L}_X\tau) \ .$$

(3) *$\mathcal{L}_X = d\iota_X + \iota_X d$ and $\mathcal{L}_X$ commutes with the exterior derivative $d$.*

*Proof.* See Theorem 20.10 in [4]. □

# Chapter 2

# Elements of Lie theory

In this chapter we state the main results of the theory of Lie groups and Lie algebras needed for our future discussion of symmetries in Hamiltonian systems. For this purpose we define a Lie group and show its tangent space at the identify has a natural Lie algebra structure, which defines the Lie algebra associated with the Lie group. Later, we show how every left-invariant vector field on a Lie group induces a one-parameter subgroup and define the exponential mapping of a Lie group. Finally, we briefly discuss representations and actions of Lie groups and Lie algebras respectively and introduce the notion of Killing vector fields. The main references for this chapter are [3], [4] and [5].

## 2.1 Lie groups and Lie algebras

In this section we introduce Lie groups and define their associated Lie algebras by endowing the tangent space at the identity with a Lie bracket. The main references for this section are Chapter 5 of [3], Chapter 4 of [4] and Section 9.2 of [5].

**Definition 2.1.1.** A *Lie group* is a group $G$ with manifold structure such that the group operations, multiplication

$$\mu : G \times G \to G\ , \quad \mu(a,b) = ab$$

and inverse

$$\iota : G \to G\ , \quad \iota(g) = g^{-1},$$

are $C^\infty$. A map $F : G \to H$ between Lie groups is a *Lie group homomorphism* if it is both $C^\infty$ and a group homomorphism.

We briefly mention the following result which gives a way to get Lie groups out of abstract subgroups of a given Lie group.

**Theorem 2.1.2.** *Let $H$ be a closed (in the topological sense) subgroup of a Lie group $G$. Then, $H$ admits a differentiable structure that makes it into a Lie group.*

*Proof.* See Theorem 5.6.8 in [3]. □

**Example 2.1.3.** In Example 1.2.7 we showed that $\mathrm{GL}(n,\mathbb{R})$ is a manifold and that both the multiplication and inversion maps are smooth. It follows that $\mathrm{GL}(n,\mathbb{R})$ is a Lie group.

We denote the left multiplication by an element $\ell_g : G \to G$, $\ell_g(a) = ga$. This is a smooth map for every $g \in G$ and has a smooth inverse $\ell_{g^{-1}}$, therefore $\ell_g$ is a diffeomorphisms of $G$ onto itself. This map also satisfies

$$\ell_a \circ \ell_b = \ell_{ab} \ . \tag{2.1}$$

Similar facts hold for the right multiplication map $r_g : G \to G$, $r_g(a) = ag$.

**Definition 2.1.4.** A vector field $X$ on a Lie group $G$ is called *left-invariant* if $\ell_{a*}X = X$ for all $a \in G$. The vector space of all left-invariant vector field on $G$ is called $L(G)$.

It can be shown that every left-invariant vector field is also smooth (see Proposition 16.8 in [4]). By Proposition 1.4.6, since every left-invariant vector field is $\ell_g$ related to itself for all $g \in G$, the bracket of two left-invariant vector fields is left-invariant. This makes $L(G)$ into a Lie subalgebra of $\mathfrak{X}(G)$. We now define the mapping

$$(\tilde{\ }) : T_eG \to L(G) \ , \quad \tilde{A}_g = \ell_{g*}A \ . \tag{2.2}$$

This map actually induces an isomorphism between the tangent space at the identity and the left-invariant vector fields, where the inverse is given by the evaluation of the vector field at the identity. Under this isomorphism, we can induce a bracket $[\cdot,\cdot]$ on $T_eG$ as follows. Given $A, B \in T_eG$ we map them to their corresponding left-invariant vector fields $\tilde{A}$ and $\tilde{B}$ and take their Lie bracket $[\tilde{A},\tilde{B}] = \tilde{A}\tilde{B} - \tilde{B}\tilde{A}$. We can evaluate this vector field at the identity to map the bracket back to $T_eG$. Thus, we define the Lie bracket $[A, B] \in T_eG$ to be

$$[A, B] = [\tilde{A},\tilde{B}]_e \ . \tag{2.3}$$

**Proposition 2.1.5.** *If $A, B \in T_eG$ and $\tilde{A},\tilde{B} \in L(G)$ are the left-invariant vector fields they generate, then*

$$[\tilde{A},\tilde{B}] = [A, B]^{\sim} \ .$$

*Proof.* Since $(\tilde{\ })$ and $(\dot{})_e$ are inverse operations we apply $(\tilde{\ })$ to both sides of Equation (2.3)

$$[A, B]^{\sim} = ([\tilde{A},\tilde{B}]_e)^{\sim} = [\tilde{A},\tilde{B}] \ .$$

□

With this bracket $[\cdot,\cdot]$, the tangent space $T_eG$ forms a Lie algebra which we call *Lie algebra of the Lie group* $G$ and denote $\mathfrak{g}$. The proposition above also shows tha the isomorphism $(\tilde{\ }) : T_eG \to L(G)$ is a Lie algebra isomorphism letting us identify $\mathfrak{g}$ with $L(G)$.

**Example 2.1.6** (SO(3) as a Lie group)**.** Consider the matrix group

$$\mathrm{O}(n) = \left\{ A \in \mathrm{GL}(n,\mathbb{R}) \,|\, AA^T = \mathbb{1} \right\} \ .$$

Using the manifold version of the implicit function theorem (Theorem 1.5.4) one can show that this is a manifold of dimension $(n^2-n)/2$ with differentiable structure compatible with the group structure, making $\mathrm{O}(n)$ into a Lie group (see Example 15.7 in [4]). We now compute its Lie algebra which we call $\mathfrak{o}(n)$. We remark that, since $\mathrm{O}(n)$ is contained in Euclidean space, its tangent vector, in particular its Lie algebra, can also be identified with matrices. Consider a curve $c(t)$ in $\mathrm{O}(n)$ such that $c(0) = \mathbb{1}$ and $c'(0) = X \in \mathfrak{o}(n)$. Since the curve is in $\mathrm{O}(n)$ we have

$$c(t)^T c(t) = \mathbb{1} \ ,$$

which we can differentiate with respect to $t$ and get the condition

$$c'(t)^T c(t) + c(t)^T c'(t) = 0 \ .$$

Evaluating this at $t = 0$ yields

$$X + X^T = 0 \ .$$

Since such a curve exists for every $X \in \mathfrak{o}(n)$, we get that $\mathfrak{o}(n)$ is a vector space of dimension $(n^2-n)/2$ contained in the vector space of skew-symmetric $n\times n$ matrices, which we denote $K_n$. Since $\dim(K_n) = (n^2-n)/2$ we get that $\mathfrak{o}(n) = K_n$. Under the identification between matrices and tangent vectors, the bracket is precisely the commutator of matrices (see Proposition 16.11 in [4]).

We now consider $\mathrm{SO}(n) = \{A \in \mathrm{O}(n) \mid \det(A) = 1\}$. It can be shown (see Proposition 9.2.5 in [5]) that $\mathrm{SO}(n)$ is a compact Lie group, which is the connected component of $\mathrm{O}(n)$ containing the identity, and that its Lie algebra, which we denote $\mathfrak{so}(n)$, coincides with $\mathfrak{o}(n)$. In particular, the Lie algebra $\mathfrak{so}(3)$ may be identified with $\mathbb{R}^3$, which becomes a Lie algebra with bracket given by the cross product. Consider the isomorphism

$$\varphi : \mathbb{R}^3 \to \mathfrak{so}(3) \ , \quad \mathbf{x} \mapsto \varphi(\mathbf{x}) = \begin{bmatrix} 0 & -x^3 & x^2 \\ x^3 & 0 & -x^1 \\ -x^2 & x^1 & 0 \end{bmatrix} \ . \tag{2.4}$$

This actually satisfies $\varphi(\mathbf{x})\mathbf{y} = \mathbf{x}\times\mathbf{y}$, which can used to show that $\varphi(\mathbf{x}\times\mathbf{y}) = [\varphi(\mathbf{x}), \varphi(\mathbf{y})]$ (see Section 9.2 in [5]). Thus, $\varphi$ is a Lie algebra isomorphism.

## 2.2 The exponential mapping

In this section we discuss the completeness of left-invariant vector fields, define the exponential mapping of a Lie group and state how it relates to Lie group homomorphisms. The main reference for this section is Section 5.3 of [3].

**Proposition 2.2.1.** *Every left-invariant vector field on a Lie group is complete*

*Proof.* See Proposition 5.3.1 in [3]. □

**Remark 2.2.2.** Let $X \in L(G)$ and let $\Phi$ be its flow. By Proposition 1.7.5/6 and left invariance we get

$$\Phi_{s+t}(\mathbb{1}) = (\Phi_t \circ \Phi_s)(\mathbb{1}) = (\Phi_t \circ \ell_{\Phi_s(\mathbb{1})})(\mathbb{1}) = (\ell_{\Phi_s(\mathbb{1})} \circ \Phi_t)(\mathbb{1}) = \Phi_s(\mathbb{1})\Phi_t(\mathbb{1}) \ ,$$

meaning that the flow of a left-invariant vector field induces an homomorphism between $\mathbb{R}$ and $G$, also called a one-parameter subgroup of $G$.

**Definition 2.2.3.** The *exponential mapping* of a Lie group $G$ is defined by

$$\exp : L(G) \to G \ , \ \ \exp(X) = \Phi_1^X(\mathbb{1}) \ ,$$

where $\Phi^X$ is the flow of $X$. If the Lie group is ambiguous we will write $\exp_G$.

**Remark 2.2.4.** We can use the properties of the flow to characterize the exponential mapping.

1. Since the null vector field $X = 0$ has the trivial flow $\Phi_t^0 = \mathbb{1}$ for all $t \in \mathbb{R}$, we get

$$\exp(0) = \mathbb{1}$$

2. If $X \in L(G)$ and consider the map $t \to \exp(tX)$. By the scaling property of flows, Proposition 1.7.5/4, for all $t \in \mathbb{R}$ we have

$$\exp(tX) = \Phi_t^X(\mathbb{1}) \ ,$$

meaning that the map $t \to \exp(tX)$ yields the maximal integral curve of $X$ through the identity. Since by Remark 2.2.2 the flow of a left-invariant vector field is a group homomorphism, we also get

$$\exp((s+t)X) = \exp(sX)\exp(tX) \ , \quad \exp(-tX) = (\exp(X))^{-1} \ .$$

The following proposition is an important result in the theory of Lie groups and Lie algebras.

**Proposition 2.2.5.** *Let $G$ and $H$ be Lie groups and let $\mathfrak{g}$, $exp_H$ and $\mathfrak{h}$, $exp_H$ denote their associated Lie algebras and exponential mappings. Let $\varphi : G \to H$ be a Lie group homomorphism, then*

(1) *The differential at the identify of $\varphi$, $(d\varphi)_e : \mathfrak{g} \to \mathfrak{h}$, is a Lie algebra homomorphism.*

(2) *We have the following commutative diagram*

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "(d\varphi)_e"] \arrow[d, "exp_G"] & \mathfrak{h} \arrow[d, "exp_H"] \\ G \arrow[r, "\varphi"] & H \end{tikzcd} \tag{2.5}$$

*Proof.* See Propositions 5.2.13 and 5.3.6 in [3]. □

## 2.3 Representations and actions

In this section we define representations of Lie groups and Lie algebras. In particular we define the adjoint and coadjoint representations of a Lie group and its Lie algebra. Later we define actions of Lie groups and Lie algebras, define Killing vector fields and show how they can be used to induce a Lie algebra action starting from a Lie group action. The main reference for this section are Chapters 5 and 6 of [3].

**Definition 2.3.1.** Let $V$ be a finite dimensional real vector space, let $G$ a Lie group and let $\mathfrak{gl}(V)$ be the Lie algebra of $\mathrm{GL}(V)$.

(1) A *representation of* $G$ on $V$ is a Lie group homomorphism $\varphi : G \to \mathrm{GL}(V)$.

(2) A *representation of* $\mathfrak{g}$ on $V$ is a Lie algebra homomorphism $\varphi : \mathfrak{g} \to \mathfrak{gl}(V)$

Every Lie groups $G$ and its Lie algebra $\mathfrak{g}$ admit representations on $\mathfrak{g}$. Consider the conjugation map of the Lie group $c_a = \ell_a \circ r_{a^{-1}}$. Since this is a diffeomorphism of $G$ onto itself, its differential at the identity $dc_a := c_{a\,*,e} : \mathfrak{g} \to \mathfrak{g}$ is a linear automorphism, $dc_a \in GL(\mathfrak{g})$. As $a \in G$ varies we get a mapping

$$\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g}) \ , \quad \mathrm{Ad}(a) := dc_a \ .$$

This is a group homomorphism which can be shown to be smooth, making it into a Lie group homomorphism (see Section 5.4 in [3]), making Ad into a representation of G on $\mathfrak{g}$. The natural representation of every Lie algebra onto itself is given by

$$\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \ , \quad \mathrm{ad}(X)Y = [X, Y] \text{ for all } Y \in \mathfrak{g} \ .$$

The Jacobi identify implies $\mathrm{ad}([X,Y]) = [\mathrm{ad}(X), \mathrm{ad}(Y)]$, hence this is a Lie algebra homomorphism and a representation.

**Definition 2.3.2.** The representations Ad and ad are called the *adjoint representations* of $G$ and $\mathfrak{g}$ respectively.

**Definition 2.3.3.** Let $G$ be Lie group and $\mathfrak{g}$ its Lie algebra. We call *coadjoint representations* of $G$ and $\mathfrak{g}$, and denote $\mathrm{Ad}^*$ and $\mathrm{ad}^*$, the representations of $G$ and $\mathfrak{g}$ on $\mathfrak{g}^*$ defined by

1. $\langle \mathrm{Ad}^*(a)\xi, Y\rangle = \langle \xi, \mathrm{Ad}(a^{-1})(Y)\rangle$ for all $a \in G$, $Y \in \mathfrak{g}$ and $\xi \in \mathfrak{g}^*$.
2. $\langle \mathrm{ad}^*(X)\xi, Y\rangle = -\langle \xi, \mathrm{ad}(X)(Y)\rangle = -\langle \xi, [X,Y]\rangle$ for all $X, Y \in \mathfrak{g}$ and $\xi \in \mathfrak{g}^*$.

**Definition 2.3.4.** An *action* of a Lie group $G$ on a manifold $M$ is a smooth mapping $\Psi : G \times M \to M$, such that the induced mapping

$$\Psi_a : M \to M \ , \quad \Psi_a(m) = \Psi(a,m) \ ,$$

satisfies $\Psi_e = \mathrm{id}_M$ and either $\Psi_a \circ \Psi_b = \Psi_{ab}$ or $\Psi_a \circ \Psi_b = \Psi_{ba}$. In the first case $\Psi$ is called a *left action* while it is called a *right action* in the second case. The triple $(M, G, \Psi)$ is referred to as a *Lie group action* and the pair $(M, \Psi)$ is called a *$G$-manifold.*
A morphism of Lie group actions $(M_1, G_1, \Psi^1)$ and $(M_2, G_2, \Psi^2)$ (both left or right) consists of a smooth map $\varphi : M_1 \to M_2$ and a Lie group homomorphism $\rho : G_1 \to G_2$ such that

$$\varphi \circ \Psi^1 = \Psi^2 \circ (\rho \times \varphi) \ .$$

If $G_1 = G_2$ and $\rho$ is the identical mapping $\varphi$ is called *equivariant* or *morphism of $G$-manifolds.*

**Example 2.3.5.** Every representation $\varphi : G \to \mathrm{GL}(V)$ of $G$ on a finite dimensional vector space $V$ defines a left action of $G$ on $V$ by $\Psi(a,v) = \varphi(a)v$ for all $a \in G$ and $v \in V$. In particular the adjoint and coadjoint representations of $G$ induce actions on $\mathfrak{g}$, called the adjoint and coadjoint action respectively.

We recall that a map between locally compact Hausdorff spaces is called *proper* if it is continuous and the preimages of compact sets are compact.

**Definition 2.3.6.** Let $(M, G, \Psi)$ be a Lie group left action.

1. Two point $m_1, m_2 \in M$ are said to be *conjugate* $m_2 = \Psi_a(m_1)$ for some $a \in G$. This forms an equivalence relation on $M$ whose equivalence classes are called the *orbits* of $G$ under $\Psi$. The orbit thorugh a given point $m$ is denoted $G \cdot m$. The set of equivalence classes, equipped with the quotient topology, is called *orbit space* of $(M, G, \Psi)$ and denoted $M/G$.
2. For $m \in M$, the subgroup $G_m = \{a \in G : \Psi_a(m) = m\}$ is called the *stabilizer* of $m$. The action is said to be *free* is all stabilizers are equal to $\{\mathbb{1}\}$.

3. $(M, G, \Psi)$ is called a *proper* Lie group action if the following mapping is proper

$$\Psi_{\text{ext}} : G \times M \to M \times M \ , \quad \Psi_{\text{ext}}(a, m) := (\psi_a(m), m) \ .$$

**Proposition 2.3.7.** *The orbit space $M/G$ of a free proper Lie group action $(M, G, \Psi)$ admits a unique smooth structure such that the natural projection $\pi : M \to M/G$ is a submersion.*

*Proof.* For the proof see Corollary 6.5.1 in [3]. □

Let $(M, G, \Psi)$ be a Lie group action and let $\mathfrak{g}$ be the Lie algebra of $G$. For $A \in \mathfrak{g}$, we have

$$\Psi_{\exp(tA)} \circ \Psi_{\exp(sA)} = \Psi_{\exp(tA)\exp(sA)} = \Psi_{\exp((s+t)A)} \ .$$

Therefore, the assignment $(m, t) \to \Psi_{\exp(tA)(m)}$ defines a one parameter group of diffeomorphism, which is a complete flow.

**Definition 2.3.8.** Let $(M, G, \Psi)$ be a Lie group action and let $\mathfrak{g}$ be the Lie algebra of $G$. The vector field $A_*$ on $M$ is said to be the *Killing vector field* generated by $A \in \mathfrak{g}$ if it corresponds to the flow $\Psi_{\exp(tA)}$.

The Killing vector field field's value at a point $m \in M$ may be expressed as

$$(A_*)_m = \frac{d}{dt}\bigg|_0 \Psi_{\exp(tA)}(m) = \frac{d}{dt}\bigg|_0 \Psi_m(\exp(tA)) = (\Psi_m)_*(A) \ .$$

Killing vector fields have the following properties.

**Proposition 2.3.9.** *Let $(M, G, \Psi)$ be a Lie group action.*

(1) *For every $a \in G$ there holds $\Psi_{a*}A_* = (Ad(a)A)_*$ if $\Psi$ is a left action and $\Psi_{a*}A_* = (Ad(a^{-1})A)_*$ if $\Psi$ is a right action.*

(2) *The mapping $\mathfrak{g} \to \mathfrak{X}(M)$ given by $A \mapsto A_*$ is an antihomomorphism of Lie algebras if $\Psi$ is a left action and a homomorphism is $\Psi$ is a right-action.*

(3) *Let $(M_i, G_i, \Psi^i)$, $i = 1, 2$, be Lie group actions and let $\varphi : M_1 \to M_2$, $\rho : G_1 \to G_2$ define a morphism between them. Then, the Killing vector field on $M_1$ generated by an element $A$ in the Lie algebra of $G_1$ is $\varphi$-related to the Killing vector field generated by $d\rho(A)$ (where $d\rho$ denotes the differential at the identity).*

(4) *Let $(M_i, \Psi^i)$, $i = 1, 2$, be $G$-manifolds and let $\varphi : M_1 \to M_2$ be equivariant. For every $A \in \mathfrak{g}$, the Killing vector fields $A_*^{M_i}$ generated by $A$ are $\varphi$-related.*

*Proof.* See Propositions 6.2.2 and 6.24 in [3]. □

**Definition 2.3.10.**

1. An *action of a Lie algebra* $\mathfrak{g}$ on a manifold $M$ is a mapping $\psi : \mathfrak{g} \to \mathfrak{X}(M)$ which is either an antihomomorphism or an homomorphism of Lie algebras. In the first case $\psi$ is called a left action and in the second it is called a right action. The triple $(M, \mathfrak{g}, \Psi)$ is referred to as a Lie algebra action and the pair $(M, \Psi)$ as a $\mathfrak{g}$-manifold.

2. A *morphism of Lie algebra actions* $(M_i, \mathfrak{g}_i, \psi_i)$, $i = 1, 2$, consists of a smooth mapping $\varphi : M_1 \to M_2$ and a Lie algebra homomorphism $\varrho : \mathfrak{g}_1 \to \mathfrak{g}_2$ such that the vector fields $\psi_1(A)$ and $\psi_2(\varrho(A))$ are $\varphi$-realted for all $A \in \mathfrak{g}_1$. If $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathfrak{g}$ and $\varrho = \mathbb{1}$, $\varphi$ is called *equivariant* or a *morphism of* $\mathfrak{g}$*-manifolds.*

By Proposition 2.3.9 we get the following Corollary.

**Corollary 2.3.11.**

(1) *The assignment of Killing vector fields, $A \mapsto A_*$, of a left or right Lie group action $(M, G, \Psi)$ induces a left or right action, respectively, of the Lie algebra of $G$, $\mathfrak{g}$, on $M$.*

(2) *If $(\varphi, \rho)$ is a morphism of Lie group actions, then $(\varphi, d\rho)$ is a morphism of the corresponding Lie algebra actions. If $\varphi$ is a morphism of $G$-manifolds, it is also a morphism of the corresponding $\mathfrak{g}$-manifolds.*

# Chapter 3

# Symplectic geometry

In this chapter we introduce the main concepts of symplectic geometry. For this reason, we first start by talking about symplectic vector spaces, which are to symplectic manifolds what regular Euclidean space is to differentiable manifolds. We then focus on symplectic manifolds, show how the symplectic structure induces Hamiltonian vector fields and the Poisson bracket and provide an important example of a symplectic manifold: the cotangent bundle. Finally, we study symmetries by defining momentum maps and proving the regular symplectic reduction theorem. The main reference for this chapter are [3] and [6].

## 3.1 Symplectic linear algebra

In this section we introduce symplectic vector spaces. These are even dimensional vector spaces equipped with an antisymmetric, non-degenerate bilinear form. The main reference for this section is Chapter 7 of [3].

Let $V$ a finite-dimensional vector space over $\mathbb{R}$ and consider an antisymmetric bilinear form $\omega \in A_2(V)$. This map induces the linear mapping

$$\omega^\flat : V \to V^* \ , \quad \langle \omega^\flat(v), u \rangle = (\iota_v \omega)(u) = \omega(v, u) \ , \ \ \forall v, u \in V \ .$$

We define the *kernel* and the *rank* of $\omega$ to be the kernel and rank of $\omega^\flat$, respectively, and say that $\omega$ is *non-degenerate* if $\omega^\flat$ is an isomorphism. If $\omega^\flat$ is an isomorphism we denote its inverse by $\omega^\sharp$.
For a subspace $W \subset V$, we define the $\omega$-orthogonal subspace

$$W^\omega = \{v \in V \, : \, \omega(v, w) = 0 \text{ for all } w \in W\}$$

**Remark 3.1.1.** By Proposition 1.8.4, if we let $\{e_i\}$ and denote a basis in $V$ and $e^{*i}$ be its dual basis in $V^*$, then

$$\omega = \sum_{i,j} \frac{1}{2}\omega_{ij} e^{*i} \wedge e^{*j} \ , \quad \omega^\flat(e_i) = \omega_{ij} e^{*j} \ , \quad \mathrm{rank}(\omega) = \mathrm{rank}(\omega_{ij}) \ ,$$

where $\omega_{ij} = \omega(e_i, e_j)$. If we let $\omega$ be non-degenerate it follows that $\mathrm{rank}(\omega) = \dim(V)$ is even. To show this let $\dim(V) = n$, then

$$\det(\omega_{ij}) = \det(-\omega_{ji}) = (-1)^n \det(\omega_{ji}) = (-1)^n \det(\omega_{ij}) \ .$$

Since we assumed $\omega$ to be non-degenerate, $\det(\omega_{ij}) \neq 0$. It follows that $n$ is even.

The following theorem characterizes antisymmetric bilinear form by a procedure akin to the Gram-Schmidt process.

**Theorem 3.1.2.** *For every antisymmetric bilinear form $\omega$ on a finite dimensional real vector space $V$ there exists an ordered basis $\{e_i\}$ in $V$ and an $n \in \mathbb{N}$ such that*

$$\omega = \sum_{i=1}^{n} e^{*i} \wedge e^{*(i+n)} \ ,$$

*that is, the matrix of coefficients of $\omega$, $\omega_{ij}$, has the form*

$$\omega_{ij} = \begin{bmatrix} J_n & 0 \\ 0 & 0 \end{bmatrix} \ ,$$

*where $J_n$ is given by*

$$J_n = \begin{bmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{bmatrix} \ .$$

*Proof.* Let $V_0 \coloneqq V$. If $\omega$ is zero on $V_0$, any basis of $V$ will yield the desired result. Otherwise, there exist $v_1$ and $u_1$ in $V$ such that $\omega(v_1, u_1) = 1$. By bilinearity both vectors are non zero, meanwhile, by antisymmetry, $v_1$ and $u_1$ are not parallel; therefore they are linearly independent. Consider now $E_1$, the vector space given by their span. We claim that

$$E_1 \cap E_1^\omega = 0 \ , \quad E_1 + E_1^\omega = V_0 \ .$$

To show the first equation consider a vector $v \in E_1 \cap E_1^\omega$. This can be decomposed as $v = \alpha v_1 + \beta u_1$ but, since $v \in E_1^\omega$, we have $\alpha = \omega(v, u_1) = 0$ and $\beta = -\omega(v, v_1) = 0$. For the second equation we consider $v \in V_0$ and write

$$v = (\omega(v, u_1)v_1 - \omega(v, v_1)u_1) + (v - \omega(v, u_1)v_1 + \omega(v, v_1)u_1) \ .$$

Since the second term is in $E_1^\omega$, the second equation follows.
We now replace $V_0$ by $V_1 = E_1^\omega$ and iterate the process until we arrive at a, possibly trivial, subspace $V_n \subset V$ on which $\omega$ vanishes. This process yields $2n$ linearly independent vectors $e_1 := v_1$, $e_{i+n} := u_i$, $i = 1, \dots, n$ that satisfy

$$\omega(e_i, e_j) = \omega(e_{i+n}, e_{j+n}) = 0 \ , \quad \omega(e_i, e_{j+n}) = \delta_{ij} \ , \quad \text{for all } i, j = 1, \dots, n \ .$$

If $V_n$ is trivial these 2n vectors form the desired basis in V. If $V_n \neq \{0\}$, we can choose a basis $\left\{e_{2n+1}, \dots, e_{\dim(V)}\right\}$ in $V_n$ and take $\left\{e_1, \dots, e_{\dim(V)}\right\}$ to be the basis in V with the desired properties. □

**Definition 3.1.3.**

(1) A *symplectic vector space* is a pair $(V, \omega)$ consisting of a real vector space $V$ and antisymmetric non-degenerate bilinear form $\omega$. A basis in $V$ for which $\omega$ has the canonical form of Theorem 3.1.2 is called *symplectic* or *canonical.*

(2) Let $(V_1, \omega_1)$ and $(V_2, \omega_2)$ be two symplectic vector spaces. A linear map $f : V_1 \to V_2$ is called *symplectic* if $f^*\omega_2 = \omega_1$. A bijective symplectic mapping is called a *symplectomorphism.*

We note that the inverse of a symplectic map is also symplectic, making symplectomorphisms the isomorphisms corresponding to the morphisms of symplectic vector spaces given by symplectic mappings.

**Example 3.1.4.** Let $V = \mathbb{R}^{2n}$. When equipped the bilinear form $\omega_0(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T J_n \mathbf{v}$, $(\mathbb{R}^{2n}, \omega_0)$ clearly defines a symplectic vector space where the standard basis is symplectic. We call $J_n$ the *standard symplectic matrix* of $\mathbb{R}^{2n}$ and the pair $(\mathbb{R}^{2n}, \omega_0)$ the *canonical symplectic vector space* in dimension $2n$.

From this example, we get an immediate corollary of Theorem 3.1.2.

**Corollary 3.1.5.** *Let $(V, \omega)$ be a symplectic vector space of dimension $2n$. Every symplectic basis of $V$ defines a symplectomorphism onto the canonical symplectic vector space $(\mathbb{R}^{2n}, \omega_0)$.*

We now focus our attention to the properties of a particular kind of subspace of a symplectic vector space.

**Definition 3.1.6.** A subspace $W$ of a symplectic vector space $(V, \omega)$ is called *Lagrangian* if $W = W^\omega$.

It can be shown (see Proposition 7.2.6 in [3]) that every Lagrangian subspace admits a Lagrangian complement, i.e. for every Lagrangian subspace $W \subset V$ there exists a subspace $W' \subset V$ that is also Lagrangian and such that $V = W \oplus W'$.

It turns out that, given a vector space $W$, the vector space $V = W \oplus W^*$ has a symplectic structure given by the form

$$\omega_{W \oplus W^*}(v \oplus \rho, u \oplus \sigma) = \rho(u) - \sigma(v) \ .$$

This form is referred to as the canonical symplectic form on $W \oplus W^*$. For further detail see Section 7.1 of [3]. We can now state the following.

**Proposition 3.1.7.** *Let $(V, \omega)$ be a symplectic vector space and let $V = W \oplus W'$ be a decomposition into complementary Lagrangian subspaces. Then, this composition induces a symplectomorphism onto the canonical symplectic vector space $(W \oplus W^*, \omega_{W \oplus W^*})$.*

*Proof.* For the proof see Proposition 7.2.9 in [3]. □

**Example 3.1.8.** A particularly interesting example is given by the following decomposition of $\mathbb{R}^{2n}$. To describe this we can decompose $\mathbb{R}^{2n}$ into the two Lagrangian complements $\mathbb{R}^n \oplus \{0\}$ and $\{0\} \oplus \mathbb{R}^n$. By the proposition above, this induces a symplectomorphism from $\mathbb{R}^{2n}$ to $\mathbb{R}^n \oplus (\mathbb{R}^n)^*$. While this decompoistion does not seem computationally advantageous, it is physically interesting as we will see later on.

Finally, we end this section by showing a necessary property for a linear map to be a symplectomorphism.

**Proposition 3.1.9.** *Let $(V, \omega)$ and $(W, \rho)$ be symplectic vector spaces and let $f : V \to W$ be a symplectic mapping. Then, the mapping is injective. In particular, if dim$(V)$ = dim$(W)$, $f$ is a symplectomorphism.*

*Proof.* Let $v \in V$ be such that $f(v) = 0$, then, since the mapping is symplectic, we have

$$\omega(v, u) = (f^*\rho)(v, u) = \rho(f(v), f(u)) = 0$$

for all $u \in V$. Since $\omega$ is non-degenerate, it follows that $v = 0$. □

## 3.2 Symplectic geometry

In this section we introduce and characterize symplectic manifolds, a class of differentiable manifolds equipped with a symplectic form, which is a 2-form that makes all the tangent spaces of the manifold into symplectic vector spaces. The main reference for this section is Section 8.1 of [3].

**Definition 3.2.1.** Let $M$ be a manifold and $\omega$ be a smooth 2-form on $M$. Every such form induces a map $\omega^\flat : TM \to T^*M$ defined by

$$\langle \omega^\flat(X), Y \rangle = \omega_p(X, Y) \ , \quad X, Y \in T_pM \ .$$

We define the *kernel* and the *rank* of of $\omega$ to be the kernel and rank of $\omega^\flat$ respectively. The 2-form $\omega$ is said to be *non-degenerate* if $\omega_p \in A_2(T_pM)$ is non-degenerate for all points $p \in M$, otherwise it is called *degenerate*.

When $\omega$ is non-degenerate, $\omega^\flat$ is actually an *isomorphism of vector bundles*. For our purposes, this just means that $\omega^\flat$ is a diffeomorphism between $TM$ and $T^*M$ which is fibrewise linear, it satifies $\pi_{TM} = \pi_{T^*M} \circ \omega^\flat$ and its inverse posses analogue properties (for further details on vector bundles and their morphisms see Section 2.2 of [3]).
In the following, whenever $\omega$ is non-degenerate, we denote the inverse of $\omega^\flat$ by $\omega^\sharp$. We will also occasionally write $X^\flat = \omega^\flat(X)$ and $\alpha^\sharp = \omega^\sharp(\alpha)$ for $X \in \mathfrak{X}(M)$ and $\alpha \in \Omega^1(M)$.

**Definition 3.2.2.**

1. A *symplectic manifold* is a pair $(M, \omega)$ consisting of a differentiable manifold and a closed non-degenerate 2-form $\omega$, called the *symplectic form* or the *symplectic structure*.
2. A smooth mapping $\Phi : M \to N$ between symplectic manifolds $(M, \omega)$ and $(N, \rho)$ is called *symplectic* or *canonical* if $\Phi^*\rho = \omega$. If $\Phi$ is a (local) diffeomorphism it is called a (local) *symplectomorphism*.

If $(M, \omega)$ is a symplectic manifold, then every tangent space forms a symplectic vector space $(T_pM, \omega_p)$. In particular, this shows that every symplectic manifold must be even dimensional. As for symplectic mappings, every such map $\Phi : M \to N$ is symplectic if and only if its differential $\Phi_{*,p} : T_pM \to T_{\Phi(p)}N$ is symplectic for all points. By Proposition 3.1.9 it follows that every symplectic mapping is also an immersion. Just as it was for the symplectic maps in the sense of symplectic vector spaces, we have that the inverse of a symplectic map is also symplectic.

**Example 3.2.3.** Every symplectic vector space $(V, \omega)$ naturally induces a corresponding symplectic manifold given by the standard differentiable structure on $V$ and the identification $T_pV \simeq V$ by which $\omega$ can be viewed as a 2-form. In particular, this makes the canonical symplectic vector space $(\mathbb{R}^{2n}, \omega_0)$ into a symplectic manifold.

The following theorem can be viewed the manifold version of Corollary 3.1.5, generalizing it to arbitrary spaces but losing the global aspect. It also tells us that $(\mathbb{R}^{2n}, \omega_0)$ is the prototypical $2n$ dimensional symplectic manifold.

**Theorem 3.2.4** (Darboux)**.** *Let $(M, \omega)$ be symplectic manifold of dimension $2n$. Then, for every $m \in M$, there exist a symplectomorphism from a neighborhood of $m$ onto an open subset of the canonical symplectic vector space $(\mathbb{R}^{2n}, \omega_0)$.*

*Proof.* See Theorem 8.1.5 in [3]. □

Consider now a symplectic manifold $(M, \omega)$. By Proposition 1.2.9, we know that every diffeomorphism onto a subset of Euclidean space can be used as a chart. We can apply this result to the symplectomorphisms given by Darboux's theorem to construct charts $(U, \phi) = (U, x^1, \ldots, x^{2n})$ such that $\omega$ takes the form

$$\omega|_U = \sum_{i=1}^{n} dx^i \wedge dx^{i+1} .$$

Such a chart takes the name of a *Darboux chart* while the coordinates are referred to as *Darboux coordinates*. By the theorem above a Darboux chart is guaranteed to exist about every point $p \in M$. As we will soon see, having such a simple expression for the symplectic form is extremely useful when doing local computations in Hamiltonian dynamics.
Since we are mostly concerned with physical applications, we can use the symplectomorphism between $(\mathbb{R}^{2n}, \omega_0)$ and the canonical symplectic structure on $\mathbb{R}^n \oplus (\mathbb{R}^n)^*$ to construct a new Darboux chart. If we compose this last symplectomorphism with $\phi$ we get a new chart such that, if we denote $q^i$ and $p_i$ the coordinate functions corresponding to the factors $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$, the symplectic form can be expressed as

$$\omega|_U = \sum_{i=1}^{n} dp_i \wedge dq^i . \tag{3.1}$$

The significance of this last expression comes from the fact that, as we will see in the next chapter, the phase space of an Hamiltonian system can naturally be expressed as a symplectic manifold. Thus, this decomposition shows how the symplectic structure of an arbitrary phase space is locally symplectomorphic to the phase space of a point particle of appropriate dimension.

## 3.3 Hamiltonian vector fields

In this section we shift our attention to Hamiltonian vector fields, a class of vector fields generated by the corresponding Hamiltonian functions via the symplectic form. The main reference for this section is Section 8.2 of [3].

The non-degeneracy of the symplectic form can be used to assign a vector field to every smooth function via the isomorphism $\omega^\sharp : T^*M \to TM$.

**Definition 3.3.1.** Let $(M, \omega)$ be a symplectic manifold and let $f \in C^\infty(M)$. The vector field

$$X_f = -(df)^\sharp = -\omega^\sharp(df)$$

is called the *Hamiltonian vector field* generated by $f$. The generating function is sometimes called the *Hamiltonian.* The set of all Hamiltonian vector field on $(M, \omega)$ is denoted $\mathfrak{X}_H(M, \omega)$.

We refer to the flow of $X_f$ as the *Hamiltonian flow* generated by $f$. Given an Hamiltonian function its corresponding vector field is implicitly defined by

$$(X_f)^\flat = X_f \lrcorner \omega = -df$$

In particular, Hamiltonian vector fields are in a 1-1 correspondence with exact forms via the isomorphism $\omega^\flat$.

**Example 3.3.2.** By using Darboux coordinates we can find a local expression for Hamiltonian vector fields. Consider a symplectic manifold $(M, \omega)$ of dimension $2n$ and let $f$ be the Hamiltonian function. In Darboux coordinates $q^i$ and $p_i$, the vector field $X_f$ can be expressed with respect to the local frame

$$X_f = \sum_{i=1}^{n} A_i \frac{\partial}{\partial p_i} + B^i \frac{\partial}{\partial q^i} \ ,$$

where $A_i$ and $B_i$ are the $2n$ coefficient functions determined by the condition $X_f \lrcorner \omega = -df$. By using Equation (3.1) we can compute the interior product to get

$$X_f \lrcorner \omega = (\sum_{i=1}^{n} A_i \frac{\partial}{\partial p_i} + B^i \frac{\partial}{\partial q^i}) \lrcorner (\sum_{j=1}^{n} dp_j \wedge dq^j) = \sum_i A_i dq^i - B^i dp_i \ .$$

By expanding $df$ in local coordinates we the relation

$$\sum_i A_i dq^i - B^i dp_i = \sum_i -\frac{\partial f}{\partial q^i} dq^i - \frac{\partial f}{\partial p_i} dp_i$$

which implies

$$X_f = \sum_{i=1}^{n} \frac{\partial f}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i} \frac{\partial}{\partial p_i}$$

Another important class of vector fields are the so called infinitesimal symmetries of the symplectic structure.

**Definition 3.3.3.** A vector field $X$ on $(M, \omega)$ is called *symplectic* if the symplectic form $\omega$ is invariant under its flow. The set of all symplectic vector fields on $(M, \omega)$ is denoted $\mathfrak{X}_{LH}(M, \omega)$.

**Proposition 3.3.4.** *The let $(M, \omega)$ be a symplectic manifold and let $X \in \mathfrak{X}(M)$. The following statements are equivalent:*

(1) *$X$ is symplectic*

(2) $\mathcal{L}_X \omega = 0$

(3) *$X^\flat$ is closed*

(4) *$X$ is locally Hamiltonian, meaning that for every $p \in M$ there exist a function on a neighborhood $U$ of $p$ such that $X|_U = X_f$.*

*Proof.* To show the equivalence of (2) and (3) we can use Theorem 1.10.6/3 and the fact that $d\omega = 0$ to get

$$\mathcal{L}_X\omega = X \lrcorner d\omega + d(X \lrcorner \omega) = dX^\flat \; .$$

To show the equivalence between (1) and (2) we let $\Phi$ denote the flow of $X$ and observe that

$$\frac{d}{ds}(\Phi_s^*\omega)_m = (\Phi_s^*(\mathcal{L}_X\omega))_m$$

for all $s \in (0, t)$ and all points $m$ in the domain of $\Phi_t$. If $X$ is symplectic, $\Phi_s^*\omega = \omega$ for all $s$, meaning that the derivative vanishes and implying (2). If (2) holds, we can integrate from 0 to $t$ and use the condition $\Phi_0 = \mathrm{id}$ to get the desired result.
Finally, the equivalence between (4) and (3) follows from the Poincaré Lemma. For further details see Proposition 8.2.4 in [3]. □

**Remark 3.3.5.** Every Hamiltonian field $X_f$ is such that $(X_f)^\flat$ is exact, therefore it is also closed. It follows that every Hamiltonian vector field is automatically symplectic.
The converse doesn't hold in general, however depending on the topology of the manifold, it is sometimes true that every closed form is exact guaranteeing that every symplectic vector field is also symplectic. For further details see Chapter 7 in [4].

**Proposition 3.3.6.** *Let $(M, \omega)$ be a symplectic manifold.*

(1) *$\mathfrak{X}_{LH}(M)$ is a Lie subalgebra of $\mathfrak{X}(M)$.*

(2) *For $Y, Z \in \mathfrak{X}_{LH}(M)$ we have that*

$$[Y, Z] = X_{\omega(Y,Z)} \; .$$

*Proof.* See Proposition 8.2.6 in [3]. □

In particular, since every Hamiltonian vector field is symplectic, the bracket of two Hamiltonian vector fields is an Hamiltonian vector field as well. This makes $\mathfrak{X}_H(M, \omega)$ into a Lie algebra.
We end this section with a result that shows how symplectic mappings are characterized in terms of how they act on Hamiltonian vector fields.

**Theorem 3.3.7.** *A smooth mapping $\Phi : M \to N$ of symplectic manifolds $(M, \omega)$ and $(N, \rho)$ is symplectic if and only if it is an immersion and the Hamiltonian vector fields*

*$X_{\Phi^* f}$ on $M$ and $X_f$ on $N$ are $\Phi$-related for all $f \in C^\infty$.*
*In particular, if $\Phi$ is a diffeomorphism it is a symplectomorphism if and only if*

$$\Phi_* X_h = X_{h \circ \Phi^{-1}} \ , \quad \textit{for all } h \in C^\infty(M) \ .$$

*Proof.* See Proposition 8.2.9 in [3]. □

## 3.4 Poisson bracket

In this section we define the Poisson bracket and prove its main properties. The main reference is Section 8.2 of [3].

In the previous section we showed that the set of Hamiltonian vector fields is a Lie algebra. Since these vector fields are generated by smooth functions it makes sense to ask if there is a Lie algebra structure on $C^\infty(M)$ that makes the mapping $f \to X_f$ into a Lie algebra homomorphism.

**Definition 3.4.1.** Let $(M, \omega)$ be a symplectic manifold and let $f, h \in C^\infty(M)$. The function

$$\{f, g\} \coloneqq \omega(X_f, X_g)$$

is called the *Poisson bracket* of $f$ and $h$.

**Remark 3.4.2.** Since by construction we have $\omega(X_f, X_g) = -X_f \lrcorner (X_g \lrcorner \omega) = X_f \lrcorner (dg) = X_f(g)$, the Poisson bracket can be rewritten as

$$\{f, g\} = X_f(g) = -X_g(f) \ . \tag{3.2}$$

**Proposition 3.4.3.** *Let $(M, \omega)$ be a symplectic manifold.*

(1) *The Poisson bracket makes $C^\infty(M)$ into a Lie algebra. With respect to this structure, the mapping $f \to X_f$ is a Lie algebra homomorphism onto $\mathfrak{X}_H(M)$.*

(2) *The Poisson bracket satisfies the Leibniz rule, that is, for all $f, g, h \in C^\infty(M)$ we have*

$$\{f, gh\} = \{f, g\}\, h + \{f, h\}\, g \ .$$

*Proof.* 1. Unpacking the various definitions, it follows that the Poisson bracket is antisymmetric and bilinear and that the assignment $f \to X_f$ is linear. Meanwhile, by Proposition 3.3.6 it is also true that

$$[X_f, X_g] = X_{\omega(X_f, X_g)} = X_{\{f,g\}} \ .$$

By this and Equation (3.2) we can show the Poisson bracket satisfies the Jacobi identity, since

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = X_f(X_g(h)) - X_g(X_f(h)) - X_{\{f,g\}}(h) = 0 \ ,$$

making $C^\infty$ into a Lie algebra and $f \to X_f$ into a Lie algebra homomorphism.
2. It follows from Equation (3.2) and that vector fields are derivations on $C^\infty(M)$. □

We end this section by showing how symplectic mappings can be characterized in terms of the Poisson brackets.

**Proposition 3.4.4.** *A smooth mapping $\Phi : M \to N$ of symplectic manifolds $(M, \omega)$ and $(N, \rho)$ is symplectic if and only if it is an immersion and for all $f, g \in C^\infty(N)$ one has*

$$\{\Phi^* f, \Phi^* g\} = \Phi^* \{f, g\} \ .$$

*Proof.* See Proposition 8.2.15 in [3]. □

## 3.5 Liouville form and the cotangent bundle

In this section we show how the cotangent of every manifold carries a canonical symplectic structure given with symplectic potential given by the Liouville form. The main references for this section are Section 8.3 of [3] and Chapter 2 of [6].

Let $Q$ be a smooth manifold. We want to show that $T^*Q$ not only can always be given a symplectic structure, but that it admits a *symplectic potential*, meaning there exists a 1-form $\theta$ such that $\omega = d\theta$ is a symplectic form on $T^*Q$.
In the following, we consider the coordinates $(T^*U, q^1, \dots, q^n, p_1, \dots, p_n)$ for the cotangent bundle.

**Definition 3.5.1.** The *canonical 1-form $\theta$*, or *Liouville form*, on $T^*Q$ is defined by

$$\langle \theta_\xi, X \rangle \coloneqq \langle \xi, \pi_*(X) \rangle \ , \tag{3.3}$$

where $\xi \in T^*Q$, $X \in T(T^*Q)$ and $\pi : T^*Q \to Q$ is the natural projection.

We note that the canonical 1-form on $T^*Q$ might also be written as

$$\theta_\xi = (\pi_{*,\xi})^* \xi = \xi \circ \pi_{*,\xi} \ ,$$

where $\xi \in T_p^*Q$ for some $p \in Q$.

We claim that the 2-form $\omega \coloneqq d\theta$ is a symplectic form on $T^*Q$. Consider a tangent vector $X \in T(T^*Q)_\xi$. In coordinates, we will have

$$\theta = \sum_{i=1}^n \alpha_i dq^i + \beta^i dp_i \ , \qquad X = \sum_{i=1}^n A^i \frac{\partial}{\partial q^i}\bigg|_\xi + B_i \frac{\partial}{\partial p_i}\bigg|_\xi \ .$$

Then, the defining equation of $\theta$ becomes

$$A^i\alpha_i(\xi) + B_i\beta^i(\xi) = p_i(\xi)A^i \ .$$

Therefore, the local expressions for the canonical 1-form and the 2-form $\omega$ are

$$\theta = \sum_{i=1}^{n} p_i dq^i \ , \quad \omega = \sum_{i=1}^{n} dp_i \wedge dq^i \ . \tag{3.4}$$

From this is follows that $\theta$ and $\omega$ are smooth and that $\omega$ is non-degenerate. Furthermore, $\omega$ is clearly closed since $d\omega = d^2\theta = 0$, proving the initial claim.

**Definition 3.5.2.** The 2-form $\omega = d\theta$ is called the *canonical symplectic structure* on $T^*Q$.

**Remark 3.5.3.** The local expression Equation (3.4) for $\theta$ is enough to intrinsically define it on $T^*Q$. Consider two charts, $(T^*U, q^1, \dots, q^n, p_1, \dots, p_n)$ and $(T^*U', q'^1, \dots, q'^n, p'_1, \dots, p'_n)$, and the changes of coordinates

$$p'_j = \sum_i \frac{\partial q^i}{\partial q'^j} p_i \ , \quad dq'^j = \sum_i \frac{\partial q'^j}{\partial q^i} q^i \ .$$

It follows that, on $(T^*U) \cap (T^*U')$

$$\theta = \sum_i p_i dq^i = \sum_j p'_j dq'^j \ .$$

Therefore, the local expression is invariant under change of coordinates. Clearly, the same will hold for the symplectic form $\omega$. In particular, the standard atlas of $T^*Q$ is made up of Darboux charts with respect to the symplectic form $\omega$. For furthere details see Section 2.2 of [6].

We now consider two manifolds of the same dimension, $Q_1$ and $Q_2$, and a diffeomorphism between them $f : Q_1 \to Q_2$. The map $f$ naturally induces a map $f_\sharp : T^*Q_1 \to T^*Q_2$ which is called the *lift* of $f$. Let $p_1 = (x_1, \xi_1) \in T^*Q_1$, where $x_1 \in Q_1$ and $\xi_1 \in T^*_{x_1}Q$. Then $f_\sharp$ is defined by

$$f_\sharp(p_1) = p_2 = (x_2, \xi_2) \ ,$$

where $x_2 = f(x_1) \in Q_2$ and $\xi_1 = f^*(\xi_2)$. It can be shown that $f_\sharp$ is a diffeomorphism (See Section 2.4 of [6]).

**Proposition 3.5.4.** *Let $f : Q_1 \to Q_2$ be a diffeomorphism between manifolds and let $\theta_i$ denote the canonical 1-form on $T^*Q_i$, $i = 1, 2$. Then, the lift $f_\sharp : T^*Q_1 \to T^*Q_2$ pulls back the tautological 1-form on $T^*Q_2$ to the tautological 1-form on $T^*Q_1$, i.e.,*

$$(f_\sharp)^*\theta_2 = \theta_1 \ .$$

*Proof.* We first remark that, if $\pi_i : T^*Q_i \to Q_i$ denotes the natural projection for $i = 1, 2$, then we have

$$\pi_1 \circ f = f_\sharp \circ \pi_2 \ .$$

Now, if we let $f(p_1) = f(x_1, \xi_1) = (x_2, \xi_2) = p_2$, we have

$$\begin{aligned}(f_\sharp^* \theta_2)_{p_2} &= (df_\sharp)^*_{p_1}(\theta_2)_{p_2} = (df_\sharp)^*_{p_1}(d\pi_2)^*_{p_2}\xi_2 = (d(\pi_2 \circ f_\sharp))^*_{p_1}\xi_2 \\ &= (d(f \circ \pi_1))^*_{p_1}\xi_2 = (d\pi_2)^*_{p_1}(df)^*_{x_1}\xi_2 = (d\pi_2)^*_{p_1}\xi_1 = (\theta_1)_{p_1} \ ,\end{aligned}$$

where we denoted the differentials as $dF_p = F_{*,p}$. □

The results above can be summarized by saying that the canonical 1-form is the *natural* symplectic potential for the cotangent bundle.

**Corollary 3.5.5.** *Let $f : Q_1 \to Q_2$ be a diffeomorphism between manifolds and let $(T^*Q_i, \omega_i)$ be symplectic manifolds, where $\omega_i$ denotes the canonical symplectic structure. Then, the lift $f_\sharp : T^*Q_1 \to T^*Q_2$ is a symplectomorphism.*

*Proof.* By the previous proposition and the fact that the pullback and exterior derivative commute (Proposition 1.10.3/3), we have that

$$\omega_1 = d\theta_1 = d(f_\sharp^*\theta_2) = f_\sharp^*(d\theta_2) = f_\sharp^*\omega_2 \ .$$

□

## 3.6 The momentum map

In this section we define momentum mappings on symplectic $G$-manifolds and show how to induce one on symplectic manifold with a $G$-invariant symplectic potential. The main reference for this section is Section 10.1 of [3].

**Definition 3.6.1.** Let $(M, \omega)$ be a symplectic manifold, $G$ be a Lie group and $\mathfrak{g}$ be a Lie algebra

1. Let $\Psi$ be a Lie group action of $G$ on $M$. We say that $\Psi$ is *symplectic* if we have

   $$\Psi_g^*\omega = \omega$$

   for all $g \in G$. In this case, the tuple $(M, \omega, \Psi)$ is called a *symplectic $G$-manifold.*

2. Let $\psi$ be a Lie algebra action of $\mathfrak{g}$ on $M$. We say that $\psi$ is *symplectic* if the vector field $\psi(A)$ is symplectic for all $A \in \mathfrak{g}$.

Let $(M, \omega, \Psi)$ be a symplectic $G$-manifold and let $\mathfrak{g}$ be the Lie algebra of $G$. Then, by Proposition 3.3.4, all Killing vector fields (Definition 2.3.8) are symplectic since for all $A \in \mathfrak{g}$ we have

$$\mathcal{L}_{A_*}\omega = \left.\frac{d}{dt}\right|_{t=0} \left((\Psi_{\exp(tA)})^*\omega\right) = \left.\frac{d}{dt}\right|_{t=0} \omega = 0 \ .$$

In particular, this shows that Lie algebra action induced by a symplectic Lie group action $\Psi$ by the assignment of Killing vector fields (Corollary 2.3.11) is a symplectic Lie algebra action.

**Definition 3.6.2.** Let $(M, \omega)$ be a symplectic manifold and let $H \in C^\infty(M)$. We say that the triple $(M, \omega, H)$ is an *Hamiltonian system* and call $H$ the *Hamiltonian function*, or simply the *Hamiltonian*.

**Definition 3.6.3.** A Hamiltonian system $(M, \omega, H)$ is called *symmetric*

(1) under a symplectic action $\Psi$ of a Lie group $G$ if $\Psi_g^* H = H$ for all $g \in G$. Such a function is called *$G$-invariant* and the set of all $G$-invariant functions is denoted $C^\infty(M)^G$.

(2) under a symplectic action $\Psi$ of a Lie algebra $\mathfrak{g}$ if $\psi(A)H = 0$ for all $A \in \mathfrak{g}$.

It is clear that, if a Hamiltonian system is symmetric under a Lie group action, it will be symmetric under the corresponding Lie algebra action. Thus, every Lie group symmetry of a Hamiltonian system induces a Lie algebra symmetry. The converse need not to be true.

As shown above, the Killing vector fields of a symplectic $G$-action are automatically symplectic, however they will not necessarily be Hamiltonian. In this special case, for every $A \in \mathfrak{g}$ there exists a smooth function $J_A \in C^\infty(M)$ such that its corresponding Hamiltonian vector field is the Killing vector field generated by $A$, i.e., $X_{J_A} = A_*$.

**Definition 3.6.4.** Let $(M, \omega)$ be a symplectic manifold and let $\Psi$ be the an action of a Lie group $G$ on $M$. A mapping $J : M \to \mathfrak{g}^*$ is called a *momentum mapping* if

$$X_{J_A} = A_* \tag{3.5}$$

for all $A \in \mathfrak{g}$, where the functions $J_A : M \to \mathbb{R}$ are given by

$$J_A(m) = \langle J(m), A \rangle \ .$$

The action $\Psi$ is called *Hamiltonian* if it is symplectic and if there exists a momentum mapping. In this case, the tuple $(M, \omega, \Psi)$ is called a *Hamiltonian $G$-manifold*.

Given a Lie group action there can be various different momentum mappings. If a mapping $J$ is fixed we may include it in the tuple, i.e., $(M, \omega, \Psi, J)$. We also remark that, by the non-degeneracy of the symplectic form, Equation (3.5) is equivalent to

$$A_* \lrcorner \omega = -dJ_A \ .$$

**Proposition 3.6.5.** *Let $(M, \omega)$ be a symplectic manifold and let $\Psi$ be an action of a Lie group $G$ on $M$. A momentum mapping for $\Psi$ exists if and only if all Killing vector fields are Hamiltonian.*

*Proof.* Suppose every Killing vector field is Hamiltonian. Then, for a given basis $\{e_i\}$ in $\mathfrak{g}$, there exists some smooth functions $f_i \in C^\infty(M)$ such that $X_{f_i} = (e_i)_*$. Now, we consider the dual basis $\{e^{*i}\}$ in $\mathfrak{g}^*$ and define the map

$$J : M \to \mathfrak{g}^* \ , \quad J(m) := \sum_i f_i(m) e^{*i} \ .$$

Then, for every $A \in \mathfrak{g}$, we get

$$J_A = \sum_i \langle e^{*i}, A \rangle f_i$$

from Definition 3.6.4. It follows that

$$X_{J_A} = \sum_i \langle e^{*i}, A \rangle X_{f_i} = \sum_i \langle e^{*i}, A \rangle (e_i)_* = \sum_i \big( \langle e^{*i}, A \rangle e_i \big) = A_* \ .$$

Therefore, $J$ is a momentum mapping for $\Psi$. The converse is a direct consequence of the definition of momentum mapping. $\square$

**Remark 3.6.6.** The proof above shows that, given a momentum mapping $J : M \to \mathfrak{g}$, the functions $J_A$ can always be choses so that the mapping $A \mapsto J_A$ is linear.

We recall that, as shown in Example 2.3.5, any given representation of a Lie group $G$ on a vector space $V$ naturally induces an action of $G$ on $V$. In particular, the coadjoint representation $\mathrm{Ad}^*$ induces an action of $G$ on $\mathfrak{g}^*$ called the *coadjoint action of* $G$.

**Definition 3.6.7.** Let $(M, \omega, \Psi, J)$ be a Hamiltonian $G$-manifold. Then we say that $J$ is *equivariant* if it is equivariant with respect to the coadjoint action of $G$.

More explicitly, if $(M, \omega, \Psi, J)$ is a Hamiltonian $G$-manifold, the momentum mapping $J$ is equivariant if

$$J \circ \Psi_g = \mathrm{Ad}^*(g) \circ J$$

for all $g \in G$.

A particularly interesting case is that of a symplectic $G$-manifolds endowed with a symplectic potential, $\theta$, *invariant* under the Lie group action, i.e., $\Psi_g^* \theta = \theta$ for all $g \in G$. The following proposition shows how an equivariant momentum mapping can be constructed for this particular class of manifolds, making them into Hamiltonian $G$-manifolds.

**Proposition 3.6.8.** *Let $(M, \omega, \Psi)$ be a symplectic $G$-manifold. If $\omega = d\theta$ for some 1-form invariant under $\Psi$, then*

$$J : M \to \mathfrak{g}^* \ , \quad \langle J(m), A \rangle := \theta_m(A_*)$$

*is an equivariant momentum mapping.*

*Proof.* For every $m \in M$, $J(m)$ is clearly a linear functional on $\mathfrak{g}$. The invariance of $\theta$ implies $\mathcal{L}_{A_*}\theta = 0$, therefore

$$0 = \mathcal{L}_{A_*}\theta = A_* \lrcorner d\theta + d(A_* \lrcorner \theta) \ .$$

It follows that $A_* \lrcorner \omega = -dJ_A$ for all $A \in \mathfrak{g}$, making $J$ into a momentum map.
To show that $J$ is equivariant we note that, by Proposition 2.3.9/1, we have

$$\theta_m(A_*) = (\Psi_g^*\theta)_m(A_*) = \theta_{\Psi_g(m)}(\Psi_{g,*}A_*) = \theta_{\Psi_g(m)}((\mathrm{Ad}(g)A)_*) \ .$$

Since this holds for all $m \in M$ and for all $A \in \mathfrak{g}$, by definition of the coadjoint representation it follows that

$$J = \mathrm{Ad}^*(g^{-1}) \circ J \circ \Psi_g \tag{3.6}$$

for all $g \in G$. This shows that $J$ is equivariant. □

An important class of Hamiltonian $G$-manifolds which admit an invariant symplectic potential is constituted by the the cotangent bundles of $G$-manifolds. Let $(Q, \psi)$ be a left $G$-manifold and let $\pi : T^*Q \to Q$ denote the canonical projection. Since $\psi_g : Q \to Q$ is a diffeomorphism, it can be lifted to a diffeomorphism $\Psi_g : T^*Q \to T^*Q$ as showed in Section 2.3. Since by definition we have

$$\langle \Psi_g(\xi), X \rangle = \langle \xi, (\psi_{g^{-1}})_* X \rangle$$

with $\xi \in T^*Q$ and $X \in T_{(\psi_g \circ \pi)(\xi)}Q$, then the assignment $g \mapsto \Psi_g$ defines a map

$$\Psi : G \times T^*Q \to T^*Q$$

which is a left $G$-action. In particular, we note that by construction

$$\pi \circ \Psi_g = \psi_g \circ \pi$$

for all $g \in G$, therefore the projection is equivariant. An action $\Psi$ on $T^*Q$ such that the natural projection is equivariant is called a lift of the action $\psi$.
By Proposition 3.5.4 and Corollary 3.5.5 both the canonical 1-form on $T^*Q$, $\theta$, and the canonical symplectic structure, $\omega = d\theta$, are invariant under the lift $\Psi$. Therefore, the tuple $(M, \omega, \Psi)$ is a symplectic $G$-manifold satisfying the assumptions of Proposition 3.6.8. This yields the following result

**Corollary 3.6.9.** *In the case of the symplectic $G$-manifold $(T^*Q, \omega, \Psi)$ associated with the $G$-manifold $(Q, \psi)$, the equivariant momentum mapping defined in Proposition 3.6.8 is given by*

$$J : T^*Q \to \mathfrak{g} , \quad \langle J(\xi), A\rangle = \langle \xi, A_*^{\psi}(\pi(\xi))\rangle ,$$

*where $A_*^{\psi}$ is the Killing vector field generated by $A$ under the action of $\psi$ on $Q$.*

*Proof.* For $A \in \mathfrak{g}$ let $A_*^{\psi}$ be the Killing vector field generated by $A$ under $\Psi$. Since the projection $\pi : T^*Q \to Q$ is equivariant, by Equation (3.3) and Proposition 2.3.9 we have

$$\langle J(\xi), A\rangle = \theta_\xi(A_*^{\Psi}) = \langle \xi, \pi_*(A_*^{\Psi}(\xi))\rangle = \langle \xi, A_*^{\psi}(\pi(\xi))\rangle$$

□

## 3.7 Regular symplectic reduction

In this section we give a proof for the regular symplectic reduction Theorem and discuss the reduction of Hamiltonian systems and the Poisson structure induced on the reduced manifold. To do so, we briefly discuss some algebric properties of tangent spaces of the orbits of Lie group actions. The main references for this section are Sections 10.2 and 10.3 [3].

We recall that an immersed submanifold consists of a pair $(N, \varphi)$ where $N$ is a smooth manifold and $\varphi$ is an injective immersion (Definition 1.5.1). When $\varphi$ is open onto its image we talk about embedded submanifold (Definition 1.5.2).

**Definition 3.7.1.** Let $(M, G, \Psi)$ be a Lie group action, $(N, \varphi)$ an immersed submanifold of $M$ and let $(H, \varrho)$ be an immersed submanifold of $G$ where $H$ is also a Lie group. We say $(N, \varphi)$ is *invariant* under $(H, \varrho)$ if, for all $g \in \varrho(H)$, we have

$$\Psi_g(\varphi(N)) \subset \varphi(N) .$$

In the following we require that a submanifold be $H$-invariant where $H \subset G$ is taken only to be a closed subgroup. This is sufficient since the differentiable structure induced on $H$ by Theorem 2.1.2 actually makes it into an embedded submanifold $(H, i)$ of the ambient Lie group $G$, where the embedding is given by the inclusion map $i : H \hookrightarrow G$ (see Theorem 5.6.8 in [3]).

**Lemma 3.7.2.** *Let $(M, G, \Psi)$ be a proper Lie group action (Definition 2.3.6/3), $H \subset G$ a closed subgroup and $(N, \varphi)$ a $H$-invariant embedded submanifold. The, restriction of $\Psi$ to $H \times N$ is a proper Lie group action.*

*Proof.* See Proposition 6.3.4 in [3] □

Consider an Hamiltonian $G$-manifold $(M, \omega, \Psi, J)$ where $\Psi$ is a proper action and $J$ is an equivariant momentum mapping. Let $m \in M$, $\mu = J(m)$ and let $G_\mu$ be the stabilizer of $\mu$ under the coadjoint action. We recall that $G \cdot m$ and $G_\mu \cdot m$ denote the orbits of $G$ and $G_\mu$ through $m$ respectively and that the $\omega_m$-orthogonal complement of a subspace $V \subset T_m M$ is denoted $V^{\omega_m}$.

**Lemma 3.7.3.** *Let $(M, \omega, \Psi, J)$ be an Hamiltonian $G$-manifold where $\Psi$ is a proper action and $J$ is an equivariant momentum mapping. Let $G_\mu$ be the stabilizer of $\mu$ under the coadjoint action. Furthermore, let $m \in M$ such that $\mu = J(m)$ is a regular value of $J$ and $M_\mu = J^{-1}(\mu)$ is an embedded submanifold of $M$ by Theorem 1.5.4. Then we have*

$$T_m(G \cdot m) \cap T_m M_\mu = T_m(G_\mu \cdot m) \tag{3.7}$$

*and*

$$T_m(G \cdot m) = (T_m M_\mu)^{\omega_m} \ , \tag{3.8}$$

*Proof.* Remark 10.2.6 in [3]. □

The expressions above are justified by the fact that $G \cdot m$ and $G_\mu \cdot m$ are submanifolds of $M$ (See Corollary 6.2.9 of [3]). In particular, $G_\mu \cdot m$ can be seen as the orbit of $m$ via the restriction of the action $\Psi$ to $G_\mu \times M$ via Lemma 3.7.2. This follows since $G_\mu$, as the stabilizer of $\mu$, is equal to the preimage of $\mu$ under the orbit mapping $\mathrm{Ad}^*_\mu : G \to \mathfrak{g}^*$ of $\mu$ and is therefore closed in $G$.
Furthermore, if we again consider the action $\Psi$, $M_\mu$ is invariant under $G_\mu$ since

$$J \circ \Psi_g(m) = \mathrm{Ad}^*(g)J(m) = \mathrm{Ad}^*(g)\mu = \mu \ , \quad \forall m \in M_\mu \ , \forall g \in G_\mu \ ,$$

where we used the assumption that $J$ is equivariant. Since $\Psi$ was taken to be proper, by Lemma 3.7.2, it restricts to a proper $G_\mu$-action on $M_\mu$

$$\Psi^\mu : G_\mu \times M_\mu \to M_\mu \ . \tag{3.9}$$

When this induced action is free, we get the regular reduction theorem. This states that we can leverage the $G_\mu$-invariance of $M_\mu$ to quotient out this symmetry and induce a symplectic structure on this quotient space in a way that's compatible with the original symplectic structure.

**Theorem 3.7.4.** *Let $(M, \omega, \Psi, J)$ be an Hamiltonian $G$-manifold with a proper action and an equivariant momentum mapping. Let $\mu \in \mathfrak{g}^*$ be a regular value of $J$ and assume that $\Psi^\mu$, the induced $G_\mu$ action, is free.*

(1) *The topological space $M_\mu/G_\mu$ carries a unique differentiable structure such that the natural projection $\pi_\mu : M_\mu \to M_\mu/G_\mu$ is a submersion.*

(2) *There exists a unique symplectic form $\omega^\mu$ on $M_\mu/G_\mu$ such that*

$$\pi_\mu^* \omega^\mu = i_\mu^* \omega \ , \tag{3.10}$$

*where $i_\mu : M_\mu \to M$ denotes the natural inclusion mapping.*

*Proof.* The first point follows directly from Proposition 2.3.7 since $\Psi^\mu$ is assumed to be free and it proper by construction. To prove point 2 we observe that, since $\pi_\mu$ is a surjective submersion, every tangent vector of $M_\mu/G_\mu$ can be written as $\pi_{\mu*}X$ for some $X \in TM_\mu$. Therefore, we define $\omega^\mu$ by

$$\omega^\mu_{\pi_\mu(m)}(\pi_\mu X, \pi_\mu Y) := \omega_m(X, Y) \tag{3.11}$$

for $m \in M_\mu$ and $X, Y \in T_m M_\mu$. To prove that $\omega^\mu$ is well defined we show that the right side does not depend on $m$, $X$ and $Y$ as long as they project down the same point and tangent vectors. We let $\tilde{m} \in M_\mu$ and $\tilde{X}, \tilde{Y} \in T_{\tilde{m}} M_\mu$ such that $\pi_\mu(\tilde{m}) = \pi_\mu(m)$, $\pi_{\mu*}(\tilde{X}) = \pi_{\mu*}(X)$ and $\pi_{\mu*}(\tilde{Y}) = \pi_{\mu*}(Y)$. It follows that $m$ and $\tilde{m}$ lie on the same orbit, thus there exists a $g \in G_\mu$ such that $m = \Psi^\mu_g(\tilde{m})$ and

$$(\Psi^\mu_g)_* \tilde{X} - X \in T_m(G_\mu \cdot m) \ , \quad (\Psi^\mu_g)_* \tilde{Y} - Y \in T_m(G_\mu \cdot m) \ .$$

Since $\omega$ is $G$-invariant we get

$$\begin{aligned}
\omega_{\tilde{m}}(\tilde{X}, \tilde{Y}) &= ((\Psi^\mu_{g^{-1}})^* \omega)_m ((\Psi^\mu_g)_* \tilde{X}, (\Psi^\mu_g)_* \tilde{Y}) \\
&= \omega_m((\Psi^\mu_g)_* \tilde{X}, (\Psi^\mu_g)_* \tilde{Y}) \\
&= \omega_m(X + (\Psi^\mu_g)_* \tilde{X} - X, Y + (\Psi^\mu_g)_* \tilde{Y} - Y) = \omega_m(Y, X) \ .
\end{aligned}$$

where in the last step we used Equation (3.8). By construction $\omega^\mu$ satifies Equation (3.10). This defining property is enough to make it both uniquely defined and smooth. This result comes from a general fact about surjective submersions where, if $\Psi : M \to N$ is a surjective submersion and $F : M \to P$ is a smooth map, then, if there exists a mapping $\hat{F} : N \to P$ such that $\hat{F} \circ \Psi = F$, $\hat{F}$ is both smooth and uniquely defined (see Remark 1.5.16/3 in [3]). Therefore, by the smoothness of $\omega \circ i_\mu$, Equation (3.10) implies the desired result. For the same reason, $d\omega = 0$ implies $d\omega^\mu = 0$. Lastly, we prove that $\omega^\mu$ is non-degenerate. Thus, let $m \in M_\mu$ and $X \in T_m M_\mu$ such that

$$\omega^\mu_{\pi_\mu(m)}(\pi_{\mu*}X, \pi_{\mu*}Y) = 0$$

for all $Y \in T_m M_\mu$. By definition of $\mu$, this implies that $\omega_m(X, Y) = 0$ for all $Y \in T_m M_\mu$, therefore $X \in (T_m M_\mu)^{\omega_m}$. By Equations (3.7) and (3.8)

$$(T_m M_\mu)^{\omega_m} \cap T_m M_\mu = T_m(G_\mu \cdot m) \ .$$

Thus, $\pi_{\mu*}X = 0$ holds, making $\omega^\mu$ non-degenerate. $\square$

This reduction procedure can be used for Hamiltonian system. This results in another Hamiltonian system on $M_\mu/G_\mu$ with an Hamiltonian related to the original by the projection.

**Proposition 3.7.5.** *Let $(M, \omega, \Psi, J)$ be an Hamiltonian $G$-manifold with a proper action and an equivariant momentum mapping and let $H \in C^\infty(M)^G$. Let $\mu \in \mathfrak{g}^*$ be a regular value of $J$ and assume that $\Psi^\mu$, the induced $G_\mu$ action, is free.*

(1) *$M_\mu$ is invariant under the flow of $X_H$ and $X_H$ restricts to a vector field $X_H^\mu$ on $M_\mu$ which is $i_\mu$-related to $X_H$, that is*
$$i_{\mu*} \circ X_H^\mu = X_H \circ i_\mu \ . \tag{3.12}$$

(2) *$H$ defines a unique smooth function $H_\mu$ on $M_\mu/G_\mu$ by*
$$H_\mu \circ \pi_\mu = H \circ i_\mu \ . \tag{3.13}$$
*The corresponding Hamiltonian vector field $X_{H_\mu}$, relative to the symplectic structure $\omega^\mu$, is $\pi_\mu$ related to $X_H^\mu$*
$$X_{H_\mu} \circ \pi_\mu = \pi_{\mu*} \circ X_H^\mu \ . \tag{3.14}$$

*Proof.* See Proposition 10.3.3 in [3]. □

The Hamiltonian system $(M_\mu/G_\mu, \omega^\mu, H_\mu,)$ is referred to as the *reduction* of the Hamiltonian system $(M, \omega, H)$.
Since the result above is independent of the choice of Hamiltonian function, we get the following.

**Corollary 3.7.6.** *The Poisson structures defined by $\omega$ and $\omega^\mu$ are compatible, in the sense that*
$$\{f_\mu, h_\mu\} \circ \pi_\mu = \{f, h\} \circ i_\mu \tag{3.15}$$
*for all functions $f, h \in C^\infty(M)^G$ and $f_\mu, h_\mu \in C^\infty(M_\mu/G_\mu)$ related by Equation (3.13), i.e., $f_\mu \circ \pi_\mu = f \circ i_\mu$ and similarly for $h$ and $h_\mu$.*

*Proof.* Using Theorem 3.7.4 and Proposition 3.7.5, we compute
$$\begin{aligned}
\{f_\mu, h_\mu\}(\pi_\mu(m)) &= \omega^\mu_{\pi_\mu(m)}(X_{f_\mu}, X_{h_\mu}) \\
&= (\pi_\mu^*\omega^\mu)_m(X_f^\mu, X_h^\mu) \qquad \text{(Equation (3.12))} \\
&= (i_\mu^*\omega)_m(X_f^\mu, X_h^\mu) \qquad \text{(Equation (3.10))} \\
&= \omega_{i_\mu(m)}(X_f, X_h) \qquad \text{(Equation (3.14))} \\
&= \{f, h\}(i_\mu(m)) \ .
\end{aligned}$$
□

Since the action $\Psi$ was assumed to be symplectic, we actually have that $\{f,h\} \in C^\infty(M)^G$ if both $f$ and $h$ are. This is shown using the relation $F^*(X \lrcorner \omega) = (F^*X) \lrcorner (F^*\omega)$, where $F$ is a diffeomorphism and $F^*X := (F^{-1})_*(X \circ F)$, from which we get

$$(\Psi_g^* X_f) \lrcorner \omega = (\Psi_g^* X_f) \lrcorner (\Psi_g^* \omega) = \Psi_g^*(X_f \lrcorner \omega) = \Psi_g^*(-df) = -df = X_f \lrcorner \omega \ .$$

By the non degeneracy of $\omega$ this implies $X_f = \Psi_g^* X_f$ for every Hamiltonian vector field. Thus,

$$\Psi_g^* \{f,h\} = \Psi_g^*(X_f \lrcorner (-X_h \lrcorner \omega)) = ((\Psi_g^* X_f) \lrcorner (-\Psi_g^* X_h \lrcorner (\Psi_g^* \omega))) = (X_f \lrcorner (-X_h \lrcorner \omega)) = \{f,h\} \ .$$

Therefore, since the Poisson parenthesis of two $G$-invariant smooth functions is also $G$-invariant, by virtue of the uniqueness of the function defined by relation Equation (3.13), the Corollary above shows that $\{f,h\}_\mu = \{f_\mu, h_\mu\}$, meaning that the Poisson parenthesis actually reduces to the parenthesis of the reduced functions.

# Chapter 4

# Hamiltonian mechanics

In this chapter we formulate Hamiltonian mechanics using the the language of symplectic geometry. We first show how symplectic structures emerge naturally from simple Hamiltonian systems such as the phase space of a point particle. This is generalized by having any phase space described by an approapriate symplectic manifold where the dynamics is governed by the flow an Hamiltonian vector field. The integral curves' equations of such vector fields locally reproduce Hamilton equations. Then, we characterize constants of motions via the Poisson bracket and discuss the Hamiltonian formulation of Noether's Theorem. Later, we show how Hamiltonian mechanics can be obtained starting from the Lagrangian formulation, based on a variational principle, via the Legendre transformation. Finally, we show how the Lie group actions of translations and rotations yield the linear and angular momentums as momentum mappings. We apply this to show how the two-body problem can be solved by means of Regular Symplectic Reduction. The main references for this chapter are [3], [7], [8] and [9].

## 4.1 Hamiltonian mechanics in $\mathbb{R}^n$

In this section we illustrate how Hamiltonian mechanics on $\mathbb{R}^n$ can naturally be expressed via symplectic geometry. In particular, we show how the Hamilton equations for the dynamics a point particle can be expressed as the integral curves of an Hamiltonian vector field by viewing its phase space as a symplectic manifold. The main references for this section is Section 9.1 of [3] and Chapter 8 of [9].

Consider a mechanical system with *configuration space* $\mathbb{R}^n$ and *phase space* $\mathbb{R}^{2n}$ with coordinates $(\mathbf{q}, \mathbf{p})$ such that $\mathbf{q}, \mathbf{p} \in \mathbb{R}^n$, where as customary we identify $T^*_{\mathbf{q}}\mathbb{R}^n$ with $\mathbb{R}^n$. The first $n$ phase coordinates represent the spatial coordinates in the configuration space, while the last $n$ phase coordinates represent the generalized momenta of the system.

From classical mechanics, we know that every point in phase space represents a *state* of the system, every function on phase space, also known as a *phase function*, represents an *observable* such that the result of a measurement of an observable $f : \mathbb{R}^{2n} \to \mathbb{R}$ in the state $(\mathbf{q}, \mathbf{p})$ is $f(\mathbf{q}, \mathbf{p})$. We remark that there is a 1-1 correspondence between states and phase points and observables and phase functions respectively. As such, these terminologies are interchangable.
In the Hamiltonian formulation of classical mechanics the dynamics of the system is governed by the *Hamiltonian function* $H$ and its corresponding *Hamilton equations*

$$\begin{cases} \dot{q}^i(t) = \dfrac{\partial H}{\partial p_i}(\mathbf{q}(t), \mathbf{p}(t)) \\ \dot{p}_i(t) = -\dfrac{\partial H}{\partial q^i}(\mathbf{q}(t), \mathbf{p}(t)) \end{cases} \qquad i = 1, \dots, n \ , \tag{4.1}$$

where $\mathbf{q}(t) = (q^1(t), \dots, q^n(t))$ and $\mathbf{p}(t) = (p_1(t, \dots, p_n(t)))$. Here, the evolution of a state $(\mathbf{q}(t_0), \mathbf{p}(t_0)) = (\mathbf{q}_0, \mathbf{p}_0)$ is described by the solution of Hamilton equations with initial condition, specified at time $t_0$, equal to $(\mathbf{q}_0, \mathbf{p}_0)$. This can be seen as a parametrized curve in phase space. In physical applications, such a solution is often assumed to be both unique and global.

**Remark 4.1.1.** As the index placement suggests, the momenta $p_i$ are to be interpreted as covectors components and $\mathbf{p}$ is to be interpreted as a covector which acts on velocity vectors. This interpretation will become more natural in light of the formulation of mechanics on manifolds.

Alternatively, if we denote $\mathbf{x} = (\mathbf{q}, \mathbf{p})$ and $\nabla_{\mathbf{x}} H(\mathbf{x}) = (\partial H/\partial q^1(\mathbf{x}), \dots, \partial H/\partial p_n(\mathbf{x}))$, then Hamilton equations can be written more compactly as

$$\dot{\mathbf{x}}(t) = J_n \nabla_{\mathbf{x}} H(\mathbf{x}(t)) \ , \tag{4.2}$$

where $J_n$ is the standard symplectic matrix and $x(t) \coloneqq (\mathbf{q}(t), \mathbf{p}(t))$. Thus, given an Hamiltonian function for $\mathbb{R}^{2n}$, the equations of motions are readily obtained via the standard symplectic structure on phase space.

This procedure can be expressed naturally in the language of symplectic geometry. Let $(\mathbb{R}^{2n}, \omega_0)$ be the canonical symplectic vector space in dimension $2n$ where $\omega_0(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T J_n \mathbf{v}$ and let $H$ be an Hamiltonian function. Here we have a global Darboux chart with coordinates $q^i$ and $p_i$. In Example 3.3.2 we computed the Hamiltonian vector field generated by $H$

$$X_H = \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} \ . \tag{4.3}$$

Thus, given a Hamiltonian function $H$, Hamilton equations coincide with the equations for the integral curves of the Hamiltonian vector field $X_H$ generated by $H$. We see that all the informations about the dynamics of a system is encoded in the Hamiltonian vector field $H$, which accounts for both the Hamiltonian function itself and the symplectic structure of the manifold.

## 4.2 Hamiltonian mechanics on a manifold

In this section we illustrate how Hamiltonian mechanics can naturally be generalized to manifolds using symplectic geometry. In particular, we describe arbitrary physical systems with an appropriate Hamiltonian system, i.e., a tuple formed by a smooth manifold, a symplectic structure and a smooth function which serves as a Hamiltonian. The main references for this section are Section 9.1 of [3] and Section 4.3 of [7].

The discussion of Section 4.1 can be generalized to have the phase space of a given system be described be any symplectic manifold. Given a Hamiltonian system $(M, \omega, H)$ we have that every point in $M$ represents a state, that every function on $M$ represents an observable and that the time evolution of a state $m \in M$ is given by the integral curve passing through $m$ of the Hamiltonian vector field $H$ where the equation for such integral curves are the Hamilton equations when expressed in Darboux coordinates.
The most prominent phase space model is taking the manifold $M$ as the cotangent bundle of a given manifold $Q$ endowed with the canonical symplectic structure. Here, the base manifold $Q$ is interpreted as the configuration space of the system and the momenta are represented by 1-forms on $Q$. This is coherent with the notion that the momenta be linear functions of the velocities since tangent vectors are naturally interpreted as velocities.

We now focus on the role that the symplectic form plays in determining the dynamics of a system. Given a function on a manifold, which we think of as a Hamiltonian function, the associated Hamiltonian vector field is determined only when we equip the manifold with a symplectic structure. Thus, the evolution of the system depends not only on the Hamiltonian but also on the symplectic structure, since the Hamiltonian vector field depends on both. We conclude that the symplectic structure encodes the kinematics of a given physical system.

**Example 4.2.1.** Take the canonical symplectic vector space as a symplectic manifold $(\mathbb{R}^6, \omega)$ with coordinates $q^i$ and $p_i$. This describes the phase space associated with the configuration space $\mathbb{R}^3$, under the identification $T^*\mathbb{R}^3 \cong \mathbb{R}^6$. Consider a particle of mass m subject to a conservative force with potential energy $U(\mathbf{q})$ such that the Hamiltonian is

$$H(\mathbf{q}, \mathbf{p}) = \frac{p^2}{2m} + U(\mathbf{q}) \; . \tag{4.4}$$

The Hamiltonian vector field is thus

$$X_H = \sum_{i=1}^{3} \frac{p_i}{m}\frac{\partial}{\partial q^i} - \frac{\partial U}{\partial q^i}\frac{\partial}{\partial p_i} \tag{4.5}$$

and equations of motions are readily found

$$\dot{\mathbf{q}}(t) = \frac{\mathbf{p}}{m}\ , \quad \dot{\mathbf{p}}(t) = -\nabla_{\mathbf{q}} U(\mathbf{q}(t))\ , \tag{4.6}$$

where we denote $\nabla_{\mathbf{q}} U(\mathbf{q}(t)) := (\partial U/\partial q^1, \ldots, \partial U/\partial q^3)(\mathbf{q}(t))$. Thus, the first equation is just the classical definition of linear momentum that can be substituted in the second to find Newton's second law

$$m\ddot{\mathbf{q}}(t) = -\nabla_{\mathbf{q}} U(\mathbf{q}(t))\ . \tag{4.7}$$

Now, assume the point particle has charge $e$ and that an external magnetic field $\mathbf{B}$ is applied to the same system. The laws of motion will be different since the the particle will also be subject to the Lorentz force, thus

$$m\ddot{\mathbf{q}}(t) = \frac{e}{m}\mathbf{p} \times \mathbf{B} - \nabla_{\mathbf{q}} U(\mathbf{q}(t))\ . \tag{4.8}$$

Since the magnetic field is not conservative, instead of modifying the Hamiltonian function to obtain the equations of motions, we modify the symplectic form and define

$$\omega_B = \omega + e(B^1 dr^2 \wedge dr^3 + B^2 dr^3 \wedge dr^1 + B^3 dr^1 \wedge dr^2)$$

It follows that the Hamiltonian vector field generated by $H$ via this new symplectic form is

$$X_H^B = X_H + \frac{e}{m}(p_2 B^3 - p_3 B^2)\frac{\partial}{\partial p_1} + \frac{e}{m}(p_3 B^1 - p_1 B^3)\frac{\partial}{\partial p_2} + \frac{e}{m}(p_1 B^2 - p_2 B^1)\frac{\partial}{\partial p_3}.$$

whose integral curves must satisfy

$$\dot{\mathbf{q}}(t) = \frac{\mathbf{p}}{m}\ , \quad \dot{\mathbf{p}}(t) = \frac{e}{m}\mathbf{p} \times \mathbf{B} - \nabla_{\mathbf{q}} U(\mathbf{q}(t))\ .$$

Thus, by encoding the presence of non-conservative fields in the symplectic form, to account for how they modify the kinematics, we can still get the equations of motions from the Hamiltonian framework where the Hamiltonian is interpreted as the total energy of the system.

We end this section with the following consequence of Theorem 3.3.7, which states that Hamilton equations are invariant under canonical transformation, i.e. transformations that preserve the symplectic structure.

**Corollary 4.2.2.** *Let $(M, \omega, H)$ be a Hamiltonian system and let $\Phi : M \to M$ be a canonical transformation, then*

$$\Phi_* X_H = X_{H\circ\Phi^{-1}} \,. \tag{4.9}$$

In particular, if the canonical transformation is given in Darboux coordinates by

$$(\mathbf{q}, \mathbf{p}) \mapsto (\mathbf{Q}(\mathbf{q}, \mathbf{p}), \mathbf{P}(\mathbf{q}, \mathbf{p})) \,,$$

then Equation (4.9) yields Hamilton equations in the variables $(\mathbf{Q}, \mathbf{P})$ with Hamiltonian $K(\mathbf{Q}, \mathbf{P}) = (H \circ \Phi^{-1})(\mathbf{Q}, \mathbf{P})$. We remark that not all transformation that leave Hamilton equations invariant are canonical.

**Example 4.2.3.** Let $(U, q^i, p_i)$ be a Darboux chart for a symplectic manifold $(M, \omega)$ and consider the transformation

$$(\mathbf{q}, \mathbf{p}, H(\mathbf{q}, \mathbf{p})) \mapsto (\mathbf{Q}(\mathbf{q}, \mathbf{p}) = \mathbf{q},\ \mathbf{P}(\mathbf{q}, \mathbf{p}) = a\mathbf{p},\ K(\mathbf{X}) = aH(\mathbf{x}(\mathbf{X}))) \,,$$

where $\mathbf{x} = (\mathbf{q}, \mathbf{p})$, $\mathbf{X} = (\mathbf{Q}, \mathbf{P})$ and $a \neq 0$ is a constant. This transformation leaves the Hamilton equations invariant but it is not symplectic. We introduce the notation $(\mathbf{Q}(t), \mathbf{P}(t)) \coloneqq \mathbf{X}(t) \coloneqq \mathbf{X}(\mathbf{x}(t))$ where $\mathbf{x}(t) = (\mathbf{q}(t), \mathbf{p}(t))$. It follows that

$$\dot{Q}^i(t) = \dot{q}^i(t) = \frac{\partial H}{\partial p_i}(\mathbf{x}(t)) = \frac{\partial K}{\partial P_i}(\mathbf{X}(t)) \,, \qquad \dot{P}_i(t) = a\dot{p}_i(t) = -a\frac{\partial H}{\partial q^i}(\mathbf{x}(t)) = -\frac{\partial K}{\partial Q^i}(\mathbf{X}(t)) \,.$$

Therefore, the transformation preserves the Hamilton equations. However, it is not canonical as it does not pull the symplectic form back to itself.

## 4.3 Constants of motion and Noether's Theorem

In this section, we shift our attention to the time evolution of observables. In particular, we define the notion of constant of motion, characterize them in terms of the Poisson bracket and prove the Hamiltonian version of Noether's theorem. The main reference for this section is Section 9.1 of [3].

**Definition 4.3.1.** Let $(M, \omega, H)$ be an Hamiltonian system. A function $f \in C^\infty(M)$ is called a *constant of motion* if

$$\frac{d}{dt} f(\gamma(t)) = 0$$

for every integral $t \to \gamma(t)$ of the Hamiltonian vector field $X_H$.

Equivalently, a smooth function $f$ is a constant of motion if its pullback $\Phi_t^* f$ under the flow $\Phi_t$ of $X_H$ is independent of $t$.

**Proposition 4.3.2.** *The Hamiltonian $H$ of a Hamiltonian system $(M, \omega, H)$ is a constant of motion*

*Proof.* Let $t \to \gamma(t)$ be an integral curve of $X_H$, then

$$\frac{d}{dt}H(\gamma(t)) = X_H(H)(\gamma(t)) = dH(X_H)(\gamma(t)) = -\omega(X_H, X_H)(\gamma(t)) = 0 \ .$$

□

This result is simply the law conservation of energy, which is expected to hold for autonomous system. The result above is a special case of the following characterization of the constant of motions in terms of the Poisson bracket.

**Proposition 4.3.3.** *Let $(M, \omega, H)$ be a Hamiltonian system,*

(1) *A smooth function $f$ is a constant of motion if and only if $\{H, f\} = 0$.*

(2) *If $f, g \in C^\infty(M)$ are constants of motion, then $\{f, g\}$ is a constant of motion as well.*

*Proof.* The first statement follows from

$$\frac{d}{dt}f(\gamma(t)) = X_H f(\gamma(t)) = \{H, f\}(\gamma(t)) \ .$$

The second point follows from the first and Jacobi identity, since

$$\{H, \{f, g\}\} = \{f, \{H, g\}\} + \{g, \{f, H\}\} = 0 \ .$$

□

Let $(M, \omega, H)$ be a Hamiltonian system and let $e \in \mathbb{R}$ be a regular value (Definition 1.5.3) of $H$. From Theorem 1.5.4 we get that $H^{-1}(e)$ is an embedded submanifold of $M$ of dimension $2n - 1$. The connected components of this submanifold are called *regular energy surfaces*. Let $\Sigma$ be one such regular energy surface and let $i : \Sigma \to M$ be the natural inclusion mapping. Since by Proposition 4.3.2 $\Sigma$ is invariant under the flow of $X_H$, then $X_H$ is tangent to $\Sigma$ in the sense of Definition 1.5.5. From Proposition 1.5.6, it follows that $X_H$ induces a unique vector field $X_H^\Sigma$, on $\Sigma$, which is $i$-related to $X_H$,

$$i_*(X_H^\Sigma) = X_H \circ i \ .$$

Thus, the conservation of energy can be exploited to reduce the dynamics to a submanifold of codimension 1 for every regular value of the Hamiltonian function.

We now prove the Hamiltonian formulation of Noether's theorem. Recall that, by Definition 3.6.4, given an Hamiltonian $G$-manifold $(M, \omega, \Psi)$, $J_A$ denotes the Hamiltonian function corresponding to the Killing vector field $A_*$ generated by $A \in \mathfrak{g}$.

**Theorem 4.3.4.** *Let $(M, \omega, \Psi, J)$ be a Hamiltonian $G$-manifold and let $H \in C^\infty(M)^G$. For $A \in \mathfrak{g}$, the function $J_A$ is a constant of motion for the Hamiltonian system $(M, \omega, H)$, i.e.,*

$$\frac{d}{dt} J_A(\gamma(t)) = 0$$

*for all integral curves $\gamma(t)$ of the Hamiltonian vector field $X_H$.*

*Proof.* For all $m \in M$ we have

$$\{J_A, H\}(m) = X_{J_A}(H)(m) = A_*(H)(m) = \frac{d}{dt} H \circ \Psi_{\exp(tA)}(m) = 0 \ ,$$

which follow from the $G$-invariance of $H$. Hence, the result follow from Proposition 4.3.3 □

With a slight abuse of language, the view of Hamiltonian mechanics presented in the previous section could be called the Schroedinger picture of classical mechanics, as the dynamics of the system is given by the evolution of its states. Alternatively, one can study the evolution of a system by studying how its observables evolve in time. Consider the Hamiltonian system $(M, \omega, H)$. Let $f \in C^\infty(M)$ and define a family of functions $f(t) = f \circ \Phi_t$ where $\Phi$ is the flow of $X_H$. We remark that, since $\Phi_{t=0} = \mathrm{id}_M$, we have that $(f(0))(x) = f(x)$ for all $x \in M$. Thus, the result of a measurement of $f$ in the state $\Phi_t(x)$ is equal to the result of a measurement of $f(t)$ in the state $x \in M$ as

$$(f(t))(x) = f(\Phi_t(x)) \ .$$

If we denote $\dot{f}(t) = (d/dt)(f(t))$, then the equation of motion for $f(t)$ can be written as

$$\dot{f}(t) = \{H, f(t)\} \ .$$

This follows from Equation (3.2) as

$$\dot{f}(t)(x) = \frac{d}{dt}\left(f \circ \Phi_t\right)(x) = (df(X_H))(\Phi_t(x)) = \{H, f\}(\Phi_t(x)).$$

Now by Remark 3.3.5 we know that the Hamiltonian vector field $X_H$ is also symplectic, therefore, by Proposition 3.3.4/2, its flow is a symplectomorphism for all $t \in I$ and some small enough interval $I$ containing 0. We can now apply Proposition 3.4.4 to get

$$(\{H, f\} \circ \Phi_t)(x) = \{H \circ \Phi_t, f \circ \Phi_t\}(x) = \{H, f \circ \Phi_t\}(x)$$

for all $x \in M$, where in the last step we used the fact that $H$ is a constant of motion. We remark that if the flow $\Phi$ is complete the interval $I$ is equal to $\mathbb{R}$ and the equation holds

for all times.
In particular, the equations of motion can be written as

$$\dot{q}^i = \{H, q^i\} \ , \qquad \dot{p}_i = \{H, p_i\} \ ,$$

where the coordinates are now viewed as observables. This can be seen as the Heinsenberg picture of classical mechanics.

## 4.4 Lagrangian mechanics

In this section, we present the Lagrangian formulation of classical mechanics in the generalized context of smooth manifolds. To do this we first define the Lagrangian as a smooth function on the tangent bundle and define its associated functional, the action. Later, we define the corresponding variational problem that yields the Euler-Lagrange equations. Finally, we discuss the main issues that arise in our local treatment of the problem. The main reference for this section is Lecture 4 in Part 1 of [8].

**Definition 4.4.1.** Let $Q$ be a smooth manifold. A *Lagrangian* of $Q$ is a smooth function of $L : TQ \to \mathbb{R}$. For any smooth curve $\gamma : [a, b] \to M$, we define the *functional* associated to $L$, also called the *action*, to be

$$\mathcal{F}_L(\gamma) = \int_a^b L(\gamma'(t))dt \ .$$

We remark that the word functional is classical and $\mathcal{F}_L$ is interpreted as a function on the set of all smooth curves $\gamma : [a, b] \to M$.

**Definition 4.4.2.** Let $Q$ be a manifold and let $\gamma : [a, b] \to Q$ be a smooth curve. A *smooth variation of $\gamma$ with fixed endpoints* is a smooth map

$$\Gamma : [a, b] \times (-\epsilon, \epsilon) \to Q \ ,$$

for some $\epsilon > 0$, such that $\Gamma(t, 0) = \gamma(t)$ for all $t \in [a, b]$ and that $\Gamma(a, s) = \gamma(a)$ and $\Gamma(b, s) = \gamma(b)$ for all $s \in (-\epsilon, \epsilon)$.

In the following, since we will consider only smooth variations with fixed endpoints, we will simply speak of variations.

**Definition 4.4.3.** Let $Q$ be a smooth manifold and let $\Gamma$ be a variation of a smooth curve $\gamma : [a, b] \to Q$. We define a function $\mathcal{F}_{L,\Gamma} : (-\epsilon, \epsilon) \to \mathbb{R}$ by setting

$$\mathcal{F}_{L,\Gamma}(s) = \mathcal{F}_L(\gamma_s)$$

where $\gamma_s(t) = \Gamma(t, s)$. We say that such a curve $\gamma$ is *$L$-critical* if ${\mathcal{F}_{L,\Gamma}}'(0) = 0$ for all variations of $\gamma$. Whenever this is the case, we say that the action $F_L$ is *stationary* at $\gamma$.

From classical mechanics, it is know that a solution of the equations of motion coincides with the curves such that the action is stationary. This means that the solution of the equations of motion of the system, whose position and velocities are represented by points in the tangent bundle $TQ$, are $L$-critical curves.
To find such curve, we choose a particularly convenient set of coordinates. Let $(U, x)$ be a chart for $Q$. This canonically extends to a chart $(TU, x, p)$ with the property that, for any curve $\gamma : [a, b] \to U$, with coordinates $y^i = x^i \circ \gamma$, the $p$-coordinates of its lift, $\gamma' : [a, b] \to TU$, are given by $p \circ \gamma' = \dot{y}$. We say that the coordinates $(x^i, p^i)$ on $TU$ are the *canonical coordinates* associated to the coordinate system $(U, x^i)$. As it is customary in physics, given a coordinate system $(U, q)$, we will denote the associated canonical coordinates on $TU$ with $(q, \dot{q})$, with an abuse of notation.
Consider a canonical coordinate system $(q, \dot{q})$ on $TU$ associated with the chart $(U, q)$ for $Q$. A Lagrangian $L$ can be expressed in local coordinates as $L(q, \dot{q})$. Thus, for a curve $\gamma : [a, b] \to M$ that lies in $U$, the associated functional $\mathcal{F}_L$ becomes

$$\mathcal{F}_L(\gamma) = \int_a^b L(y(t), \dot{y}(t))dt \ .$$

It can be shown (see Lecture 4 in Part 1 of [8]) that the curve $\gamma$ is $L$-critical if and only if $y(t) = (x \circ \gamma)(t)$ satifies the system of differential equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(y(t), \dot{y}(t))\right) - \frac{\partial L}{\partial q^i}(y(t), \dot{y}(t)) = 0 \tag{4.10}$$

for $i = 1, \ldots, n$. These are obviously the *Euler-Lagrange equations*.
As is, the main issue with this approach is that these equations give a necessary and sufficient condition for a curve to be $L$-critical only when it lies in a coordinate neighbourhood. Also, we would like to express the equations in a way that does not depend on the choice of coordinates. In the following, we want to remedy this issues.

## 4.5 The Legendre transform

In this section we obtain a coordinate-free form of the Euler-Lagrange equations on manifolds and obtain the Hamiltonian formulation in our generalized context starting the Lagrangian formulation. To do this we induce a symplectic structure on the tangent bundle and define a Hamiltonian system on it such that the equations of motion are equivalent to the Euler-Lagrange equations in an appropriate sense. Finally, we show that this system is equivalent to the Hamiltonian formulation on the cotangent bundle for a suitably defined Hamiltonian function. The main reference for this section is Section 9.1 of [3].

As stated previously, the most prominent phase space model is that of a cotangent bundle, $M = T^*Q$, where the base manifold $Q$ is the configuration space of the system. Given the importance of this model, we want to derive the Hamiltonian formalism starting from the Lagrangian formulation of mechanics.

Let $\pi_T : TQ \to Q$ denote the canonical bundle projection and $L : TQ \to \mathbb{R}$ a Lagrangian. We extend $L$ to a mapping

$$\tilde{L} : TQ \to Q \times \mathbb{R} \,, \qquad \tilde{L}(X) = (\pi_T(X), L(X)) \,.$$

Here, we treat $Q \times \mathbb{R}$ as a vector bundle over $Q$ with canonical projection $\rho : Q \times \mathbb{R} \to Q$. We define the *fibre derivative* of $L$ to be the mapping

$$FL : TQ \to \mathrm{Hom}(TQ, Q \times \mathbb{R}) \cong T^*Q \,, \qquad (FL)(X) := (\tilde{L}_{\pi_T(X)})_{*,X} \,, \tag{4.11}$$

where $\tilde{L}_q := \tilde{L}|_{\pi_T^{-1}(q)}$ is the restriction of $\tilde{L}$ to the tangent space over $q \in Q$. For further details on vector bundles and the construction $\mathrm{Hom}(TQ, Q \times \mathbb{R})$ see sections 2.2 and 2.4 in [3]. It is clear that $FL$ is both fiber-preserving and smooth. Also, if we identity $(FL)(X)$ with a cotangent vector we have

$$\langle (FL)(X), Y \rangle = \left.\frac{d}{dt}\right|_0 L(X + tY)$$

for $X, Y$ in the same tangent space. If the fibre derivative $FL : TQ \to T^*Q$ is a diffeomorphism, it is called the *Legendre transformation induced by* $L$ and the Lagrangian is called *hyperregular*. We now compute the local form for the fibre derivative. Let $X, Y \in T_qQ$ for some $q \in Q$ such that, in coordinates $(q, \dot{q})$, they're given by $X = (q, x^1, \ldots, x^n)$ and $Y = (q, y^1, \ldots, y^n)$, then

$$\langle (FL)(X), Y \rangle = \left.\frac{d}{dt}\right|_0 L(X + tY) = \sum_i^n \frac{\partial L}{\partial \dot{q}^i}(X) y^i = \left( \sum_i^n \frac{\partial L}{\partial \dot{q}^i}(X) dq^i \right) Y \,.$$

Since $Y$ was arbitrary, it follows that

$$FL(q, \dot{q}) = \sum_i \frac{\partial L}{\partial \dot{q}^i}(q, \dot{q}) dq^i \tag{4.12}$$

and the fibre derivative is given by the mapping

$$(q, \dot{q}) \mapsto (q, p = \frac{\partial L}{\partial \dot{q}^i}(q, \dot{q})) \,,$$

where $(q, p)$ are the local coordinates induces on $T^*Q$ by the coordinates $q$. When $L$ is hyperregular, this is the local representative of the Legendre transform known from classical mechanics.

**Remark 4.5.1.** A Lagrangian $L$ is called regular if the fibre derivative $FL$ is a local diffeomorphism. This holds if and only if

$$\det\left(\frac{\partial^2 L}{\partial \dot q^i \partial \dot q^j}\right) \neq 0 \ ,$$

where $(q, \dot q)$ are local coordinates on $TQ$ (see Remark 9.1.1 in [3]). In particular, if $L$ is hyperregular the condition above has to hold globally.

Now, assume that $L$ is hyperregular and let $\omega$ be the canonical symplectic form on $T^*Q$. We endow $TQ$ with a symplectic structure given by $\omega_L := (FL)^*\omega$ such that $FL$ becomes a symplectomorphism. Next we define the *energy function*

$$E : TQ \to \mathbb{R} \ , \qquad E(X) := \langle (FL)(X), X\rangle - L(X) \tag{4.13}$$

and consider the Hamiltonian vector field $X_E$ generated by $E$ on $TQ$

$$X_E \lrcorner \omega_L = -dE \ .$$

In local coordinates, the equations for the integral curves of $X_E$ turn out to be equivalent to the Euler-Lagrange equations, which casts them in a coordinate-free form.

**Proposition 4.5.2.** *Consider an hyperregular Lagrangian $L$. Then, a smooth curve $\gamma : I \to TQ$ is an integral curve of the Hamiltonian vector field $X_E$, generated by the energy function defined in Equation* (4.13)*, if and only if there is a curve $c : I \to Q$ that satifies the Euler-Lagrange equations and whose tangent lift is $\gamma$.*

*Proof.* In local coordinates $(q, \dot q)$, we have by Equation (4.12)

$$E(q, \dot q) = \sum_i \dot q^i \frac{\partial L}{\partial \dot q^i}(q, \dot q) - L(q, \dot q) \ ,$$

from which it follows that

$$dE = \sum_{i,k} \dot q^i \left(\frac{\partial^2 L}{\partial q^k \partial \dot q^i} dq^k + \frac{\partial^2 L}{\partial \dot q^k \partial \dot q^i} d\dot q^k\right) - \sum_i \frac{\partial L}{\partial q^i} dq^i \ .$$

Meanwhile, from the definition of $\omega_L$ and Equation (4.12) we get

$$\omega_L = \sum_i d(FL^* p_i) \wedge d(FL^* q^i) = \sum_i d\left(\frac{\partial L}{\partial \dot q^i}\right) \wedge dq^i \ ,$$

or more explicitly

$$\omega_L = \sum_{i,k} \frac{\partial^2 L}{\partial q^k \partial \dot q^i} dq^k \wedge dq^i + \frac{\partial^2 L}{\partial \dot q^k \partial \dot q^i} d\dot q^k \wedge dq^i \ .$$

Let $X_E$ be given by

$$X_E = \sum_i A^i \frac{\partial}{\partial q^i} + B^i \frac{\partial}{\partial \dot{q}^i} ,$$

where $A^i$ and $B^i$ are functions on $TQ$, then

$$X_E \lrcorner \omega_L = \sum_{i,k} A^k \frac{\partial^2 L}{\partial q^k \partial \dot{q}^i} dq^i - A^i \frac{\partial^2 L}{\partial q^k \partial \dot{q}^i} dq^k - A^i \frac{\partial^2 L}{\partial \dot{q}^k \partial \dot{q}^i} d\dot{q}^k + B^k \frac{\partial^2 L}{\partial \dot{q}^k \partial \dot{q}^i} dq^i .$$

Now, from the equation $X_E \lrcorner \omega_L = -dE$, we can determine the functions $A^i$ and $B^i$. By comparing the two sides, is clear that

$$A^i = \dot{q}^i$$

from which the equation simplifies to

$$\sum_k \frac{\partial^2 L}{\partial \dot{q}^k \partial \dot{q}^i} B^k = \frac{\partial L}{\partial q^i} - \sum_k \frac{\partial^2 L}{\partial q^k \partial q^i} \dot{q}^k$$

for all $i$. From Remark 4.5.1, it follows that the functions $B^i$ are determined as the matrix $\left[\partial^2 L / \partial \dot{q}^k \partial \dot{q}^i\right]$ is always invertible. Now, let $\gamma(t)$ be an integral curve for $X_E$, i.e., $\gamma'(t) = X_{E,\gamma(t)}$. Then, if we express $\gamma$ in local coordinates $\gamma(t) = (x(t), v(t))$, we have the differential equations

$$\dot{x}^i(t) = A^i(\gamma(t)) , \qquad \dot{v}(t) = B^i(\gamma(t)) .$$

The first set of equations then states that $\dot{x}^i(t) = v^i(t)$, while the second set can be written as

$$\frac{\partial L}{\partial q^i}(\gamma(t)) - \sum_k \frac{\partial^2 L}{\partial q^k \partial \dot{q}^i}(\gamma(t)) \dot{x}^k - \frac{\partial^2 L}{\partial \dot{q}^k \partial \dot{q}^i}(\gamma(t)) \dot{v}^k = 0 .$$

By using that $\dot{x}^i(t) = v^i(t)$, this is equivalent to

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(x(t), \dot{x}(t))\right) - \frac{\partial L}{\partial q^i}(x(t), \dot{x}(t)) = 0 .$$

Thus, if the curve $\gamma(t)$ is an integral curve of $X_E$ it is precisely the tangent lift of the curve $c(t) = (\pi_T \circ \gamma)(t)$ which solves the Euler-Lagrange equations, where $\pi_T : TQ \to Q$ is the natural projection.

Conversely, let $c(t)$ be a curve on $Q$ that satifies the Euler-Lagrange equations and consider its lift $c'(t)$. We want to show that $c'(t)$ is an integral curve of $X_E$. If we expand it in canonical coordinates we have $c'(t) = (y(t), \dot{y}(t))$, meaning that it satifies first set of equations, $\dot{y}^i(t) = A^i(c'(t))$, since $\dot{q}^i(c'(t)) = \dot{y}^i(t)$. The second condition is shown to hold by expanding the Euler-Lagrange equation and reversing the steps above. □

Finally, we define the Hamiltonian of the system by

$$H : T^*Q \to \mathbb{R} \;, \qquad H \coloneqq E \circ (FL)^{-1} \;. \tag{4.14}$$

It follows that

$$((FL)^{-1})^*(X_E \lrcorner \omega_L) = -dH \;.$$

Meanwhile, since $FL$ is a symplectomorphism, Theorem 3.3.7 implies that

$$(FL)_* X_E = X_H \;.$$

Thus, when the Lagrangian is hyperregular the Lagrangian and Hamiltonian formulations are equivalent. In particular, $(TQ, \omega_L, E)$ is a Hamiltonian system equivalent to $(T^*Q, \omega, H)$.
We end this section by writing the Hamiltonian function is local coordinates. Let $q$ be coordinates on $Q$ that induce the coordinates $(q, \dot{q})$ and $(q, p)$ on $TQ$ and $T^*Q$ respectively. Then

$$H(q,p) = \sum_i \frac{\partial L}{\partial \dot{q}^i}(q, \dot{q}(q,p))\dot{q}^i(q,p) - L(q, \dot{q}(q,p)) \;,$$

where $\dot{q}(q,p)$ is determined by solving the system $p_i = \frac{\partial L}{\partial \dot{q}^i}$ in terms of $\dot{q}^i$, where we use the fact that $FL$ is a diffeomorphism. Thus, when $Q = \mathbb{R}^n$, the Hamiltonian function is the Legendre transform of $L$ known from classical mechanics.

## 4.6 Examples of momentum maps

In this section we show some important examples of momentum mappings. In particular, we show how the left Lie group actions of translation and rotations on Euclidean space respectively yield the momentum and angular momentum as momentum mappings. The main reference for this section is Section 10.1 of [3].

Recall that, by Corollary 3.6.9, given a $G$-manifold $(Q, \psi)$, there is a standard procedure to endow the symplectic $G$-manifold $(T^*Q, \omega, \Psi)$ with a momentum mapping. In the following examples, we show how the momentum mapping coincides with the momentum and angular momentum known from classical mechanics for an appropriate choice of Lie group actions. This fact explains the origin of the name momentum mapping.

**Example 4.6.1.** Consider the base manifold $Q = \mathbb{R}^3$ and let $G$ be Lie group $\mathbb{R}^3$, with the group operation given by the sum of two vectors, act on $Q$ by translations. Thus, we define the left action $\psi : G \times Q \to Q$

$$\psi_{\mathbf{a}}(\mathbf{x}) \coloneqq \mathbf{x} + \mathbf{a} \;.$$

Under the identification $T^*Q \cong \mathbb{R}^3 \times \mathbb{R}^3$, the induced action on $T^*Q$ is given by

$$\Psi_{\mathbf{a}}(\mathbf{x}, \mathbf{p}) = (\mathbf{x} + \mathbf{a}, \mathbf{p}) \ .$$

We now want to find an explicit form for the Killing vector fields generated by the elements of $\mathfrak{g} \in \mathbb{R}^3$. To do this we use the result of Remark 2.3.4/1 in [3], which states that for any open subset $M$ of a vector space $V$ the vector fiels on $M$ can be represented as maps from $M$ to $V$. In particular, $L(G)$, the left-invariant vector fields on $G$, is formed by maps of the form $X : G \to \mathbb{R}^3$. Let $X \in L(G)$ and consider its flow $\Phi_t^X$. By Proposition 1.7.5/5, the left-invariance of $X$ implies that that

$$\Phi_t^X(\mathbf{a}) = \Phi_t^X \circ l_{\mathbf{a}}(\mathbf{0}) = l_{\mathbf{a}} \circ \Phi_t^X(\mathbf{0}) = \mathbf{a} + \Phi_t^X(\mathbf{0})$$

for all $\mathbf{a} \in \mathbb{R}^3$. Therefore,

$$X(\mathbf{a}) = \left.\frac{d}{dt}\right|_0 \left(\Phi_t^X(\mathbf{a})\right) = \left.\frac{d}{dt}\right|_0 \left(\mathbf{a} + \Phi_t^X(\mathbf{0})\right) = X(\mathbf{0}) \ ,$$

meaning that the left-invariant vector fields corresponds to constant mappings. This identifies $L(G)$ with $\mathbb{R}^3$. Under this identification, the equation for the integral curve of the left-invariant vector field corresponding to $\mathbf{X} \in \mathbb{R}^3$ is given by $\dot{\mathbf{c}}(t) = \mathbf{X}$. Thus, the solution with initial condition $\mathbf{v} \in \mathbb{R}^3$ is $\mathbf{c}(t) = \mathbf{v} + t\mathbf{X}$. Therefore, the expontial mapping is the identity since we have

$$\exp(t\mathbf{X}) = \Phi_t^{\mathbf{X}}(\mathbf{0}) = t\mathbf{X} \ .$$

It follows that the Killing vector field generated by $\mathbf{b} \in \mathfrak{g} \cong \mathbb{R}^3$ is

$$\mathbf{b}_*^{\psi}(\mathbf{x}) = \left.\frac{d}{dt}\right|_0 \psi_{\exp(t\mathbf{b})}(\mathbf{x}) = \left.\frac{d}{dt}\right|_0 (\mathbf{x} + t\mathbf{b}) = \mathbf{b} \ .$$

Finally, we can apply Corollary 3.6.9 to obtain

$$\langle J(\mathbf{x}, \mathbf{p}), \mathbf{b}\rangle = \langle \mathbf{p}, b_*^{\psi}(\mathbf{x})\rangle = \mathbf{p} \cdot \mathbf{b} \ ,$$

which yields

$$J(\mathbf{x}, \mathbf{p}) = \mathbf{p} \ . \tag{4.15}$$

Thus, the momentum mapping associated with translations is precisely the momentum from classical mechanics.

**Example 4.6.2.** We considered the action of spatial translations on the phase space of a point particle. This can be generalized to an arbitrary number of point particles, whose configuration and phase space are given by an appropriate number of copies of $Q$ and $T^*Q$ respectively. Here, the action is given by the simultaneous translation

$$\psi_{\mathbf{a}}(x) = \Psi_{\mathbf{a}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) = (\mathbf{x}_1 + \mathbf{a}, \ldots, \mathbf{x}_N + \mathbf{a}) \ .$$

Thus, the Killing vector field generated by $\mathbf{b} \in \mathfrak{g} \cong \mathbb{R}^3$ is

$$\mathbf{b}_*^\psi(x) = \frac{d}{dt}\bigg|_0 \psi_{\exp(t\mathbf{b})}(x) = \frac{d}{dt}\bigg|_0 (\mathbf{x}_1 + t\mathbf{b}, \dots, \mathbf{x}_N + t\mathbf{b}) = (\mathbf{b}, \dots, \mathbf{b}) .$$

By Corollary 3.6.9 we have

$$\langle J(x,p), \mathbf{b} \rangle = \langle (\mathbf{p}_1, \dots, \mathbf{p}_N), b_*^\psi(x) \rangle = \mathbf{p}_1 \cdot \mathbf{b} + \cdots + \mathbf{p}_N \cdot \mathbf{b} ,$$

from which we conclude that the momentum map is exactly the total momentum of the system, i.e.,

$$J(x,p) = \sum_i^N \mathbf{p}_i .$$

By Noether's Theorem (Theorem 4.3.4), it is clear that when the Hamiltonian of the system is invariant under translations the total momentum of the system is conserved.

**Example 4.6.3.** Let $Q = \mathbb{R}^3$ and let $G = \mathrm{SO}(3)$ act by rotations

$$\psi_g(\mathbf{x}) = g\mathbf{x} .$$

Thus, under the identification $T^*Q \cong \mathbb{R}^3$ it induces the action $\Psi : G \times T^*Q \to T^*Q$ is given by

$$\Psi_g(\mathbf{x}, \mathbf{p}) = (g\mathbf{x}, g\mathbf{p}) .$$

We observe that the action $\psi$ is associated with the identical representation $\varphi : \mathrm{SO}(3) \to \mathrm{GL}(\mathbb{R}^3)$, $\varphi(g) = g$, in the sense of Example 2.3.5. It follows that the Killing vector field generated by $A \in \mathfrak{so}(3)$ is

$$A_*^\psi(\mathbf{x}) = \frac{d}{dt}\left(\psi_{\exp(tA)}(\mathbf{x})\right) \frac{d}{dt}\left(\exp(tA)(\mathbf{x})\right) = A\mathbf{x} . \tag{4.16}$$

Thus, the momentum mapping satifies

$$\langle J(\mathbf{x}, \mathbf{p}), A \rangle = \langle \mathbf{p}, A_*^\psi(\mathbf{x}) \rangle = \mathbf{p} \cdot A\mathbf{x} .$$

Recall now the Lie algebra isomorphism defined by Equation (2.4) in Example 2.1.6. Denote $\mathbf{A} \in \mathbb{R}^3$ the vector corresponding to $A$ under this isomorphism and observe that

$$\mathbf{p} \cdot A\mathbf{x} = \mathbf{p} \cdot (\mathbf{A} \times \mathbf{x}) = (\mathbf{x} \times \mathbf{p}) \cdot \mathbf{A} .$$

This finally shows that the momentum mapping associated with rotations is

$$J(\mathbf{x}, \mathbf{p}) = \mathbf{x} \times \mathbf{p} , \tag{4.17}$$

which is precisely the angular momentum. By Noether's Theorem (Theorem 4.3.4), it is clear that when the Hamiltonian of the system is invariant under rotations the angular momentum of the system is conserved.

## 4.7 Kepler's problem

In this section we apply Regular Symplectic Reduction to the two-body problem. The problem can be simplified by means of two symplectic reductions. The first coincides with the choice of the center-of-mass as a reference frame while the second leverages the rotational invariance of the system to yield a one dimensional problem. The main references for this section are Section 10.6 of [3] and Section 8.1 in [7].

Consider the standard two-body problem. This is the problem of two point particles of masses $m_1$ and $m_2$ which interact gravitationally. From classical mechanics, we know that this highly symmetric problem can be reduced to a one dimensional problem. Traditionally, this is done by choosing the center-of-mass reference frame to reduce the problem to that of a 3-dimensional point particle in a gravitational potential. Then, by exploiting the rotational invariance of the system, this can be further reduced to a one-dimensional problem with effective potential given by a gravitational and a centrifugal term. We will show that these two steps can be formalized in the Hamiltonian setting by means of two symplectic reductions.

The system is described by the Hamiltonian system $(M_0, \omega_0, H_0)$, where $M = T^*(\mathbb{R}^6 \backslash \{0\})$, $\omega_0$ is the canonical symplectic structure on $M$ and the Hamiltonian function is

$$H_0(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2) = \frac{1}{2}\left(\frac{\mathbf{p}_1^2}{m_1} + \frac{\mathbf{p}_2^2}{m_2}\right) - \frac{G m_1 m_2}{|\mathbf{r}_1 - \mathbf{r}_2|} \ .$$

Consider the proper action of simultaneous translation defined in Example 4.6.2 and corresponding momentum mapping

$$J_0(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2) = \mathbf{p}_1 + \mathbf{p}_2 \ .$$

This is a submersion, therefore, $M_{0,\mathbf{a}} = J_0^{-1}(\mathbf{a})$ is an embedded submanifold of $M_0$ for every $\mathbf{a} \in \mathbb{R}^3$. Also, since the conjugation map (see Section 2.3) on $\mathbb{R}^3$ is the identity, it follows that $J$ is equivariant and that the induced $\mathbb{R}^3_{\mathbf{a}}$-action (Equation (3.9)) is free. Finally, since the Hamiltonian is clearly invariant under simultaneous translations, we can apply the results of Theorem 3.7.4 and Proposition 3.7.5. Thus, the dynamics is reduced to the quotient space $M_{0,\mathbf{a}}/\mathbb{R}^3_{\mathbf{a}} \cong T^*(\mathbb{R}^3 \setminus \{0\})$ where we identify all points that differ by simultaneous translations in the set of points with total momentum $\mathbf{a}$. We introduce the coordinates $(\mathbf{q}, \mathbf{p})$ and $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p})$ on $M_{0,\mathbf{a}}/\mathbb{R}^3_{\mathbf{a}}$ and $M_{0,\mathbf{a}}$ respectively, where $\mathbf{q} = \mathbf{q}_1 - \mathbf{q}_2$ and

$$\mathbf{p}_1 = \frac{m_1}{m_1 + m_2}\mathbf{a} + \mathbf{p} \ , \qquad \mathbf{p}_2 = \frac{m_2}{m_1 + m_2}\mathbf{a} - \mathbf{p} \ .$$

Now, if we consider the inclusion mapping $i_{\mathbf{a}} : M_{0,\mathbf{a}} \to M_0$ we have that

$$i_{\mathbf{a}}^*\omega_0 = \sum_{i=1}^{3} d(i_{\mathbf{a}}^* p_{1i}) \wedge d(i_{\mathbf{a}}^* q_1^i) + d(i_{\mathbf{a}}^* p_{2i}) \wedge d(i_{\mathbf{a}}^* q_2^i) =$$

$$= \sum_{i=1}^{3} dp_i \wedge dq_1^i - dp_i \wedge dq_2^i = \sum_{i=1}^{3} dp_i \wedge d(q_1^i - q_2^i)$$

Thus, the reduced symplectic structure is

$$\omega_0^{\mathbf{a}} = \sum_i dp_i \wedge dq^i$$

as it satifies $\pi_{\mathbf{a}}^*\omega_0^{\mathbf{a}} = i_{\mathbf{a}}^*\omega_0$. Meanwhile, for the Hamiltonian we have

$$(H_0 \circ i_{\mathbf{a}})(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}) = \frac{1}{2}\left(\frac{\mathbf{a}^2}{M} + \frac{\mathbf{p}^2}{\mu}\right) - \frac{GM\mu}{|\mathbf{q}_1 - \mathbf{q}_2|}$$

where $M = m_1 + m_2$ is the total mass and $\mu = m_1 m_2/(m_1 + m_2)$ is the reduced mass. Thus, the reduced Hamiltonian $H_{0,\mathbf{a}}$ on $M_{0,\mathbf{a}}/\mathbb{R}^3_{\mathbf{a}}$ is

$$H_{0,\mathbf{a}}(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\left(\frac{\mathbf{a}^2}{M} + \frac{\mathbf{p}^2}{\mu}\right) - \frac{GM\mu}{q} ,$$

where $q = |\mathbf{q}|$. From the structure of the reduced phase space and the reduced Hamiltonian, it is clear that the dynamics of the two-body problem gets reduced to that of a 3-dimensional point particle subject to a Coulombic potential.

**Remark 4.7.1.** As $\mathbf{a}$ varies, the symplectic structure remains the same while the only change in Hamiltonian is in the the constant term $\mathbf{a}^2/2M$. Therefore, the dynamics on the reduced phase space is independent of the total momentum and we can conveniently set $\mathbf{a}$ equal to zero without loss of generality. Physically, this choice corresponds to choosing the center-of-mass reference frame.

In the following, we will set $\mathbf{a} = \mathbf{0}$ and denote $M \coloneqq \mathbf{M}_{0,\mathbf{a}}$, $\omega \coloneqq \omega_o^{\mathbf{a}}$ and $H \coloneqq H_{0,\mathbf{a}}$. For simplicity's sake, we will set all other constants equal to 1, thus the Hamiltonian becomes

$$H(\mathbf{r}, \mathbf{p}) = \frac{\mathbf{p}^2}{2} - \frac{1}{q} .$$

The Hamiltonian is clearly invariant under the the lift of the action of SO(3) on $\mathbb{R}^3$, which was examined in Example 4.6.3. The corresponding momentum mapping is the angular momentum of the particle

$$J(\mathbf{q}, \mathbf{p}) = \mathbf{q} \times \mathbf{p} .$$

We observe that every non zero value of $J$ is regular, letting us reduce the dynamics to the set $M_{\mathbf{L}_0}/\mathrm{SO}_{\mathbf{L}_0}$ for every $\mathbf{L}_0 \in \mathbb{R}^3 \setminus \{0\}$. The case with zero angular momentum can be studied using a procedure known as Singular Symplectic Reduction (See Sections 10.5 and 10.6 in [3]).
In the case of a classical Lie group, $G \subset \mathrm{End}(V)$, where $V$ is a vector space, we have that $\mathrm{Ad}(a)A = aAa^{-1}$ for all $a \in G$ and $A \in \mathfrak{g}$ (see Example 5.4.6 in [3]). In the case of SO(3), if we consider the isomorphism $\varphi : \mathbb{R}^3 \to \mathfrak{so}(3)$ of Example 2.1.6, given that for $a \in \mathrm{SO}(3)$ and $\mathbf{x} \in \mathbb{R}^3$ we have

$$\varphi(a\mathbf{x}) = a\varphi(\mathbf{x})a^T = \mathrm{Ad}(a)\varphi(\mathbf{x}) \ .$$

it follows that $\varphi$ identifies the adjoint representation of SO(3) with the identical representation on $\mathbb{R}^3$. This implies that under the identification $\mathfrak{so}(3)^* \cong \mathbb{R}^3$, the coadjoint action is

$$\mathrm{Ad}^*(a)\mathbf{b} = a\mathbf{b} \qquad a \in \mathrm{SO}(3),\ \mathbf{b} \in \mathfrak{so}(3)^* \cong \mathbb{R}^3$$

It follows that the coadjoint orbits of $\mathbf{b}$ are labeled by $b = |\mathbf{b}|$ and that its stabilizer is $\mathrm{SO}(3)_{\mathbf{b}} \cong \mathrm{SO}(2)$.
We can finally perform the second reduction. Consider $\mathbf{L}_0$ such that $l = |\mathbf{L}_0| > 0$, making it a regular value of $J$. Thus, the level set

$$J^{-1}(\mathbf{L}_0) = \{(\mathbf{q}, \mathbf{p}) \in M \ : \ \mathbf{q} \times \mathbf{p} = \mathbf{L}_0\}$$

is an embedded submanifold of $M$ which is acted upon freely the stabilizers of $\mathrm{SO}(3)_{\mathbf{L}_0}$. Furthermore, the action is proper and the momentum mapping is equivariant since

$$(J \circ \Psi_a)(\mathbf{q}, \mathbf{p}) = a\mathbf{q} \times a\mathbf{p} = a(\mathbf{q} \times \mathbf{p}) = (\mathrm{Ad}^*(a) \circ J)(\mathbf{q}, \mathbf{p}) \ .$$

Therefore, we can apply the Regular Reduction Theorem which yields the reduced phase space

$$\hat{M}_l = J^{-1}(\mathbf{L}_0)/\mathrm{SO}(3)_{\mathbf{L}_0} \ .$$

We aim to describe $\hat{M}_l$ explicitly. We first note that the constants of motion $\mathbf{L}_0$ reduce the dynamics to the plane orthogonal to the angular momentum, which we take to be $\mathbf{L}_0 = (0, 0, l)$ without loss of generality. Then, the level set $J^{-1}(\mathbf{L}_0)$ is contained in the phase space

$$\tilde{M} = T^*(\mathbb{R}^2 \setminus \{0\}) \cong (\mathbb{R}^+ \times \mathrm{SO}(2)) \times \mathbb{R}^2 \ .$$

Since $\mathbf{q} \neq 0$, we can use polar coordinates $(q, \phi)$ to describe the $q^1$-$q^2$ plane. We denote

$$p_q = p_1 \cos(\phi) + p_2 \sin(\phi) \ , \qquad p_\phi = -qp_1 \sin(\phi) + qp_2 \cos(\phi)$$

and observe that the transformation $(q^1, q^2, p_1, p_2) \mapsto (q, \phi, p_q, p_\phi)$ is a canonical transformation of $\tilde{M}$. In terms of these new coordinates, we have

$$J^{-1}(\mathbf{L}_0) = \left\{(q, \phi, p_q, p_\phi) \in \tilde{M} \colon p_\phi = l\right\} \cong (\mathbb{R}^+ \times \mathrm{SO}(2)) \times \mathbb{R}$$

and
$$H\big|_{J^{-1}(\mathbf{L}_0)} = \frac{1}{2}\left(p_q^2 + \frac{l^2}{q^2}\right) - \frac{1}{q}\ .$$
Furthermore, the action of $\mathrm{SO}(3)_{\mathbf{L}_0}$ in these new coordinates is given by $\phi \mapsto \phi + \alpha$ and the functions $q$, $p_q$, $p_\phi$ and $H|_{J^{-1}(\mathbf{L}_0)}$ are invariant under this actions. Finally, the reduced phase space $\hat{M}_l$ is isomorphic to $\mathbb{R}^+ \times \mathbb{R}$ endowed with the symplectic form $\omega_l = dp_q \wedge dq$ and the reduces Hamiltonian is given by
$$H_l = \frac{1}{2}\left(p_q^2 + \frac{l^2}{q^2}\right) - \frac{1}{q}\ .$$
Therefore, the problem is reduced to a one-dimensional problem with effective potential
$$U(q) = \frac{l^2}{2q^2} - \frac{1}{r}\ .$$
From classical mechanics, it is know that the orbits in configuration space are ellipses for energy values $h < 0$, hyperbolas for $h > 0$ and parabolas for $h = 0$. This can also be shown by purely algebraic arguments by using the hidden simmetry provided by the Lenz-Runge vector (see Example 10.6.3 in [3]).

# Bibliography


[1] Mark Gotay and James Isenberg. "The Symplectification of Science (Symplectic Geometry as a Basis of Physics and Mathematics)". In: *Gazette des Mathématiciens* 54 (Jan. 1992).

[2] Alan Weinstein. "Symplectic Geometry". In: *Bulletin of the American Mathematical Society* 5.1 (1981), pp. 1–13. ISSN: 1088-9485, 0273-0979.

[3] Gerd Rudolph and Matthias Schmidt. *Differential Geometry and Mathematical Physics: Part I. Manifolds, Lie Groups and Hamiltonian Systems.* Theoretical and Mathematical Physics. Dordrecht: Springer Netherlands, 2013. ISBN: 978-94-007-5344-0 978-94-007-5345-7.

[4] Loring W. Tu. *An Introduction to Manifolds.* Universitext. New York, NY: Springer New York, 2011. ISBN: 978-1-4419-7399-3 978-1-4419-7400-6.

[5] Jerrold E. Marsden and Tudor S. Ratiu. *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems.* Red. by Jerrold E. Marsden et al. Vol. 17. Texts in Applied Mathematics. New York, NY: Springer New York, 1999. ISBN: 978-1-4419-3143-6 978-0-387-21792-5.

[6] Ana Cannas Da Silva. *Lectures on Symplectic Geometry.* Vol. 1764. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. ISBN: 978-3-540-42195-5 978-3-540-45330-7.

[7] Stephanie Frank Singer. *Symmetry in Mechanics.* Boston, MA: Birkhäuser Boston, 2004. ISBN: 978-0-8176-4145-0 978-1-4612-0189-2.

[8] Daniel S. Freed, ed. *Geometry and Quantum Field Theory: This Volume Contains the Lecture Notes from the Graduate Summer School Program on the Geometry and Topology of Manifolds and Quantum Field Theory, Held June 22 - July 20, 1991, in Park City, Utah.* Repr. IAS Park City Mathematics Series 1. Princeton: American Math. Soc., Inst. for Advanced Study. 459 pp. ISBN: 978-0-8218-0400-1.

[9] Antonio Fasano and Stefano Marmi. *Analytical Mechanics: An Introduction.* Oxford Graduate Texts. Oxford: Oxford University Press, 2006. ISBN: 978-0-19-850802-1.