Symplectic Reduction and Applications in Hamiltonian Mechanics

Rezaei Ghavamabadi, L.

L.S.C. Rezaei Ghavamabadi

Symplectic Reduction and Applications in Hamiltonian Mechanics

Bachelor thesis

July 2023

Thesis supervisors: **dr. F. Pasquotto**
dr. S.P. Patil

Leiden University
Mathematical Institute

Contents

1 Introduction **1**

I Symplectic Preliminaries **2**

2 Conventions **2**

3 Symplectic linear algebra **3**

4 Symplectic geometry **5**
4.1 Definitions and the Darboux Theorem 5
4.2 Symplectic structure of the cotangent bundle 6

5 Hamiltonian mechanics **8**
5.1 Vector fields and flows 8
5.2 Phase space 9
5.3 The simple pendulum 13

II Momentum Maps and Symplectic Reduction **17**

6 Lie theory **17**
6.1 Basic definitions and properties 17
6.2 Classical groups and computations 22
6.3 Adjoint and coadjoint representations 25

7 Reduction **26**
7.1 Actions and momentum maps 26
7.2 The momentum map for an N-particle system 31
7.3 The Marsden-Weinstein and Noether Theorems 32
7.4 An example of reduction 35
7.5 Singular reduction 36

III Applications **38**

8 Symmetry breaking **38**
8.1 Reduction of SU(2) 38
8.2 Reduction of SO(3) 40

9 Closing Remarks **41**

IV Appendix **42**

A Elastic surface deformations **42**

References **45**

1 Introduction

Physicists often face problems which exhibit some kind of symmetry, be it a spherically symmetric gravitational field, a translationally symmetric magnetic field, or something more abstract altogether. In these situations, one can often reformulate the problem in a different set of coordinates that makes it easier to work with. The reformulated problem has fewer degrees of freedom.

Well-known examples of such techniques arise from celestial mechanics; in an N-body problem, one can position the frame of reference in such a way that one particle is stationary. Similarly, spherically symmetric potentials are often studied as 1-dimensional problems, which is also a case of reducing the degrees of freedom by exploiting symmetries of the problem.

In this report we study the underlying mathematics of reducing the degrees of freedom of a physical system. To that end, we require a rigorous formulation of classical mechanics which adopts the language of differential geometry. This is done using *symplectic manifolds*, which are objects that can be seen as a generalisation of the concept of phase space in physics. Within this framework, we apply a technique known as *symplectic reduction* to reduce the degrees of freedom of a system. To restrict the scope of this text, the reader is assumed to already be familiar with differential geometry.

We will begin by laying the groundworks for the theory of symplectic geometry. In Section 3 we study some preliminary linear algebra, including a linear version of what is known as the Darboux Theorem. This section additionally serves as a microcosm of the following Section 4, which gives the definition of symplectic manifolds and mentions the geometric Darboux Theorem, though it is not proven in this text. Here some interesting properties of cotangent bundles are also brought to light.

Section 5 concludes this part of the thesis with a symplectic-geometric reformulation of certain systems in classical mechanics. Here the relation between phase spaces and symplectic manifolds is explained, and we treat a small example of the simple pendulum using this new formulation.

The second part of this report is dedicated to Lie theory and symplectic reduction, the latter of which will ultimately be our goal. Lie theory is necessary to describe smooth symmetries of symplectic manifolds, and instrumental in defining symplectic reduction. Hence we spend some time working out the involved mathematics. After formulating some notable theorems regarding symplectic reduction, we begin applying the theory to physical problems. This is the focus of the third and final part.

Part I
Symplectic Preliminaries

2 Conventions

The following sections are heavily inspired by [3], in particular some of the materials up to chapter 6 and some of the material from chapter 18 onwards. We will cover part of the contents of these lecture notes, and work with the same preliminary knowledge about smooth manifolds. For a comprehensive overview of the general theory of smooth manifolds, we refer to [9]. To avoid ambiguity, below is a list of conventions and terminology that will be used throughout this report.

smooth	infinitely differentiable
S^n	n-sphere
$\mathbb{T}^n$	n-torus
$\mathbb{RP}^n, \mathbb{CP}^n$	real resp. complex projective space (manifolds of dimensions n resp. $2n$)
T_pN	tangent space at p to a smooth manifold N
TN	tangent bundle of a smooth manifold N
T_p^*N	cotangent space at p to a smooth manifold N
T^*N	cotangent bundle of a smooth manifold N
$\Omega^k(N)$	space of differential k-forms on a smooth manifold N
$\mathcal{X}(N)$	space of smooth vector fields on a smooth manifold N
$C^\infty(N)$	space of smooth mappings $N \to \mathbb{R}$ on a smooth manifold N
$(\mathrm{d}\omega)_p$	exterior derivative of a differential form ω at a point p on the manifold
$(\mathrm{d}F)_p$	differential of a smooth map F at a point p in its domain
$F^*\omega$	pullback of a differential form ω along a smooth map F
$X \lrcorner\, \omega$	contraction of a differential form ω along a vector field X
Stab_x	stabiliser of a point x under a group action
$\mathrm{Span}(S)$	span of a subset S of a vector space V

Furthermore, the term "smooth manifold" will often be shortened to just "manifold". If the smoothness property is at any point dropped, this will be explicitly stated.

3 Symplectic linear algebra

In the following sections we lay the groundworks for the study of symplectic geometry. This begins in a linear algebraic context, which will define part of the symplectic structures on manifolds.

Definition 3.1 (Symplectic vector space)**.** *Let V be a finite-dimensional real vector space. A **linear symplectic form** is a bilinear form $\Omega : V \times V \to \mathbb{R}$ that is skew-symmetric and nondegenerate. In this case, we call (V, Ω) a **symplectic vector space**.*

For the nondegeneracy property we may use one of the equivalent formulations

$$\forall v \in V : \Omega(v, \cdot) = 0 \Rightarrow v = 0$$

or that

$$\alpha : V \to V^*, v \mapsto \Omega(v, \cdot)$$

is an isomorphism.

The skew-symmetry of the symplectic form leads to a decomposition of the vector space into two halves. This will be further explored after the example below, and will have far-ranging effects down the line.

Example 3.2. *Consider $\mathbb{R}^4$ with basis (e_1, e_2, f_1, f_2). Denote the dual basis as $(e_1^*, e_2^*, f_1^*, f_2^*)$ and define a symplectic form Ω_0 on $\mathbb{R}^4$ by*

$$\Omega_0 = e_1^* \wedge f_1^* + e_2^* \wedge f_2^*.$$

By explicit computation we find

$$\Omega_0(e_i, f_j) = \delta_{ij} \text{ and } \Omega_0(e_i, e_j) = 0 = \Omega_0(f_i, f_j),$$

where δ_{ij} is 1 if $i = j$ and 0 otherwise. We will show that all linear symplectic forms can be written in a similar way.

Any bilinear form gives rise to a concept of complements; given a linear subspace $Y \subseteq V$, we define the **symplectic complement** with respect to Ω as

$$Y^\Omega := \{v \in V \mid \forall y \in Y : \Omega(v, y) = 0\}.$$

Unfortunately, this is not always complementary to Y; in the example above, we would find that

$$\text{Span}(e_1)^{\Omega_0} = \text{Span}(e_1, e_2, f_2),$$

which contains $\text{Span}(e_1)$. For Y and Y^Ω to be complementary, we need Y to be a **symplectic subspace** of V, meaning that $\Omega|_{Y \times Y}$ must be nondegenerate. Let us first formulate the theorem we want to prove.

Theorem 3.3. *Let* (V, Ω) *be a symplectic vector space. Then it has even dimension* $\dim V = 2n$ *and a* ***Darboux basis*** $(x_1, \dots, x_n, y_1, \dots, y_n)$ *with respect to which* Ω *can be written as*

$$\Omega = \sum_{j=1}^{n} x_j^* \wedge y_j^*,$$

where $v^* := \Omega(v, \cdot)$ *for* $v \in V$.

This is a fairly important theorem, and also has a geometric analogue. This will be discussed at the end of the section. First, we show some properties of symplectic vector spaces in order to prove the above.

Proposition 3.4. *Let* (V, Ω) *be a symplectic vector space with a subspace* $Y \subseteq V$.

(1) $\dim Y + \dim Y^\Omega = \dim V$.

(2) If Y *is symplectic,* $Y \cap Y^\Omega = 0$ *and so* $V = Y \oplus Y^\Omega$.

Proof. For (1), consider the linear map

$$\begin{aligned} \varphi : V &\to Y^* \\ v &\mapsto \Omega(v, \cdot)|_Y, \end{aligned}$$

whose kernel is $\ker \varphi = Y^\Omega$. Since the dimensions of the kernel and image of φ sum to $\dim V$, we want it to be surjective so that (1) is satisfied. Recall that the nondegeneracy of Ω gives an isomorphism $\alpha : V \to V^*$. Then we obtain φ as $\varphi = \iota^* \circ \alpha$, where ι^* is the dual mapping of the inclusion $\iota : Y \hookrightarrow V$. ι^* and α are both surjective, so φ must be too. We obtain

$$\dim V = \dim Y^* + \dim Y^\Omega = \dim Y + \dim Y^\Omega.$$

In (2), take any element $y \in Y \cap Y^\Omega$. Then since $y \in Y^\Omega$, we have that $\Omega|_{Y \times Y}(y, \cdot) = \Omega(y, \cdot)|_Y = 0$. Since $\Omega|_{Y \times Y}$ is symplectic, it follows that $y = 0$, so $Y \cap Y^\Omega = 0$. □

It turns out that (2) in the above is actually an equivalence. For our purposes, we just need the consequence that $V = Y \oplus Y^\Omega$ is a decomposition into *symplectic* subspaces. As one might imagine, this lends itself to an iterative proof of Theorem 3.3.

Corollary 3.5. *If* Y *is a symplectic subspace of* (V, Ω)*, then so is* Y^Ω.

Proof. We begin with the claim that $(Y^\Omega)^\Omega = Y$. From Proposition 3.4(1) it follows that

$$\begin{aligned} \dim (Y^\Omega)^\Omega + \dim Y^\Omega &= \dim V \\ &= \dim Y + \dim Y^\Omega, \end{aligned}$$

and hence

$$\dim (Y^\Omega)^\Omega = \dim Y.$$

Furthermore, if $y \in Y$, then for all $v \in Y^\Omega$ we have $\Omega(y, v) = 0$. So by definition of the symplectic complement, $y \in (Y^\Omega)^\Omega$. Therefore

$$Y \subseteq (Y^\Omega)^\Omega$$

and since the dimensions are equal, $(Y^\Omega)^\Omega = Y$.[1]

Now we need to show that Y^Ω is indeed a symplectic subspace, or equivalently, that $\Omega|_{Y^\Omega \times Y^\Omega}$ is nondegenerate. Let $v \in Y^\Omega$ be given and suppose $\Omega|_{Y^\Omega \times Y^\Omega}(v, \cdot) = 0$. Then $\Omega(v, \cdot)|_{Y^\Omega} = 0$, so $v \in (Y^\Omega)^\Omega = Y$. But now v is in the intersection $Y \cap Y^\Omega$, which is 0 by Proposition 3.4(2). Hence $v = 0$ and $\Omega|_{Y^\Omega \times Y^\Omega}$ is indeed a linear symplectic form. □

Now we are ready to prove Theorem 3.3

Proof of Theorem 3.3. The case $V = 0$ is trivial, so suppose $\dim V > 0$. First note that $\dim V \neq 1$, since Ω vanishes on 1-dimensional subspaces due to its skew-symmetry and would hence be degenerate. So $\dim V \geqslant 2$.

Now fix an $x_1 \in V \backslash \{0\}$, and let $y_1 \in V$ be such that $\Omega(x_1, y_1) \neq 0$. The existence of such y_1 is required by the nondegeneracy of Ω. Now normalise y_1 such that $\Omega(x_1, y_1) = 1$. This makes $W = \operatorname{Span}(x_1, y_1)$ a symplectic subspace of V with

$$\Omega|_{W \times W} = x_1^* \wedge y_1^*.$$

Due to Proposition 3.4(2) and Corollary 3.5 we can now decompose

$$V = W \oplus W^\Omega$$

as a direct sum of symplectic subspaces. The process above can be applied iteratively to W^Ω in order to extract 2-dimensional symplectic subspaces until $\dim W^\Omega < 2$. With W^Ω being symplectic, it cannot have dimension 1, so this in fact yields V as a direct sum of 2-dimensional symplectic subspaces with the desired Darboux basis. □

Theorem 3.3 shows that symplectic vector spaces of the same dimension are not only isomorphic in the traditional sense, but also necessarily have the same symplectic structure. In the following section we formulate a geometric version of the same theorem, although we will not prove it. Essentially, it states that all symplectic manifolds locally have the same structure.

4 Symplectic geometry

4.1 Definitions and the Darboux Theorem

With the symplectic linear algebra we have discussed, we can now give a definition of symplectic manifolds and the associated isomorphisms.

[1] This result holds for any subspace of a symplectic vector space.

Definition 4.1 (Symplectic manifold). *Let M be a smooth manifold with a differential 2-form $\omega \in \Omega^2(M)$ satisfying the conditions*

(1) ω is closed

(2) $\forall p \in M : \omega_p$ is nondegenerate[2] *as a bilinear form $T_pM \times T_pM \to \mathbb{R}$*

*Then we call ω a **symplectic form** and (M, ω) a **symplectic manifold**.*

Definition 4.2 (Symplectomorphism). *Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds. A **symplectomorphism** from M_1 to M_2 is a diffeomorphism $\varphi : M_1 \to M_2$ which preserves the symplectic forms: $\varphi^*\omega_2 = \omega_1$.*

Note that the nondegeneracy of the symplectic form makes the tangent spaces of a symplectic manifold into symplectic vector spaces. Hence symplectic manifolds, just like symplectic vector spaces, are always even-dimensional. The closedness is a less intuitive condition, but important for technical reasons. For one, it is necessary for the Darboux Theorem to hold, which is the smooth version of Theorem 3.3, and will be essential in our treatment of classical mechanics. The standard example for a symplectic manifold is the following.

Example 4.3. *$\mathbb{R}^{2n}$ with coordinates $(x_1, \ldots, x_n, y_1, \ldots, y_n)$ has a standard symplectic form*

$$\omega_0 = \sum_{j=1}^{n} \mathrm{d}x_j \wedge \mathrm{d}y_j.$$

There is a lot of interesting mathematics we could get into regarding classification, but this is not necessary for the problems we will be studying. It is therefore also beyond the scope of this report to prove the Darboux Theorem, even though we will be using it extensively.

Theorem 4.4 (Darboux). *Let (M, ω) be a $2n$-dimensional symplectic manifold. Given a point $p \in M$, there exists an open neighbourhood U of p with local coordinates $(x_1, \ldots, x_n, y_1, \ldots, y_n)$ such that on U*

$$\omega = \sum_{j} \mathrm{d}x_j \wedge \mathrm{d}y_j.$$

For a proof of this theorem, we refer to the first three parts of [3].

4.2 Symplectic structure of the cotangent bundle

The cotangent bundle of an arbitrary manifold can always be given a smooth and symplectic structure. These will be the most notable symplectic manifolds to us in the study of classical mechanics. We will see later that cotangent bundles correspond to the concept of phase space in physics. First, let us construct the smooth structure on the cotangent bundle.

Proposition 4.5. *Let N be an n-manifold, then T^*N is a $2n$-manifold.*

[2] So the mapping $T_pM \to T_p^*M$ given by $v \mapsto \omega_p(v, \cdot)$ is a linear isomorphism.

Proof. We cover N with coordinate charts $(U, x_1, \ldots, x_n)$, so that at any $p \in U$, we find a basis $((\mathrm{d}x_1)_p, \ldots, (\mathrm{d}x_n)_p)$ for T_p^*N. In terms of this basis, elements of T_p^*N can be written as $\xi = (\xi_1, \ldots, \xi_n)$, which in turn induces a local coordinate chart

$$(T^*U, x_1, \ldots, x_n, \xi_1, \ldots, \xi_n)$$

on T^*N. The coordinates $(x_1, \ldots, x_n, \xi_1, \ldots, \xi_n)$ are called the **cotangent coordinates** associated to $(x_1, \ldots, x_n)$. We need to show that these coordinates induce a smooth structure. Given two charts $(U, x_1, \ldots, x_n)$ and $(U', x'_1, \ldots, x'_n)$, $p \in U \cap U'$ and $\xi \in T_p^*N$ we find

$$\xi = \sum_i \xi_i (\mathrm{d}x_i)_p = \sum_i \xi_i \sum_j \frac{\partial x_i}{\partial x'_j} (\mathrm{d}x'_j)_p = \sum_j \left(\sum_i \xi_i \frac{\partial x_i}{\partial x'_j} \right) (\mathrm{d}x'_j)_p.$$

It follows that the transition function (defined by the components of ξ in the $\mathrm{d}x'_j$-basis) is smooth. □

From the above, we might want to define a symplectic form on T^*N by

$$\omega = \sum_j \mathrm{d}x_j \wedge \mathrm{d}\xi_j.$$

This is indeed a symplectic form, since it has the form given in the Darboux Theorem. But is it coordinate-independent? Well, it turns out that it is not only that, but also canonical. To see this, consider the 1-form

$$\alpha = \sum_j \xi_j \mathrm{d}x_j.$$

We find that $\omega = -\mathrm{d}\alpha$. It is not immediately clear that α is coordinate-indepent, but with some computation we would find that it is given by the following coordinate-free definition.

Definition 4.6 (Tautological 1-form and canonical symplectic form)**.** *Given a manifold N with cotangent bundle projection $\pi : T^*N \to N$, we define the **tautological 1-form** or **Liouville 1-form** $\alpha \in \Omega^1(T^*N)$ at $(p, \xi) \in T^*N$ by*

$$\alpha_{(p,\xi)} := (\mathrm{d}\pi)^*_{(p,\xi)} \xi = \xi \circ (\mathrm{d}\pi)_{(p,\xi)}.$$

*The **canonical symplectic form** on T^*N is then defined as*

$$\omega = -\mathrm{d}\alpha.$$

To make sense of the definition of α, note that π has a differential

$$(\mathrm{d}\pi)_{(p,\xi)} : T_{(p,\xi)}(T^*N) \to T_pN,$$

whose pullback is

$$(\mathrm{d}\pi)^*_{(p,\xi)} : T_p^*N \to T^*_{(p,\xi)}(T^*N).$$

From this we can see that $(\mathrm{d}\pi)^*_{(p,\xi)}\xi$ indeed defines a 1-form. We can easily check that

$$\mathrm{d}\pi\left(\frac{\partial}{\partial x_j}\right) = \frac{\partial}{\partial x_j} \text{ and } \mathrm{d}\pi\left(\frac{\partial}{\partial \xi_j}\right) = 0$$

so indeed $\alpha(\frac{\partial}{\partial \xi_j}) = 0$, and

$$\begin{aligned}\alpha\left(\frac{\partial}{\partial x_j}\right) &= (\xi \circ \mathrm{d}\pi)\left(\frac{\partial}{\partial x_j}\right)\\ &= \xi\left(\frac{\partial}{\partial x_j}\right)\\ &= \xi_j,\end{aligned}$$

as expected.

5 Hamiltonian mechanics

Hamiltonian mechanics is a particular treatment of classical mechanics that lends itself to both elegant and intuitive formulations of mechanical problems. It revolves around the study of a class of functions called Hamiltonians. They can be viewed as the total energy of a mechanical system, that being the sum of the kinetic and potential energy. In this section we define Hamiltonians and lay the groundworks for Hamiltonian mechanics.

5.1 Vector fields and flows

Vector field flows are going to be a fairly important class of objects, so this section is dedicated to a brief recap of the topic before we begin working with them. We will only work with *complete* vector fields, primarily because this makes the flow a group action. There is also a physical justification for this, discussed in Remark 5.4.

Consider a vector field X on a manifold N. We denote its flow by $\exp tX$ for $t \in \mathbb{R}$, a notation which will make more sense once we touch upon Lie Theory later on. Because of the notation, we may also call the flow of X its **exponential map**. It is defined such that $\{\exp tX : N \to N \mid t \in \mathbb{R}\}$ is the unique smooth family of diffeomorphisms satisfying the following system of differential equations:

$$\begin{cases} \exp tX|_{t=0} &= \mathrm{id} \\ \frac{\mathrm{d}}{\mathrm{d}t}\exp tX &= X \circ \exp tX \end{cases}$$

This is fortunately quite intuitive; it assigns to the exponential map the role of defining an action of $\mathbb{R}$ on N. This action essentially moves a point on N along the vector field X for a certain amount of "time" $t \in \mathbb{R}$. The first equation says that moving a point for a time $t = 0$

leaves it in place, and the second equation makes it so that the velocity with which $\exp tX$ moves points corresponds to the value of X at those points.

Flows also give rise to a notion of taking derivatives along a vector field. This is known as the **Lie derivative**, defined as follows:

$$\begin{aligned} \mathcal{L}_X : \Omega^k(N) &\to \Omega^k(N) \\ \omega &\mapsto \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left[(\exp tX)^*\omega\right] \end{aligned}$$

In the case $k = 0$, this just turns out to be the same as taking a partial derivative along X. Along with the Lie derivative of course comes Cartan's magic formula, which will be a great companion to us.

Lemma 5.1 (Cartan's Magic Formula). *For $X \in \mathcal{X}(N)$ and $\omega \in \Omega^k(N)$ we have*

$$\mathcal{L}_X\omega = X \lrcorner \mathrm{d}\omega + \mathrm{d}[X \lrcorner \omega].$$

This can be proven by induction on k using the identity

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

For details, see [9, p. 373].

5.2 Phase space

In this section we work towards a Hamiltonian formulation of mechanical systems.

To begin, let (M, ω) be a symplectic manifold and $H \in C^\infty(M)$ a smooth function. Since ω is nondegenerate, we have an isomorphism

$$\begin{aligned} \mathcal{X}(M) &\xrightarrow{\sim} \Omega^1(M) \\ X &\mapsto X \lrcorner \omega. \end{aligned}$$

So there is a unique vector field $X_H \in \mathcal{X}(M)$ such that

$$\mathrm{d}H = X_H \lrcorner \omega.$$

Definition 5.2. *In the above, we call H a **Hamiltonian function** and X_H the unique associated **Hamiltonian vector field**.*

Note that

$$\begin{aligned} \mathcal{L}_{X_H} H &= X_H \lrcorner \mathrm{d}H \\ &= X_H \lrcorner (X_H \lrcorner \omega) \\ &= \omega(X_H, X_H) \\ &= 0, \end{aligned}$$

due to the skew-symmetry of ω. It follows that H does not change along integral curves of X_H; in other words, the flow of X_H **preserves** H. This is a valuable insight into the associated physics; since H can be viewed as the total energy of a system, we can view the integral curves as the possible trajectories an object can take with constant energy. Hence, energy is conserved on integral curves of X_H. This becomes clearer when we look at Hamilton's equations towards the end of this section.

Example 5.3. *Consider the symplectic manifold $(S^2, \mathrm{d}\theta \wedge \mathrm{d}h)$ with coordinates given by the angle $\theta \in [0, 2\pi)$ and height $h \in [-1, 1]$. Define the Hamiltonian $H(\theta, h) = h$ as the height function. With*

$$X_H \lrcorner (\mathrm{d}\theta \wedge \mathrm{d}h) = \mathrm{d}h.$$

we find $X_H = \frac{\partial}{\partial\theta}$. This has flow $\rho_t(\theta, h) = (\theta + t, h)$. We see that the level curves $H^{-1}(h)$ indeed turn out to be the integral curves at the respective heights h.

Note that S^2 is compact, and X_H complete. This is no coincidence; on compact manifolds, all vector fields are complete. *See Theorem 9.16 of [9, p. 216] for a proof of this fact.*

Remark 5.4. *Earlier we remarked that we will only need complete vector fields in our study of symplectic reduction. We already saw that this makes the flow a group action, which is certainly important, but this is also justified from a physics standpoint. Above we found that integral curves of Hamiltonian vector fields represent trajectories allowed in the mechanical system defined by the Hamiltonian. If this vector field were incomplete, then those trajectories may be discontinuous. This amounts to non-physical solutions without proper care.*

Example 5.5. *As an example of the remark above, consider a particle of mass m moving in $\mathbb{R}^3 \backslash \{0\}$ under a Coulomb potential*

$$V(r) = -\frac{1}{r},$$

expressed in spherical coordinates. Since we have not yet discussed the mathematical language required to describe this system, let us settle for the physicist's formulation; the Hamiltonian is the sum of kinetic and potential energy, written as

$$H(r, p) = \frac{p^2}{2m} - \frac{1}{r},$$

where p is the total momentum. The trajectory of the particle will have a constant energy E, so setting $H(r, p) = E$ we find

$$p = \sqrt{2mE + \frac{2m}{r}}.$$

Now we can see that we run into a great deal of trouble around the origin, as the momentum diverges to infinity. The reason can be understood through both a mathematical and physical lens; on the one hand, the Hamiltonian vector field corresponding to H is not complete. And on the other hand, approaching the problem via classical mechanics is not physically sound; our particle travels at relativistic velocities (i.e., close to the speed of light) near the origin, *and hence requires a relativistic treatment. To speak in informal terms for a moment, relativistic effects should stretch out the time required to reach the origin infinitely; hence, a proper relativistic approach would lead to a well-defined complete Hamiltonian vector field.*

Notice that Hamiltonian vector fields preserve the symplectic form. This motivates the following definition of symplectic vector fields, which play a role later when we define smooth actions on manifolds.

Definition 5.6. *Let (M, ω) be a symplectic manifold and $X \in \mathcal{X}(M)$ a vector field which preserves ω (so $\mathcal{L}_X \omega = 0$). Then we call X a* ***symplectic vector field****.*

Since $\mathrm{d}\omega = 0$, we have $\mathcal{L}_X \omega = \mathrm{d}[X \lrcorner \omega]$ (with Cartan's magic formula). Hence X is symplectic if and only if $X \lrcorner \omega$ is closed. An analogous statement holds for Hamiltonian vector fields and exactness (by definition).

We now have the equivalences

$$X \text{ is Hamiltonian} \iff X \lrcorner \omega \text{ is exact}$$

and

$$\begin{aligned} X \text{ is symplectic} &\iff X \lrcorner \omega \text{ is closed} \\ &\iff \mathcal{L}_X \omega = 0 \\ &\iff \rho_t^* \omega = \omega \text{ for all } t. \end{aligned}$$

The last equivalence follows from the fact that $\mathcal{L}_X \omega = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \rho_t^* \omega$ and $\rho_0^* \omega = \mathrm{id}^* \omega = \omega$.

Now that we have defined Hamiltonians, we want to consider them in a physical context. To this end, consider an arbitrary n-dimensional manifold N with local coordinates $(q_1, \ldots, q_n)$. Every point in N represents a particular configuration of our mechanical system. For instance, if we look at a particle in three dimensions, then we would take $N = \mathbb{R}^3$. In any case, we call N the **configuration space** of our system, and often encodes positions. A Hamiltonian takes not only positions, but also momenta as arguments. Hence, we want to look at the **phase space** T^*N, with cotangent coordinates denoted as $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ locally, or (q, p) for short. The phase space is the space of all configuration coordinates and the corresponding momenta.

We now want to derive the equations governing a Hamiltonian system. Consider the canonical symplectic form $\omega = -\mathrm{d}\alpha$, which locally takes the form

$$\omega = \sum_j \mathrm{d}q_j \wedge \mathrm{d}p_j$$

and a Hamiltonian $H \in C^\infty(T^*N)$ with associated Hamiltonian vector field X_H. Call the components Q_j and P_j such that

$$X_H = \sum_j \left(Q_j \frac{\partial}{\partial q_j} + P_j \frac{\partial}{\partial p_j} \right).$$

We want to determine these components. We compute

$$\mathrm{d}H = \sum_j \left(\frac{\partial H}{\partial q_j} \mathrm{d}q_j + \frac{\partial H}{\partial p_j} \mathrm{d}p_j \right)$$

and find

$$\begin{aligned} X_H \lrcorner\, \omega &= \sum_j X_H \lrcorner\, (\mathrm{d}q_j \wedge \mathrm{d}p_j) \\ &= \sum_j ((X_H \lrcorner\, \mathrm{d}q_j) \wedge \mathrm{d}p_j - \mathrm{d}q_j \wedge (X_H \lrcorner\, \mathrm{d}p_j)) \\ &= \sum_j \left(Q_j \mathrm{d}p_j - P_j \mathrm{d}q_j \right). \end{aligned}$$

The requirement $X_H \lrcorner\, \omega = \mathrm{d}H$ gives the components Q_j and P_j of X_H such that

$$X_H = \sum_j \left(\frac{\partial H}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial H}{\partial q_j} \frac{\partial}{\partial p_j} \right).$$

Let $(q(t), p(t))$ be an integral curve of X_H. Then from the properties of vector field flows, we find the system of differential equations

$$\begin{cases} \dfrac{\mathrm{d}q_j}{\mathrm{d}t}(t) = \dfrac{\partial H}{\partial p_j} \\ \dfrac{\mathrm{d}p_j}{\mathrm{d}t}(t) = -\dfrac{\partial H}{\partial q_j} \end{cases}$$

using $X_H(q_j(t)) = \frac{\partial H}{\partial p_j}$ and $X_H(p_j(t)) = -\frac{\partial H}{\partial q_j}$. This system of equations is known as **Hamilton's equations**.

Note that these equations are equivalent with Newton's second law. If we consider a point mass m moving through $N = \mathbb{R}^3$, then we can denote the Hamiltonian as

$$H(q,p) = \frac{p^2}{2m} + V(q),$$

where V is a potential, which describes the forces acting on the mass. Then

$$\dot{q}_j = \frac{\partial H}{\partial p_j} = \frac{p_j}{m} \iff \dot{p} = m\ddot{q},$$

and

$$\dot{p}_j = -\frac{\partial H}{\partial q_j} = -\frac{\partial V}{\partial q_j} \iff \dot{p} = -\nabla V.$$

Hence we obtain Newton's second law

$$m\ddot{q} = -\nabla V,$$

with on the left the net force and on the right the forces arising from the potential.

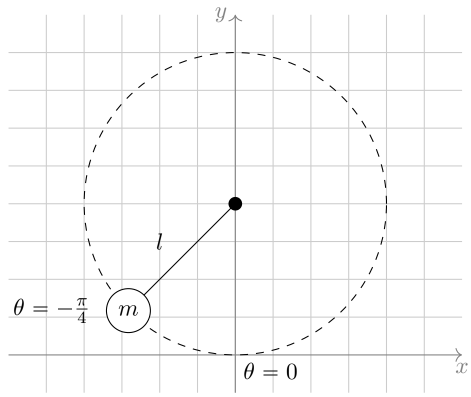


Figure 1: The configuration space of the simple pendulum.

5.3 The simple pendulum

This example is adapted from Homework 13 of [3].

Here we treat the mechanical system of the simple pendulum. This is described by a mass m hung from a fixed point by a rigid rod of length l. Our configuration space is then S^1. We can choose our coordinate $\theta \in (-\pi, \pi)$ such that $\theta = 0$ corresponds to the equilibrium position; see Figure 1. In this problem we only consider a downward pointing gravitational force acting on the mass.

We begin by constructing the Hamiltonian for our system. We begin by computing the momentum of the mass, by embedding S^1 into $\mathbb{R}^2$ through

$$\begin{aligned} \iota : S^1 &\hookrightarrow \mathbb{R}^2 \\ \theta &\mapsto (l \sin\theta, l(1-\cos\theta)), \end{aligned}$$

so $\theta = 0$ corresponds to $(0,0)$. If we consider a trajectory $(x(t), y(t))$ on $\iota[S^1]$ for now, then the corresponding squared momentum is

$$p^2 = m^2(\dot{x}^2 + \dot{y}^2) = m^2(l^2\dot{\theta}^2\cos^2\theta + l^2\dot{\theta}^2\sin^2\theta) = m^2l^2\dot{\theta}^2.$$

Note that the angular momentum of the mass is given by

$$\xi = pl = ml^2\dot{\theta}.$$

Since our system exhibits rotational motion, we take ξ as the cotangent coordinate on T^*S^1. This gives the kinetic energy as

$$K = \frac{p^2}{2m} = \frac{\xi^2}{2ml^2}.$$

The gravitational potential is of the form

$$V = mgy = mgl(1-\cos\theta),$$

with g the gravitational acceleration. Now we find our Hamiltonian as

$$H(\theta, \xi) = \frac{\xi^2}{2ml^2} + mgl(1 - \cos\theta).$$

Note that Hamilton's equations can now be computed as

$$\dot{\theta} = \frac{\xi}{ml^2} \text{ and } \dot{\xi} = -mgl\sin\theta.$$

Combining these equations gives

$$\ddot{\theta} + \frac{g}{l}\sin\theta = 0,$$

which is the well-known equation of motion of the simple pendulum.

We may also study the level curves of H, which contain the trajectories of constant energy; set $H = E$, then we have

$$\frac{\xi^2}{2ml^2} + mgl(1 - \cos\theta) = E,$$

which can be rewritten to

$$\xi^2 = 2m^2l^3g\left(\frac{E}{mgl} - 1 + \cos\theta\right).$$

This has solutions for both $\xi \geqslant 0$ and $\xi \leqslant 0$. We set $\alpha^2 = 2m^2l^3g$ and $\beta = \frac{E}{mgl} - 1$ so

$$\xi^2 = \alpha^2(\beta + \cos\theta).$$

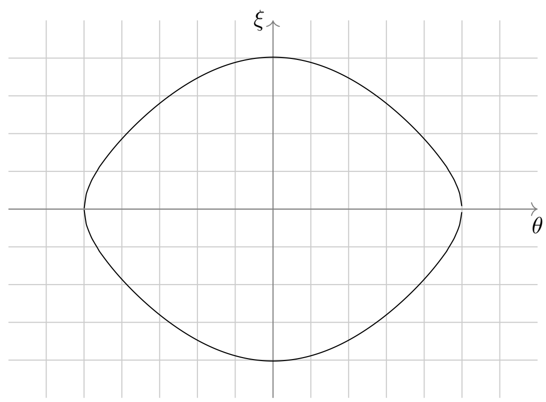
Note that we chose our potential such that its lowest value is 0; hence only values $\beta \geqslant -1$ give physical solutions. Three level curves are plotted in Figure 2 for different values of α and β.

In the plots, we see that some level curves are connected, while others consist of two connected components. We noted that there is a '$\xi \geqslant 0$' and '$\xi \leqslant 0$' component due to the expression $\xi^2 = \alpha^2(\beta + \cos\theta)$. So the connectedness of the level curve depends on whether it contains a point with $\xi = 0$. These points have $\beta + \cos\theta = 0$, so $H^{-1}(E)$ is connected if and only if $-1 \leqslant \beta \leqslant 1$. This connectedness has a very clear physical meaning; if $\beta \leqslant 1$, then the mass does not have enough energy to go *over* the uppermost point of the pendulum.[3] In most cases, it is pulled back down before it reaches the top, flipping the sign of ξ. If $\beta > 1$, the mass moves so fast that its direction never changes.

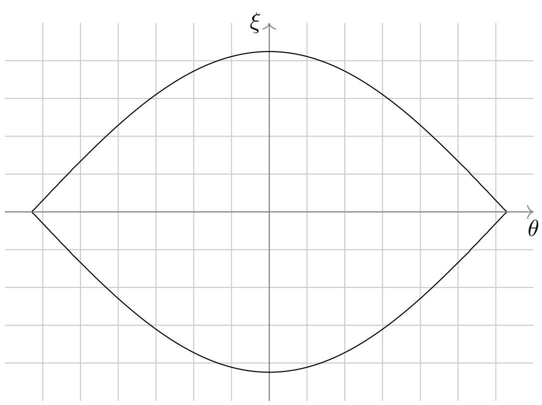
Two special cases arise when $\beta = \pm 1$; we can compute that $\beta = -1$ is the case $E = 0$, and $\beta = 1$ is the case $E = 2mgl$ (see the corresponding plot in Figure 2). In the former, our mass never moves since it has no energy; it remains at rest in $\theta = 0$. In the latter case, we note that $2mgl$ is the potential energy at the topmost position. If the mass has this as its total energy, then its kinetic energy must vanish at the top; hence, with this exact energy, the particle always ends up at rest in the point $\theta = \pi$.

To summarise, the simple pendulum has the following types of trajectories:

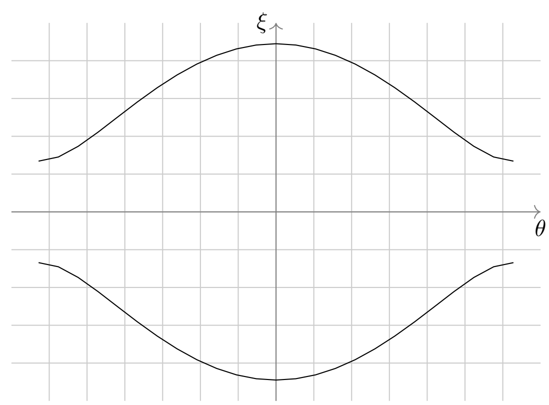
[3] We technically did not define the coordinates $\theta = \pm\pi$, but our discussion remains valid if we identify $S^1 \cong \mathbb{R}/2\pi\mathbb{Z}$. This does not change any of the expressions we have derived, due to the periodicity of cos and sin.



(a) $\alpha = 1.5$, $\beta = 0.8$



(b) $\alpha = 1.5$, $\beta = 1$



(c) $\alpha = 1.5$, $\beta = 1.2$

Figure 2: The level curves of H, for given values of α and β. Note that the lower two graphs should be considered as living on a cylinder where their left and right ends are joined, since the coordinates $\theta = \pm\pi$ correspond to the same point on S^1.

- $E = 0$. The mass is always at rest in $\theta = 0$.
- $0 < E < 2mgl$. The mass oscillates indefinitely around $\theta = 0$, and never reaches $\theta = \pi$.
- $E = 2mgl$. The mass rotates in one direction, and stops when it reaches $\theta = \pi$.
- $E > 2mgl$. The mass rotates indefinitely in one direction.

Part II
Momentum Maps and Symplectic Reduction

6 Lie theory

This section discusses the basics of the theory of Lie groups, which are groups with a compatible smooth structure. We need the smoothness to define smooth actions, which are required for symplectic reduction. The treatment of this topic is heavily inspired by Chapters 2 and 3 of [7], and Chapter 8 of [9].

6.1 Basic definitions and properties

Definition 6.1 (Lie group)**.** *A **Lie group** is a group G with a compatible manifold structure, i.e., its group operation and inversion are smooth.*

Naturally, a **Lie group morphism** is a smooth group homomorphism between Lie groups and a **Lie subgroup** is simultaneously a subgroup and a submanifold of a Lie group. To every Lie group we associate the tangent space at its identity element as an algebra.

Definition 6.2. *The tangent space $\mathfrak{g} = T_1 G$ is called the **Lie algebra** of the Lie group G.*

Right now, it might not be clear why the tangent space should have the structure of an algebra, but this will be cleared up once we define a multiplication on it. We will define this operation on $\mathfrak{g}$ through a correspondence with vector fields on G, but it is also possible to do this more directly. We will not explore this approach, but curious readers are referred to [7] for a more in-depth study of Lie theory.

Example 6.3. *The following objects have a Lie group structure:*

(1) Euclidean space $(\mathbb{R}^n, +)$ with pointwise addition.

(2) The unit group $(\mathbb{R}^*, \cdot)$ of $\mathbb{R}$.

(3) The general linear group $\mathrm{GL}(n, K)$ (with K being $\mathbb{R}$ or $\mathbb{C}$); all of the typical matrix groups can be given a compatible smooth structure, and make for some of the most important Lie groups we will discuss. The following section is dedicated to these so-called classical groups.

(4) The unit circle S^1 as a subset of $\mathbb{C}$ with the usual multiplication.

The reason we are interested in Lie groups is that they allow us to define smooth actions on manifolds. For instance, say we want to study a physical system in which a rigid body undergoes a rotation. Then it would not do to just allow for rotations over a disparate collection of angles; we need to include rotations over *any* angle, hence the necessity of smoothness. With the presence of this property, we will of course exploit it to its full extent.

Definition 6.4 (Smooth action). *Let G be a Lie group and N a manifold. A group action of G on N is a* ***smooth action*** *if the evaluation map*

$$\begin{aligned} G \times N &\to N \\ (g, p) &\mapsto g \cdot p \end{aligned}$$

is smooth.

For now, we just want to look at actions of G on itself, particularly the **left action**, defined for $g \in G$ by the left multiplication

$$L_g : G \to G, h \mapsto gh.$$

This allows us to express the Lie algebra of G as the collection of **left-invariant vector fields**, that is, vector fields $X \in \mathcal{X}(G)$ for which

$$(\mathrm{d}L_g)_h X_h = X_{gh}$$

for all $g, h \in G$. This can be abbreviated as

$$(L_g)_* X = X.$$

Proposition 6.5. *The Lie algebra $\mathfrak{g} = T_1 G$ of a Lie group G can be identified with the vector space $\mathcal{X}(G)^G$ of left-invariant vector fields through*

$$\begin{aligned} \mathcal{X}(G)^G &\to \mathfrak{g} \\ X &\mapsto X_1. \end{aligned}$$

Proof. A left-invariant vector field is uniquely determined by its value at 1, since for any $g \in G$,

$$X_g = ((L_g)_* X)_g = (\mathrm{d}L_g)_1 X_1.$$

This makes the given mapping a linear isomorphism, with the remark that it is indeed linear. □

Remark 6.6. *From now on, we may also identify $\mathfrak{g}$ with $\mathcal{X}(G)^G$.*

Lemma 6.7. *All left-invariant vector fields on a Lie group are complete.*

Proof. See Theorem 9.18 of [9]. □

We now want to define a *Lie bracket* on our left-invariant vector field. This is essentially a map which behaves similarly to the way the commutator does on matrices. In fact, we can define such an operation for any manifold N; any vector field $X \in \mathcal{X}(N)$ defines a *derivation* $C^\infty(N) \to C^\infty(N), f \mapsto Xf$ by taking the derivative of a function along X. Note:

$$Xf = \mathcal{L}_X f.$$

On the space of derivations (defined in the theorem below), we can define a Lie bracket

$$[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X,$$

and we would like to pull this back to $\mathcal{X}(N)$.

Theorem 6.8. *Let N be a manifold and $\mathcal{D}(N)$ the vector space of derivations on N, those being the linear maps $\delta : C^\infty(N) \to C^\infty(N)$ which obey the Leibniz rule $\delta[fg] = f\delta g + g\delta f$. Then the mapping*

$$\mathcal{X}(N) \to \mathcal{D}(N)$$
$$X \mapsto \mathcal{L}_X$$

is a linear isomorphism.

Proof. The mapping is quite clearly linear, since $(\lambda X + Y)f = \lambda X f + Y f$ for $\lambda \in \mathbb{R}$ and $X, Y \in \mathcal{X}(N)$. We will show that the mapping is also invertible, by assigning to each derivation a vector field.

Given $\delta \in \mathcal{D}(N)$ and $p \in N$, recall that T_pN is the vector space of derivations at p. We construct a vector field $X \in \mathcal{X}(N)$ such that for all $p \in N$

$$X_p = \delta|_p,$$

where $\delta|_p : C^\infty(N) \to \mathbb{R}, f \mapsto (\delta f)(p)$ is the derivation at p defined by δ. Since δ is defined on smooth functions, $\delta|_p$ must be smooth. Hence X is well-defined as a smooth vector field on N.

Now we only need to show that $\delta = \mathcal{L}_X$. For any $f \in C^\infty(N)$, we have that

$$(\delta f)(p) = \delta|_p f = X_p f = (Xf)(p) = (\mathcal{L}_X f)(p),$$

so indeed $\delta = \mathcal{L}_X$, which concludes the proof. □

Corollary 6.9. *The Lie algebra $\mathfrak{g}$ of a Lie group G has a Lie bracket $[\cdot, \cdot]$, which sends left-invariant vector fields $X, Y \in \mathfrak{g}$ to the unique vector field $[X, Y] \in \mathfrak{g}$ with the property that*

$$\mathcal{L}_{[X,Y]} = \mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X.$$

Lie algebras are in fact defined more generally as (real) vector spaces $\mathfrak{g}$ endowed with a **Lie bracket** $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, which is defined as a bilinear, skew-symmetric operation which satisfies the **Jacobi identity**

$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$$

for all $X, Y, Z \in \mathfrak{g}$.

Now that we more or less know what the Lie algebra of a Lie group is, we can move towards a definition of the exponential map on general Lie groups. This not only reveals a justification for the terminology, but also lays down some important groundwork for the definition of symplectic reduction.

Definition 6.10. *A* ***one-parameter subgroup*** *of a Lie group G is a Lie group morphism $\phi : \mathbb{R} \to G$.*

It may seem a bit odd to give a mapping the name of subgroup, but the reasoning for this is that we want to view ϕ as the integral curve of some vector field through the identity element of G. So while the subgroup is really just the image of ϕ, we need the extra information that the derivative of ϕ gives to get the full picture. This gives rise to yet another identification with the Lie algebra.

Theorem 6.11. *There is a one-to-one correspondence between the collection of one-parameter subgroups of G and its Lie algebra $\mathfrak{g}$.*

Proof. To a one-parameter subgroup ϕ of G we can uniquely assign a left-invariant vector field $X^\phi \in \mathfrak{g}$ for which ϕ is an integral curve. We do need to show that this is well-defined, but once that is done, we find that X^ϕ is uniquely determined by its left-invariance through

$$X^\phi_g = (\mathrm{d}L_g)_e X^\phi_e = (\mathrm{d}L_g)_e \dot{\phi}(0),$$

where the identity element of G is written as e to avoid confusion with $1 \in \mathbb{R}$ (not an identity element!).

Let us now take a point $\phi(t)$ on ϕ for a fixed $t \in \mathbb{R}$. Then

$$X^\phi_{\phi(t)} = (\mathrm{d}L_{\phi(t)})_e \circ (\mathrm{d}\phi)_0(1) = \mathrm{d}[L_{\phi(t)} \circ \phi]_0(1),$$

where we used the definition $\dot{\phi}(x) = (\mathrm{d}\phi)_x(1)$. Now we need to use that ϕ is a homomorphism, so

$$L_{\phi(t)} \circ \phi = \phi \circ L_t,$$

where $L_t : \mathbb{R} \to \mathbb{R}$ is given by $x \mapsto x + t$. So

$$\begin{aligned} X^\phi_{\phi(t)} &= \mathrm{d}[\phi \circ L_t]_0(1) \\ &= (\mathrm{d}\phi)_t \circ (\mathrm{d}L_t)_0(1) \\ &= (\mathrm{d}\phi)_t \circ \dot{L}_t(0) \\ &= (\mathrm{d}\phi)_t(1) \\ &= \dot{\phi}(t). \end{aligned}$$

So indeed, there is a unique left-invariant vector field associated to each one-parameter subgroup. The other way around is almost trivial, since left-invariant vector fields are complete; their flows correspond to one-parameter subgroups. □

Definition 6.12 (Exponential map)**.** *Let ϕ_X be the unique one-parameter subgroup generated by $X \in \mathfrak{g}$. Then the* ***exponential map*** *on G is defined as*

$$\begin{aligned} \exp : \mathfrak{g} &\to G \\ X &\mapsto \phi_X(1). \end{aligned}$$

Now for any $t \in \mathbb{R}$, the integral curve of tX through $e \in G$ should be the same as that of X, except we move through it at a different rate: $\phi_{tX}(s) = \phi_X(ts)$. So we have

$$\exp tX = \phi_X(t).$$

Example 6.13. *On the unit group $\mathbb{R}^*$, let $E \in \mathcal{X}(\mathbb{R}^*)$ be the left-invariant vector field with $E_1 = 1$ (where we identify $T_1\mathbb{R}^* \cong \mathbb{R}$). Then*

$$E_x = (\mathrm{d}L_x)_1 E_1 = (\mathrm{d}L_x)_1(1) = \dot{L}_x(1) = x,$$

so

$$\dot{\phi}_E = \phi_E.$$

With $\phi_E(0) = 1$ we find $\phi_E(t) = e^t$. Any left-invariant vector field can be written as $X = xE$ for $x \in \mathbb{R}$ so

$$\exp tX = \phi_X(t) = \phi_E(tx) = e^{tx}.$$

This example illustrates one aspect of the relation between exponentials and exponential maps; it has everything to do with the differential equations that define them. In the following section, we show that a similar relation holds for matrix groups.

Example 6.14. *In this example we derive the form of the exponential map on the additive Lie group $\mathbb{R}$. A one-parameter subgroup of $\mathbb{R}$ is a smooth group homomorphism*

$$\phi : \mathbb{R} \to \mathbb{R}.$$

We note that for $n, m \in \mathbb{Z}$ with $n \neq 0$ we have $\phi(m/n) = m\phi(1/n)$ and $\phi(1) = \phi(n/n) = n\phi(1/n)$ due to additivity. Hence

$$\phi(m/n) = \frac{m}{n}\phi(1),$$

which defines ϕ on $\mathbb{Q}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, this extends to

$$\phi(x) = x\phi(1)$$

for any $x \in \mathbb{R}$. Hence ϕ is the multiplication by $\phi(1)$.

Now identify $T_0\mathbb{R} \cong \mathbb{R}$ so that the one-parameter subgroup associated to $X \in \mathbb{R}$ is the multiplication by X. Then

$$\exp tX = tX.$$

Example 6.15. *We compute the exponential map of S^1. Through the embedding into $\mathbb{C}$, we identify*

$$\mathfrak{g} = T_1 S^1 \cong \{1 + ix \mid x \in \mathbb{R}\} \cong i\mathbb{R}.$$

The left-invariant vector fields of S^1 can be viewed as the vector fields whose integral curve has constant speed. So the integral curves are of the form

$$\phi_{ix}(t) = e^{ixt}.$$

Thus,

$$\exp(ix) = e^{ix}.$$

6.2 Classical groups and computations

In physics, a handful of matrix groups crops up repeatedly. These are known as the **classical groups**. This term mainly refers to the following groups:

- $\mathrm{GL}(n, K) = \{A \in \mathrm{M}(n, K) \mid \det A \neq 0\}$ for $K = \mathbb{R}, \mathbb{C}$
- $\mathrm{SL}(n, K) = \{A \in \mathrm{GL}(n, K) \mid \det A = 1\}$
- $\mathrm{O}(n) = \{A \in \mathrm{GL}(n, \mathbb{R}) \mid A^{-1} = A^T\}$
- $\mathrm{SO}(n) = \{A \in \mathrm{O}(n) \mid \det A = 1\}$
- $\mathrm{U}(n) = \{A \in \mathrm{GL}(n, \mathbb{C}) \mid A^{-1} = A^*\}$, where $A^* = \overline{A}^T$ is the conjugate transpose of A
- $\mathrm{SU}(n) = \{A \in \mathrm{U}(n) \mid \det A = 1\}$

Here "G" stands for general, "L" for linear, "S" for special, "O" for orthogonal and "U" for unitary. In the case that $K = \mathbb{R}$, we may drop the field from the notation, so that for example $\mathrm{GL}(n) = \mathrm{GL}(n, \mathbb{R})$. All of the above are Lie groups, via their identification with a subset of $\mathrm{M}(n, K)$, which is diffeomorphic with K^{n^2}. The smoothness of the general linear groups follows quite easily from the smoothness of the determinant function; if we take a matrix in $\mathrm{GL}(n, K)$, slightly nudging matrix entries changes its determinant by a proportionally small amount, so the determinant should still be nonzero. Hence smoothness is implied. The smoothness of the other classical groups can be inferred by showing that they are submanifolds of a general linear group.

Example 6.16 (Lie algebra of $\mathrm{GL}(n, K)$)**.** *Let the matrix entries X_{ij} act as global coordinates on $\mathrm{GL}(n, K)$. Then any element A of the Lie algebra $\mathfrak{gl}(n, K) = T_{I_n}\mathrm{GL}(n, K)$ can be written as*

$$A = \sum_{i,j} A_{ij} \left.\frac{\partial}{\partial X_{ij}}\right|_{I_n} \in T_{I_n}\mathrm{GL}(n, K).$$

for $A_{ij} \in K$. This gives an identification of A with the matrix in $\mathrm{M}(n, K)$ with entries A_{ij}. It turns out that $\mathfrak{gl}(n, K)$ is in fact $\mathrm{M}(n, K)$ with the commutator $[\cdot, \cdot]$. For a computation, see the proof of Proposition 8.41 from [9, p. 193–195].

As above, the Lie algebra of a classical group is traditionally denoted in lowercase Fraktur letters.

Example 6.17. *This example uses the method of Example 8.47 of [9, p. 197].*

We are going to determine the Lie algebras of $\mathrm{O}(n)$ and $\mathrm{U}(n)$, although our approach is applicable to a wider collection of Lie groups. Let $\phi : \mathrm{M}(n, K) \to \mathrm{M}(n, K)$ be a self-inverse smooth map which leaves $\mathrm{GL}(n, K)$ invariant so that we have a Lie group

$$G := \{A \in \mathrm{GL}(n, K) \mid A^{-1} = \phi(A)\}.$$

and define a smooth map

$$\begin{aligned}\Phi : \mathrm{GL}(n,K) &\to \mathrm{GL}(n,K)\\ A &\mapsto A\phi(A).\end{aligned}$$

We now find G as the level set $\Phi^{-1}(I_n)$ and consequentially its Lie algebra as

$$\mathfrak{g} = T_{I_n}G = \ker(\mathrm{d}\Phi)_{I_n}.$$

We compute the differential of Φ explicitly. For $A \in \mathfrak{gl}(n,K)$, consider a path

$$\begin{aligned}\gamma_A : (-\varepsilon,\varepsilon) &\to \mathrm{GL}(n,K)\\ t &\mapsto I_n + tA\end{aligned}$$

for ε sufficiently small. Then

$$\begin{aligned}(\mathrm{d}\Phi)_{I_n}(A) &= (\mathrm{d}\Phi)_{I_n}(\dot\gamma_A(0))\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\big[\Phi\circ\gamma_A(t)\big]\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\big[(I_n+tA)\phi(I_n+tA)\big]\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\big[(I_n+tA)(I_n+t\phi(A))\big]\\ &= A+\phi(A).\end{aligned}$$

So this gives

$$\mathfrak{g} = \{A \in \mathfrak{gl}(n,K) \mid A = -\phi(A)\},$$

and it follows that $\mathfrak{o}(n)$ and $\mathfrak{u}(n)$ are the algebras of skew-symmetric and skew-Hermitian matrices, respectively.

In the example above, we computed Lie algebras via the identification with the tangent space at the identity. In the next example, we use the identification with the collection of one-parameter subgroups. To this end, we first prove the following proposition regarding exponential maps on matrix groups.

Proposition 6.18. *Consider a Lie subgroup $G \subseteq \mathrm{GL}(n,K)$ with Lie algebra $\mathfrak{g} \subseteq \mathfrak{gl}(n,K)$, and let $X \in \mathfrak{g}$ be a left-invariant vector field. Then its exponential map takes the form* $\exp tX = e^{tX}$.

Proof. Let L_A be the left multiplication on G by $A \in G$, and consider a path $\gamma_B(t) = I_n + tB$ for $B \in \mathfrak{g}$ as in the previous example. Then

$$(\mathrm{d}L_A)_{I_n}(B) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\big[L_A\circ\gamma_B(t)\big] = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\big[A(I_n+tB)\big] = AB.$$

Now let $E \in \mathfrak{g}$ be the left-invariant vector field with $E_{I_n} = I_n$ so that

$$E_A = (\mathrm{d}L_A)_{I_n} E_{I_n} = A.$$

It follows that the one-parameter subgroup generated by E will have to satisfy $\dot{\phi}_E = E_{\phi_E} = \phi_E$, so

$$\phi_E(t) = e^{tI_n}.$$

Now the vector field generated by $X \in T_{I_n}G$ takes the form (with a slight abuse of notation)

$$X_A = (\mathrm{d}L_A)_{I_n} X = AX,$$

so X generates the vector field EX. Therefore, its exponential map becomes

$$\exp tEX = \phi_{EX}(t) = e^{tX}.$$

□

Example 6.19. *Let $G \subseteq \mathrm{GL}(n, K)$ be a Lie subgroup. We will compute the Lie algebra $\mathfrak{sg}$ of*

$$\mathrm{S}G := \{A \in G \mid \det A = 1\}.$$

From Proposition 6.18, it follows that all one-parameter subgroups of $\mathrm{S}G$ take the form $t \mapsto e^{tX}$ for $X \in \mathfrak{sg}$. The determinant of any element in $\mathrm{S}G$ must be 1, so using the identity

$$\det e^A = e^{\operatorname{Tr} A}$$

we find that the condition that $\det e^{tX} = 1$ for all $t \in \mathbb{R}$ is equivalent to

$$\operatorname{Tr} X = 0.$$

Hence all matrices in $\mathfrak{g}$ with vanishing trace are contained in $\mathfrak{sg}$. The other way around, if $X \in \mathfrak{g}$ has $\operatorname{Tr} X = 0$, then $\exp tX \in \mathrm{S}G$. Theorem 2.29 of [7, p. 17] has the consequence that the restriction of $\exp$ to some neighbourhood of $0 \in \mathfrak{gl}(n, K)$ is a diffeomorphism. So for t sufficiently small, we can invert $\exp tX$ to obtain tX as an element of $\mathfrak{sg}$. Subsequently, all of its scalar multiples remain in the Lie algebra, so we may conclude that $X \in \mathfrak{sg}$. In summary,

$$\mathfrak{sg} = \{X \in \mathfrak{g} \mid \operatorname{Tr} X = 0\}.$$

With that, we can write down the Lie algebras of all the classical groups mentioned:

- $\mathfrak{gl}(n, K) = \mathrm{M}(n, K)$
- $\mathfrak{sl}(n, K) = \{A \in \mathfrak{gl}(n, K) \mid \operatorname{Tr} A = 0\}$
- $\mathfrak{o}(n) = \{A \in \mathfrak{gl}(n) \mid A^T = -A\}$
- $\mathfrak{so}(n) = \{A \in \mathfrak{o}(n) \mid \operatorname{Tr} A = 0\}$
- $\mathfrak{u}(n) = \{A \in \mathfrak{gl}(n, \mathbb{C}) \mid A^* = -A\}$
- $\mathfrak{su}(n) = \{A \in \mathfrak{u}(n) \mid \operatorname{Tr} A = 0\}$

Along with $\mathbb{R}^n$, these tend to be the only Lie algebras we come across in our studies, so it is helpful to get familiar with them.

6.3 Adjoint and coadjoint representations

Let G be a Lie group with Lie algebra $\mathfrak{g}$. Note that G acts on itself smoothly by conjugation:

$$\begin{aligned}\psi_g : G &\to G\\ h &\mapsto ghg^{-1}.\end{aligned}$$

The evaluation $(g, h) \mapsto ghg^{-1}$ is indeed quite clearly smooth since the group operation and inversion of G are smooth. This conjugation then gives rise to a representation

$$\begin{aligned}\mathrm{Ad} : G &\to \mathrm{GL}(\mathfrak{g})\\ g &\mapsto \mathrm{Ad}_g := (\mathrm{d}\psi_g)_e\end{aligned}$$

known as the **adjoint representation** of G on $\mathfrak{g}$, where we identify $\mathfrak{g}$ with T_eG. This representation will be used to generate vector fields. There is also a similar notion of a **coadjoint representation** of G on $\mathfrak{g}^*$ denoted by $\mathrm{Ad}^* : G \to \mathrm{GL}(\mathfrak{g}^*)$. This is defined using the natural pairing

$$\begin{aligned}\langle\cdot,\cdot\rangle : \mathfrak{g}^* \times \mathfrak{g} &\to \mathbb{R}\\ (\xi, X) &\mapsto \xi(X).\end{aligned}$$

Then Ad^*_g for $g \in G$ is defined uniquely by the relation

$$\langle \mathrm{Ad}^*_g \xi, X\rangle = \langle \xi, \mathrm{Ad}_{g^{-1}} X\rangle.$$

The vector field on $\mathfrak{g}$ **generated by** $X \in \mathfrak{g}$ **for the adjoint representation** is defined at $Y \in \mathfrak{g}$ by

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Ad}_{\exp tX} Y.$$

On matrix groups $G \subseteq \mathrm{G}L(n, K)$ we would find

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Ad}_{\exp tX} Y &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp tX \cdot Y \cdot \exp -tX\\ &= X \exp tX \cdot Y \cdot \exp -tX + \exp tX \cdot Y \cdot \exp -tX \cdot -X\big|_{t=0}\\ &= XY - YX\\ &= [X, Y].\end{aligned}$$

It turns out that the same relation holds on any Lie group (although the justification is a bit more complicated). Completely analogous to the vector field generated for the adjoint representation, any $X \in \mathfrak{g}$ **generates a vector field for the coadjoint representation** denoted by $X^{\#}$. For $\xi \in \mathfrak{g}^*$ and $Y \in \mathfrak{g}$ we compute (using the bilinearity of $\langle \cdot, \cdot \rangle$)

$$\begin{aligned} \langle X^{\#}_{\xi}, Y \rangle &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \langle \mathrm{Ad}^*_{\exp tX} \xi, Y \rangle \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \langle \xi, \mathrm{Ad}_{\exp -tX} Y \rangle \\ &= \langle \xi, -[X, Y] \rangle \\ &= \langle \xi, [Y, X] \rangle. \end{aligned}$$

The adjoint and coadjoint actions are a rather technical aspect of the theory, but still tie back into physics. The orbits of the coadjoint action can be endowed with a natural symplectic structure, and have a deep connection with quantum physics. This goes far beyond the scope of this thesis, but the reader is referred to [2] for a discussion of the topic.

7 Reduction

7.1 Actions and momentum maps

Let (M, ω) be a symplectic manifold and $\psi : \mathbb{R} \to \mathrm{Diff}(M)$ a smooth action. We can view ψ as the flow of a unique complete vector field $X = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \psi_t$ on M. In fact, this describes a one-to-one correspondence between complete vector fields on M and smooth $\mathbb{R}$-actions on M.

Now what if we consider the flow of a symplectic vector field as an action of $\mathbb{R}$ on M? Recall, a symplectic vector field $X \in \mathcal{X}(M)$ preserves ω, so its flow must do the same. It then follows that $\mathbb{R}$ *acts by symplectomorphism* on M; for every $t \in \mathbb{R}$ we have $(\exp tX)^*\omega = \omega$. The notion of symplectic actions in fact generalises to the actions of arbitrary Lie groups:

Definition 7.1 (Symplectic action)**.** *Let (M, ω) be a symplectic manifold and G a Lie group. A **symplectic action** of G on M is a smooth action $\psi : G \to \mathrm{Diff}(M)$ such that $\psi_g^*\omega = \omega$ or all $g \in G$.*[4]

So now we have a one-to-one correspondence between symplectic actions of $\mathbb{R}$ on M and symplectic vector fields on M. So naturally the question arises, is there a similar notion of *Hamiltonian actions*? For $\mathbb{R}$-actions, the answer is completely analogous to the preceding situations, where an action ψ is Hamiltonian if it is the flow of a Hamiltonian vector field. However, when we try to generalise this to an arbitrary Lie group, we run into trouble; this definition of a Hamiltonian action is entirely dependent on the vector field it generates, a luxury unavailable for more complicated Lie groups.

To define Hamiltonian actions, we associate a so-called *momentum map* to a symplectic action, which will have associated to it Hamiltonians whose Hamiltonian vector fields are constructed using the action.

[4]Equivalently, we may define a symplectic action as a smooth action whose image is contained in $\mathrm{Sympl}(M, \omega)$, the group of symplectomorphisms of (M, ω).

Definition 7.2 (Hamiltonian action). *Let (M, ω) be a symplectic manifold, G a Lie group with Lie algebra $\mathfrak{g}$ and*

$$\psi : G \to \mathrm{Sympl}(M, \omega)$$

*a symplectic action. We call ψ a **Hamiltonian action** if there exists an associated **momentum map**, that is, a mapping*

$$\mu : M \to \mathfrak{g}^*$$

satisfying

(1) For $X \in \mathfrak{g}$, define

- $\mu^X : M \to \mathbb{R}, p \mapsto \langle \mu(p), X \rangle$
- $X^\# \in \mathcal{X}(M)$ *given by* $X^\#_p := \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \psi_{\exp tX}(p)$

Then $\mathrm{d}\mu^X = X^\# \lrcorner\, \omega$

(2) μ is equivariant[5] *with respect to ψ and* Ad^*:[6]

$$\mu \circ \psi_g = \mathrm{Ad}^*_g \circ \mu$$

for all $g \in G$.

*If the above holds, then we call (M, ω, G, μ) a **Hamiltonian G-space**.*

In this definition, we call μ^X the **X-component** of μ. Before we proceed with more complicated examples of Hamiltonian actions, it is instructive to study the definition in the case of $\mathbb{R}$-actions.

Example 7.3. *In this example we show that Hamiltonian actions of $\mathbb{R}$ behave appropriately with respect to Hamiltonian vector fields. Let*

$$\psi : \mathbb{R} \to \mathrm{Sympl}(M, \omega)$$

be a symplectic action which generates a Hamiltonian vector field

$$X_p = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \psi_t(p).$$

Let $\mu : M \to \mathbb{R}$ be a Hamiltonian of this vector field, and identify $\mathbb{R}$ with its Lie coalgebra. We show that this makes μ a momentum map for ψ.

[5] A function $f : X \to Y$ between G-spaces is said to be *equivariant* with respect to the G-actions if $f(g \cdot x) = g \cdot f(x)$ for all $g \in G$ and $x \in X$.

[6] If G is abelian, then Ad and Ad^* become trivial. Hence the equivariance condition simplifies to invariance: $\mu \circ \psi_g = \mu$.

For $v \in \mathbb{R}$, we find that $\mu^v(p) = \langle \mu(p), v \rangle = v\mu(p)$, so

$$\mathrm{d}\mu^v = v\mathrm{d}\mu = vX \lrcorner\, \omega.$$

So we need $v^\# = vX$. This checks out:

$$\begin{aligned} v_p^\# &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \psi_{\exp tv}(p) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \psi_{tv}(p) \\ &= v\,\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \psi_s(p) \\ &= vX_p. \end{aligned}$$

Here we used the result that $\exp tv = tv$ from Example 6.14 and changed variables from t to $s = tv$. So indeed

$$\mathrm{d}\mu^v = v^\# \lrcorner\, \omega.$$

As for the equivariance condition, since $\mathbb{R}$ is abelian, the coadjoint action becomes trivial so we require μ to be invariant:

$$\mu \circ \psi_t = \mu$$

for all $t \in \mathbb{R}$. Recall that a Hamiltonian is constant along its Hamiltonian vector field:

$$\mathcal{L}_X \mu = X \lrcorner\, \mathrm{d}\mu + \mathrm{d}[X \lrcorner\, \mu] = 0.$$

But we can also write this Lie derivative as

$$\mathcal{L}_X \mu = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mu \circ \psi_t,$$

since ψ_t is the flow along X. This derivative is zero, so $\mu \circ \psi_t$ is equal for all t, in particular for $t = 0$:

$$\mu \circ \psi_t = \mu \circ \psi_0 = \mu.$$

Therefore, our definition of a Hamiltonian action is indeed a generalisation of a more natural concept of Hamiltonian $\mathbb{R}$-actions.

The following example gives a momentum map for a translation action on $T^*\mathbb{R}^3$. This can be thought of as a symmetry arising from, for instance, a constant electric field pointing along the z-axis in $\mathbb{R}^3$.

Example 7.4. *Consider the cotangent bundle $T^*\mathbb{R}^3$ with standard cotangent coordinates $(x, \xi) = (x_1, x_2, x_3, \xi_1, \xi_2, \xi_3)$ and define an action of $\mathbb{R}^2$ by*

$$a \cdot (x, \xi) := (x + a, \xi) \text{ for } a \in \mathbb{R}^2 \subseteq \mathbb{R}^3.$$

Here we identify $\mathbb{R}^2$ with the (x_1, x_2)-plane in $\mathbb{R}^3$. Let us construct a momentum map for the action.

Since the exponential map on $\mathbb{R}^2$ is the identity map (cf. Example 6.14), so $X = (X_1, X_2) \in T_0\mathbb{R}^2 \cong \mathbb{R}^2$ generates a vector field

$$X^{\#}_{(x,\xi)} = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} tX \cdot (x, \xi) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (x + tX, \xi) = X_1 \frac{\partial}{\partial x_1}\bigg|_{(x,\xi)} + X_2 \frac{\partial}{\partial x_2}\bigg|_{(x,\xi)}.$$

This gives a contraction with the canonical symplectic form

$$X^{\#} \lrcorner\, \omega = X_1 \mathrm{d}\xi_1 + X_2 \mathrm{d}\xi_2,$$

so we can validly define a momentum map by

$$\mu^X(x, \xi) = X_1\xi_1 + X_2\xi_2 = X \cdot \xi$$

(where X is embedded into $\mathbb{R}^3$ as $(X_1, X_2, 0)$). This is clearly invariant under the $\mathbb{R}^2$-action, and so satisfies the equivariance condition.

In the example above, we found a momentum map on a cotangent bundle defined by X-components of the form $X^{\#} \lrcorner\, \alpha$, where α is the tautological 1-form. We can actually always lift smooth actions on a manifold to Hamiltonian actions on the cotangent bundle, as shown by the lemma below. This result is exceedingly useful in classical mechanics; once we formulate the procedure of symplectic reduction, this can be used to reduce a phase space by symmetries of the configuration space.

Lemma 7.5 (Cotangent lift of smooth action)**.** *Let N be a manifold with a smooth action of G. Then the G-action can be lifted to a Hamiltonian G-action on T^*N given by*

$$g \cdot (x, \xi) = (g \cdot x, ((\mathrm{d}g)_x^{-1})^* \xi)$$

for $(x, \xi) \in T^*N$ and $g \in G$, with momentum map

$$\begin{aligned} \mu : T^*N &\to \mathfrak{g}^* \\ p &\mapsto (X \mapsto \mu^X(p)), \end{aligned}$$

where $\mu^X = X^{\#} \lrcorner\, \alpha$ for the tautological 1-form α on T^*N.

Before we prove this lemma, let us dissect the expression of $g \cdot \xi$. For a start, $g \in G$ can be seen as a diffeomorphism

$$g : N \to N$$

with differential

$$(\mathrm{d}g)_x : T_xN \to T_{gx}N$$

at $x \in N$. Since $\xi \in T_x^*N$ is a function on T_xN, we obtain

$$\xi \circ ((\mathrm{d}g)_x)^{-1} \in T_{gx}^*N$$

as desired.

Proof of Lemma 7.5. First note that, using Cartan's magic formula, for any $X \in \mathfrak{g}$ we find

$$\mathrm{d}\mu^X = \mathcal{L}_{X^\#}\alpha - X^\# \lrcorner\, \mathrm{d}\alpha = \mathcal{L}_{X^\#}\alpha + X^\# \lrcorner\, \omega,$$

where $\omega = -\mathrm{d}\alpha$ is the canonical symplectic form on T^*N. So we need $\mathcal{L}_{X^\#}\alpha$ to be 0. It will suffice to show that α is G-invariant, since $X^\#$ is generated by the action of G. In order to do so, note that the projection $\pi : T^*N \to N$ is G-equivariant:

$$(g^{-1} \circ \pi \circ g)(p) = (g^{-1} \circ \pi)(g \cdot p) = g^{-1} \cdot (gx) = x = \pi(p),$$

where $p = (x, \xi) \in T^*N$. Note that in this computation, g is in turn viewed as a diffeomorphism of N and one of T^*N. Now we use the fact that $\alpha_p = \xi \circ (\mathrm{d}\pi)_p$ to compute

$$\begin{aligned}
(g^*\alpha)_p &= \alpha_{g\cdot p} \circ (\mathrm{d}g)_p \\
&= (g \cdot \xi) \circ (\mathrm{d}\pi)_{g\cdot p} \circ (\mathrm{d}g)_p \\
&= ((\mathrm{d}g)_x^{-1})^*\xi \circ (\mathrm{d}\pi)_{g\cdot p} \circ (\mathrm{d}g)_p \\
&= \xi \circ (\mathrm{d}g)_x^{-1} \circ (\mathrm{d}\pi)_{g\cdot p} \circ (\mathrm{d}g)_p \\
&= \xi \circ \mathrm{d}[g^{-1} \circ \pi \circ g]_p \\
&= \xi \circ (\mathrm{d}\pi)_p \\
&= \alpha_p,
\end{aligned}$$

so

$$g^*\alpha = \alpha.$$

Hence, $\mathcal{L}_{X^\#}\alpha = 0$ and

$$\mathrm{d}\mu^X = X^\# \lrcorner\, \omega.$$

Now we need to show that μ satisfies the equivariance property of momentum maps. This is a bit more involved, but it more or less comes down to combining three pieces of information. The first one is a direct consequence of the G-equivariance of π:

$$(\mathrm{d}\pi)_{g\cdot p} \circ (\mathrm{d}g)_p = (\mathrm{d}g)_x \circ (\mathrm{d}\pi)_p. \tag{$*$}$$

The second result we need is an alternative way to write the X-component of μ:

$$\langle \mu(p), X \rangle = \alpha_p(X_p^\#) = \langle \alpha_p, X_p^\# \rangle. \tag{$**$}$$

The final piece is a bit more complicated:

$$\begin{aligned}
(\mathrm{Ad}_g X)_p^\# &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp(t\mathrm{Ad}_g X) \cdot p \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (g \cdot \exp tX \cdot g^{-1}) \cdot p \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g \cdot (\exp tX \cdot (g^{-1} \cdot p)) \\
&= (\mathrm{d}g)_{g^{-1}\cdot p}(X^\#_{g^{-1}\cdot p}).
\end{aligned} \tag{$***$}$$

This all culminates in the following computation:

$$\begin{aligned}
\langle \mu(g\cdot p), X\rangle &= \langle (g\cdot\xi)\circ(\mathrm{d}\pi)_{g\cdot p}, X^{\#}_{g\cdot p}\rangle \quad \text{by } (**)\\
&= \langle g\cdot\xi, (\mathrm{d}\pi)_{g\cdot p}(X^{\#}_{g\cdot p})\rangle\\
&= \langle \xi\circ(\mathrm{d}g)_x^{-1}, (\mathrm{d}\pi)_{g\cdot p}(X^{\#}_{g\cdot p})\rangle\\
&= \langle \xi, (\mathrm{d}g)_x^{-1}\circ(\mathrm{d}\pi)_{g\cdot p}(X^{\#}_{g\cdot p})\rangle\\
&= \langle \xi, (\mathrm{d}\pi)_p\circ(\mathrm{d}g)_p^{-1}(X^{\#}_{g\cdot p})\rangle \quad \text{by } (*)\\
&= \langle \xi, (\mathrm{d}\pi)_p((\mathrm{Ad}_{g^{-1}}X)^{\#}_p)\rangle \quad \text{by } (***)\\
&= \langle \xi\circ(\mathrm{d}\pi)_p, (\mathrm{Ad}_{g^{-1}}X)^{\#}_p\rangle\\
&= \langle \alpha_p, (\mathrm{Ad}_{g^{-1}}X)^{\#}_p\rangle\\
&= \langle \mu(p), \mathrm{Ad}_{g^{-1}}X\rangle \quad \text{by } (**)\\
&= \langle \mathrm{Ad}^*_g\mu(p), X\rangle.
\end{aligned}$$

So it follows that

$$\mu\circ g = \mathrm{Ad}^*_g\circ\mu,$$

which concludes the proof. □

7.2 The momentum map for an N-particle system

This section treats Example 1.15 of [5].

In this section we construct a Hamiltonian structure on the phase space $(T^*\mathbb{R}^3)^N$ of an N-particle system in three dimensions. This will reveal a relation between the momentum map and the physical notion of momentum. The action we will be looking at is one from the group $G = \mathbb{R}^3 \rtimes \mathrm{O}(3)$ with Lie algebra $\mathfrak{g} = \mathbb{R}^3 \rtimes \mathfrak{o}(3)$. On $\mathbb{R}^3$, this group acts as the translations and rotations. Multiplication in G is defined by

$$(v, A)\cdot(w, B) = (v + Aw, AB)$$

and it acts on $\mathbb{R}^3$ by

$$(v, A)\cdot x = Ax + v.$$

We lift this action to $T^*\mathbb{R}^3$ (with the canonical symplectic form ω_0) through Lemma 7.5. Call the momentum map defined in the lemma $\nu : T^*\mathbb{R}^3 \to \mathfrak{g}^*$.

Now let $q = (q_1, q_2, q_3)$ be the standard coordinates on $\mathbb{R}^3$, and denote the corresponding cotangent coordinates as (q, p). Note that $\mathfrak{o}(3)$ can be identified with $\mathbb{R}^3$ equipped with the cross product through the mapping

$$\begin{pmatrix} 0 & -z & y\\ z & 0 & -x\\ -y & x & 0\end{pmatrix} \mapsto \begin{pmatrix} x\\ y\\ z\end{pmatrix}$$

and hence we can identify

$$\mathfrak{g} \cong \mathbb{R}^3 \times \mathbb{R}^3 \cong \mathfrak{g}^*.$$

So taking $X = (X_1, X_2) \in \mathfrak{g}$ we find

$$X_q = X_1 + X_2 \times q$$

and so

$$\begin{aligned} \nu^X(q,p) &= (X^\# \lrcorner \alpha)(q,p) \\ &= p \cdot X_q \\ &= p \cdot (X_1 + X_2 \times q) \\ &= p \cdot X_1 + (q \times p) \cdot X_2 \\ &= (p, q \times p) \cdot X. \end{aligned}$$

Now it follows that the momentum map can be expressed as

$$\nu(q,p) = (p, q \times p).$$

This already reveals some of the implied physics of the momentum map, but we first want to extend this to an N-particle system. This is just a matter of taking the Cartesian product of N copies of $T^*\mathbb{R}^3$. Write the coordinates of the jth copy as (q^j, p^j) and let $\pi_j : (T^*\mathbb{R}^3)^N \to T^*\mathbb{R}^3$ be the projection onto this component. Writing the canonical symplectic form on $T^*\mathbb{R}^3$ as ω_0, we can define the symplectic form on $(T^*\mathbb{R}^3)^N$ as

$$\omega = \sum_j \pi_j^* \omega_0.$$

If we have G act diagonally on $(T^*\mathbb{R}^3)^N$, the coresponding momentum map becomes

$$\mu = \sum_j \nu \circ \pi_j,$$

which, explicitly, becomes

$$\mu(q,p) = \sum_j (p^j, q^j \times p^j).$$

Now we see that the "linear" component of μ is the total linear momentum while its "angular" component is the total angular momentum of the system. Hence the momentum map becomes a sort of generalisation of linear and angular momentum related to the action of G.

7.3 The Marsden-Weinstein and Noether Theorems

So far, we have discussed the actions with which we can apply symplectic reduction, but we have not yet touched upon the actual reduction. That loose end is tied up in this section, where we formulate two major theorems of symplectic reduction. The first one, the Marsden-Weinstein Theorem, guarantees that we have a well-defined quotient manifold. The second theorem is Noether's Theorem, a famous result widely quoted in physics. Here we give a symplectic-geometric formulation of the theorem.

Before we state the first theorem, we need the following definition.

Definition 7.6. *Let $\pi : E \to B$ be a smooth fibre bundle and G a Lie group. We call E with the bundle projection a* ***principal G-bundle*** *if the following conditions are met:*[7]

(1) G preserves the fibres of π

(2) G acts freely, transitively and faithfully on the fibres

Note that with this definition, we can identify B with the orbit space E/G. One might think that any manifold equipped with a smooth action can be turned into a principal bundle, but this is not necessarily true. Just take a look at the following example, where the quotient of a manifold over a smooth action is not even Hausdorff.

Example 7.7. *This example is due to [9, p. 542]. We consider a smooth $\mathbb{R}$-action on the torus $\mathbb{T}^2$, defined by*

$$x \cdot (z_1, z_2) = (e^{2\pi i x} z_1, e^{2\pi i x \alpha} z_2),$$

where we take $\alpha \in \mathbb{R}$ to be irrational. Now consider the point $(1, 1) \in \mathbb{T}^2$, and let $n, m \in \mathbb{Z}$. Then

$$(n + m/\alpha) \cdot (1, 1) = (e^{2\pi i (n+m/\alpha)}, e^{2\pi i (n\alpha+m)}) = (e^{2\pi i \frac{m}{\alpha}}, e^{2\pi i n\alpha}).$$

Since the orbits of irrational rotations form dense subsets of S^1, the orbit of $(1, 1)$ must be dense in $\mathbb{T}^2$. So any open neighbourhood U of $(1, 1)$ corresponds to a subset $\tilde{U} \subseteq \mathbb{T}^2/\mathbb{R}$ which is the quotient of a dense subset in $\mathbb{T}^2$. Since closed subsets of $\mathbb{T}^2/\mathbb{R}$ are quotients of closed subsets of $\mathbb{T}^2$, it follows that $\tilde{U}$ inherits the denseness of U. In other words, it must *overlap with any open subset of $\mathbb{T}^2/\mathbb{R}$, so this quotient cannot be Hausdorff.*

So we need some additional restrictions on our G-space before we can say that its quotient is a manifold. A sufficient condition on the action is called *properness*, which roughly means that it behaves like an action of a compact Lie group. In our statement of the Marsden-Weinstein Theorem, we will set the stronger condition that G is compact, but the reader is encouraged to consult Chapter 21 of [9] for a discussion of proper actions.

Let us now state the first theorem. The formulation is taken from [3].

Theorem 7.8 (Marsden-Weinstein). *Let (M, ω, G, μ) be a Hamiltonian G-space, where G is compact*[8] *and acts freely on $\mu^{-1}(0)$. Then*

(1) The orbit space $M_{\text{red}} = \mu^{-1}(0)/G$ is a smooth manifold

[7] Equivalently, we may require that the fibres of the bundle are G, with the G-action on E lifted from the left multiplication on the fibres.

[8] As mentioned before, the theorem also holds if G acts properly and is not necessarily compact.

(2) $\pi : \mu^{-1}(0) \to M_{\text{red}}$ *is a principal G-bundle*

(3) *There is a symplectic form* ω_{red} *on* M_{red} *such that* $\iota^*\omega = \pi^*\omega_{\text{red}}$*, where* $\iota : \mu^{-1}(0) \hookrightarrow M$ *is the inclusion*

Here, we may also write $M//G$ instead of M_{red}. This quotient is known as the **reduced space** or **Marsden-Weinstein quotient**. A proof of the theorem is beyond the scope of our discussion but can be found in Chapter 23 of [3] and Section 4.3 of [1].

One might think that the compactness condition is quite a heavy restriction here; after all, in the previous section we looked at an action of $\mathbb{R}^3 \rtimes \mathrm{O}(3)$, which is decidedly not compact. However, cases like this can often be engineered to fit the framework of symplectic reduction anyways. In this particular example, the quotient $\mathbb{R}^{3N}/\mathbb{R}^3$ is just $\mathbb{R}^{3N-3}$ (think of the reduction as centering the coordinate system on the first particle), which is still a manifold and behaves as we want with respect to the Marsden-Weinstein Theorem. It will be a recurring theme that we can slightly relax the condition of the theorem to fit our needs.

In addition to the Marsden-Weinstein Theorem, we need an additional result which brings a Hamiltonian system on M down to $M//G$. This result is Noether's Theorem:

Theorem 7.9 (Noether)**.** *Let* (M, ω, G, μ) *be a Hamiltonian G-space and* $H \in C^\infty(M)$ *a smooth map with Hamiltonian vector field* X_H*. Then*

$$H \text{ is } G\text{-invariant} \iff \mu \text{ is constant on integral curves of } X_H.$$

Proof. Let $X \in \mathfrak{g}$ be a vector field, then we can compute

$$\begin{aligned} \mathcal{L}_{X_H}\mu^X &= X_H \lrcorner \, \mathrm{d}\mu^X \\ &= X_H \lrcorner \, (X^\# \lrcorner \, \omega) \\ &= -X^\# \lrcorner \, (X_H \lrcorner \, \omega) \\ &= -X^\# \lrcorner \, \mathrm{d}H \\ &= -\mathcal{L}_{X^\#}H. \end{aligned}$$

Now note that

$$\begin{aligned} H \text{ is } G\text{-invariant} &\iff \forall X \in \mathfrak{g} : \mathcal{L}_{X^\#}H = 0 \\ &\iff \forall X \in \mathfrak{g} : \mathcal{L}_{X_H}\mu^X = 0 \\ &\iff \mu \text{ is constant on integral curves of } X_H. \end{aligned}$$

This proves the theorem. □

This theorem provides an easy way to verify the G-invariance of a Hamiltonian. If a Hamiltonian is G-invariant, then it descends down to the symplectic quotient $M//G$ as a reduced Hamiltonian, as shown by the following proposition.

Proposition 7.10. *Let (M, ω, G, μ) be a Hamiltonian G-space satisfying the conditions of the Marsden-Weinstein Theorem. Suppose $H \in C^\infty(M)$ is G-invariant. Then there is a unique reduced Hamiltonian $H_{\text{red}} \in C^\infty(M_{\text{red}})$ such that*

$$\pi^* H_{\text{red}} = \iota^* H.$$

Proof. The mapping $\iota^* H$ is simply the restriction $H|_{\mu^{-1}(0)} : \mu^{-1}(0) \to \mathbb{R}$. Since H is G-invariant, the existence and unicity of H_{red} are ensured by a universal property of quotients:

$$\begin{array}{ccc} \mu^{-1}(0) & \xrightarrow{\iota^* H} & \mathbb{R} \\ & {}_{\pi}\searrow \quad \nearrow_{\exists! H_{\text{red}}} & \\ & \mu^{-1}(0)/G & \end{array}$$

□

7.4 An example of reduction

Consider the punctured complex space $M = \mathbb{C}^n \backslash \{0\}$, a $2n$-dimensional manifold. It has a symplectic form

$$\omega = \sum_{j=1}^{n} r_j \mathrm{d}r_j \wedge \mathrm{d}\theta_j,$$

expressed in polar coordinates. Consider the S^1-action on M by

$$t \cdot (z_1, \ldots, z_n) = (t \cdot z_1, \ldots, t \cdot z_n),$$

where we take S^1 as the unit circle in $\mathbb{C}$. We claim that this action is Hamiltonian with momentum map

$$\begin{aligned} \mu : M &\to \mathbb{R} \\ z &\mapsto -\frac{1}{2}(\|z\|^2 - 1). \end{aligned}$$

To verify this claim, note that the vector field induced by $X \in \mathbb{R}$ is given by

$$X^\# = X \sum_j \frac{\partial}{\partial \theta_j},$$

and

$$X^\# \lrcorner\, \omega = -X \sum_j r_j \mathrm{d}r_j = \mathrm{d}\left[-\frac{1}{2} X \sum_j r_j^2 \right] = \mathrm{d}\mu^X.$$

Furthermore, we find that μ is quite clearly S^1-invariant. Hence it follows that μ is indeed a momentum map for the action.

Now we have $\mu^{-1}(0) = S^{2n-1}$. Since S^1 acts freely, we can perform Marsden-Weinstein reduction. Note that our action resembles scalar multiplication by $\mathbb{C}^*$ in the following manner:

$$S^{2n-1}/S^1 \cong (\mathbb{C}^n \backslash \{0\})/\mathbb{C}^*.$$

It follows that the symplectic quotient is a projective space:

$$M//S^1 \cong \mathbb{CP}^{n-1}.$$

Incidentally, this also shows that complex projective space has a symplectic structure.

7.5 Singular reduction

In physical applications, Marsden-Weinstein reduction often does not suffice due to the requirement that the action be free. We need to also allow almost-free group actions with isolated fixed points. This extension of Marsden-Weinstein reduction is known as **singular reduction**. Let us consider an example of how this works in practice. We treat the problem of the two-dimensional harmonic oscillator. This is the physical system consisting of a single particle in $\mathbb{R}^2$ in the potential

$$V(x) = \frac{1}{2}k\|x\|^2,$$

with a non-negative constant k. We begin by reducing the phase space by a rotational symmetry.

We consider the standard action of SO(2) on $\mathbb{R}^2$, and lift it to $T^*\mathbb{R}^2$ via Lemma 7.5. We denote the usual cotangent coordinates as (x_1, x_2, ξ_1, ξ_2). If we identify $T^*\mathbb{R}^2 \cong \mathbb{R}^2 \times \mathbb{R}^2$, our action is the same on both factors $\mathbb{R}^2$ of this product. Though $\mathrm{SO}(2) \cong S^1$ is compact, the requirement of the Marsden-Weinstein Theorem that the action be free, is violated; the point $x_0 = (0,0) \in \mathbb{R}^2 \times \mathbb{R}^2$ is fixed under the action. However, SO(2) does act freely on $T^*\mathbb{R}^2\backslash\{x_0\}$, so we can reduce this submanifold and deal with the singularity in x_0 later.

Let us first find the level set $\mu^{-1}(0)$ of the momentum map. Note that the components $\mu^X = X^\# \lrcorner\, \alpha$ for $X \in \mathfrak{so}(2)$ take the form

$$\mu^X(x,\xi) = \langle \xi, \tilde{X}_x \rangle,$$

where

$$\tilde{X}_x := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp tX \cdot x.$$

Then

$$\mu^{-1}(0) = \{(x,\xi) \in T^*\mathbb{R}^2 \mid \forall X \in \mathfrak{so}(2) : \langle \xi, \tilde{X}_x \rangle = 0\}.$$

Note that $\tilde{X}_x$ depends linearly on X, and so describes a linear subspace of $T_x\mathbb{R}^2$. Hence ξ must be on the (notably one-dimensional) orthogonal complement[9] of this subspace. This is quite a familiar structure; it is (almost) the structure of a vector bundle. Only at $x = (0,0) \in \mathbb{R}^2$ do

[9] Here we identify $T\mathbb{R}^2 \cong T^*\mathbb{R}^2$.

we find that ξ can take any value. Still, if we remove $0 \in T^*_{(0,0)}\mathbb{R}^2$ from the cotangent fibre, it reduces down to $\mathbb{R}\backslash\{0\}$ under the SO(2)-action. In short, we find

$$(\mu^{-1}(0)\backslash\{x_0\})/\mathrm{SO}(2) \cong T^*\mathbb{R}\backslash\{x_0\},$$

if we identify x_0 with the origin of $T^*\mathbb{R}$.

Now let us see what happens if we introduce x_0 back into the cotangent bundle. The problem of reducing $\mathbb{R}^2$ by rotations has been studied before in [10], where the authors show that the quotient is not a manifold, but something more general: an *orbifold*. Orbifolds are defined very similarly to manifolds, but instead of being modeled after $\mathbb{R}^n$, they are modeled after group quotients of $\mathbb{R}^n$. We repeat the essence of the approach given in this paper.

We choose a particularly convenient subset of $\mu^{-1}(0)$, given by

$$\Lambda = \{(x, 0, \xi, 0) \in T^*\mathbb{R}^2\} \cong T^*\mathbb{R}.$$

We remark two things; firstly, the SO(2)-orbit of Λ is all of $\mu^{-1}(0)$. Secondly, by having $\mathbb{Z}_2$ act on Λ by $(x, \xi) \mapsto (-x, -\xi)$, we can effectively view $\mathbb{Z}_2$ as a subgroup of SO(2). As a consequence of these two facts, we can say

$$T^*\mathbb{R}^2//\mathrm{SO}(2) \cong \Lambda/\mathbb{Z}_2 \cong T^*\mathbb{R}/\mathbb{Z}_2.$$

So our symplectic quotient can be viewed as the cotangent bundle $T^*(0, \infty)$ with half of the cotangent space $T^*_0\mathbb{R}$ glued on. We have neglected some minor details here, such as the smoothness of this identification (and what that means exactly). For a rigorous discussion of these peculiarities, see the paper [10].

Now let us define a Hamiltonian on $T^*\mathbb{R}^2$. It will take the form

$$H(x_1, x_2, \xi_1, \xi_2) = \frac{1}{2m}(\xi_1^2 + \xi_2^2) + \frac{1}{2}k(x_1^2 + x_2^2).$$

This Hamiltonian is clearly invariant under the SO(2)-action, and descends to the symplectic quotient. To find an expression of the reduced Hamiltonian, define a projection

$$\begin{aligned} \pi : \mu^{-1}(0) &\to T^*\mathbb{R}/\mathbb{Z}_2 \\ (x_1, x_2, \xi_1, \xi_2) &\mapsto \left[\left(\sqrt{x_1^2 + x_2^2}, \mathrm{sgn}(\xi_1)\sqrt{\xi_1^2 + \xi_2^2}\right)\right], \end{aligned}$$

where we set $\mathrm{sgn}(0) := 1$ and the square brackets represent the equivalence class of the object inside. Write the induced coordinates on $T^*\mathbb{R}/\mathbb{Z}_2$ as (x, ξ) (cf. the definition of Λ), then it follows that

$$H_{\mathrm{red}}(x, \xi) = \frac{\xi^2}{2m} + \frac{1}{2}kx^2,$$

which is the Hamiltonian of the one-dimensional harmonic oscillator.

Part III
Applications

8 Symmetry breaking

Symmetry breaking is a phenomenon in modern physics where a system exhibiting a particular symmetry, collapses to a system with a lesser degree of symmetry. This takes many different forms, both simple and more advanced. A basic example is the introduction of a uniform force field to the system of a free particle in a spherically symmetric potential. Before the uniform field is introduced, the Hamiltonian takes the form

$$H(q,p) = \frac{p^2}{2m} + V(q),$$

with p being the momentum of the particle, r its distance from the origin, and V the spherical potential. This Hamiltonian has an SO(3)-symmetry, but note what happens when we introduce the additional potential; we can write the modified Hamiltonian as

$$H'(q,p) = \frac{p^2}{2m} + V(q) + \langle q, d\rangle,$$

where $d \in \mathbb{R}^3$ is a constant vector pointing along the uniform force field. H' is now no longer SO(3)-symmetric, but instead is only invariant under a subgroup of symmetries. Specifically, H' has an SO(2)-symmetry given by the rotations around d. This is known as *explicit symmetry breaking*.

There are a few other forms of symmetry breaking, but it always comes down to some Lie group "breaking down" to a subgroup because the system transitions to a less symmetric state. This is a subject rooted deeply in quantum field theory, so we will not delve further into the physics.[10] Instead, we spend the remainder of this section applying symplectic reduction to two Lie groups of particular relevance to field theory. This reduction is carried out in an attempt to reduce Lie groups to particular subgroups.

8.1 Reduction of SU(2)

Here we base our problem on the symmetry breaking from SU(2) to U(1). Ideally, we want to symplectically reduce $T^*\mathrm{SU}(2)$ to $T^*\mathrm{U}(1)$. Unfortunately, we have not found a way to perform this reduction exactly. However, we present a calculation that *almost* gives the desired result.

To begin, we note that SU(2) is diffeomorphic to S^3. It can be checked (by computing inverses and determinants) that

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix} \middle| z, w \in \mathbb{C},\ |z|^2 + |w|^2 = 1 \right\}.$$

[10]The reader is referred to Chapter 44 of [8] for an introductory overview of symmetry breaking.

This is clearly diffeomorphic to the unit sphere in $\mathbb{C}^2$, which is S^3.

Now consider S^3 as the unit sphere in $\mathbb{R}^4$. Through the embedding

$$\begin{aligned} \mathrm{SO}(3) &\hookrightarrow \mathrm{SO}(4) \\ A &\mapsto \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}, \end{aligned}$$

we can define an action of SO(3) on S^3 descended from the natural SO(4)-action. This action consists of the rotations of S^3 that keep the last coordinate invariant.

We now lift this action to T^*S^3. Since S^3 is diffeomorphic to a Lie group, it is parallelisable: $T^*S^3 = S^3 \times \mathbb{R}^3$. Given cotangent coordinates (x, ξ), the action lifts to the diagonal action

$$A \cdot (x, \xi) = (A \cdot x, A \cdot \xi).$$

Here SO(3) acts on $\mathbb{R}^3$ by the usual matrix multiplication. The lift comes with a momentum map μ with components

$$\mu^X(x, \xi) = \langle \xi, \tilde{X}_x \rangle$$

for $X \in \mathfrak{so}(3)$, where we define

$$\tilde{X}_x := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp tX \cdot x.$$

We found earlier that $\mathfrak{so}(3) = \mathfrak{o}(3)$ is isomorphic to $\mathbb{R}^3$ with the cross product. For $x \in S^3 \subseteq \mathbb{R}^4$ we introduce the notation $x = (x', x'')$ with $x' \in \mathbb{R}^3$ and $x'' \in \mathbb{R}$ representing the components of x. Now we can write

$$\mu^X(x, \xi) = \langle \xi, X \times x' \rangle = \langle X, x' \times \xi \rangle$$

since SO(3) only acts on the first three components of x, and $\mathfrak{so}(3)$ acts on $\mathbb{R}^3$ by the cross product. Now it follows that

$$\mu^{-1}(0) = \{(x, \xi) \in S^3 \mid x' \times \xi = 0\}.$$

At the poles, $N = (0, 0, 0, 1)$ and $S = (0, 0, 0, -1)$, we find $(N, 0)$ and $(S, 0)$ as fixed points of the action, while it is free everywhere else. Hence we may reduce

$$Z = \mu^{-1}(0) \backslash \{(N, 0), (S, 0)\}$$

to the Marsden-Weinstein quotient

$$Z/\mathrm{SO}(3) = T^*[-1, 1] \backslash \{(\pm 1, 0)\}.$$

Now set

$$\Lambda = \{((0, 0, a, b), (0, 0, \xi)) \in \mu^{-1}(0) \mid a, b, \xi \in \mathbb{R}\} \cong T^*S^1$$

and have $\mathbb{Z}_2$ act on it by

$$(a, b, \xi) \mapsto (-a, b, -\xi).$$

Λ contains at least one element of each orbit of the SO(3)-action on $\mu^{-1}(0)$, and $\mathbb{Z}_2$ acts on it as a subgroup of SO(3) (generated by a 180°-rotation). Hence, we find

$$T^*S^3//\mathrm{SO}(3) \cong \Lambda/\mathbb{Z}_2 \cong T^*S^1/\mathbb{Z}_2.$$

This can be viewed as the manifold $T^*[-1, 1]$ with the outer two cotangent fibres replaced by half-spaces.

Evidently, we did not manage to reduce $T^*\mathrm{SU}(2)$ to $T^*\mathrm{U}(1)$; the resulting orbifold differs from it in the boundaries. Furthermore, it is not the cotangent space of a Lie group and therefore cannot be viewed as a collection of symmetries.

8.2 Reduction of SO(3)

Though our method fails for SU(2), it can be applied to SO(3) for a more satisfactory result. In this case, the question is based on the symmetry breaking from SO(3) to SO(2). In Exercises 2.11 through 2.14 of [7] it is shown that there is a Lie group morphism $\mathrm{SU}(2) \to \mathrm{SO}(3)$ which is also a 2-sheeted covering. Notably, SO(3) is diffeomorphic to $\mathbb{RP}^3$ and that this covering corresponds to the usual quotient map $S^3 \to \mathbb{RP}^3$.

We remark that the action defined in the previous section descends to $\mathbb{RP}^3$. Denote the antipodal equivalence on S^3 by $\sim$. Then for $x, y \in S^3$ we have

$$x \sim y \iff \|x - y\| = 2.$$

Since SO(3) acts on S^3 by isometry, it must respect $\sim$. Hence the action descends to $S^3/\sim \cong \mathbb{RP}^3$. It can be shown that the momentum map $\mu : T^*S^3 \to \mathfrak{so}(3)^*$ also descends to a unique momentum map $\tilde{\mu} : T^*\mathbb{RP}^3 \to \mathfrak{so}(3)^*$ of the induced action (this is true more generally in the presence of a quotient mapping which is equivariant with respect to the group action).

Denote the induced projection of cotangent bundles as $\pi : T^*S^3 \to T^*\mathbb{RP}^3$ so that

$$\mu = \tilde{\mu} \circ \pi.$$

Then the zero level of $\tilde{\mu}$ is

$$\tilde{\mu}^{-1}(0) = \pi[\mu^{-1}(0)]$$

and it is spanned by the orbits of the elements of

$$\tilde{\Lambda} = \{([0 : 0 : a : b], (0, 0, \xi)) \in \tilde{\mu}^{-1}(0)\} \cong T^*\mathbb{RP}^1.$$

$\mathbb{Z}_2$ acts on $\tilde{\Lambda}$ by $([a : b], \xi) \mapsto ([-a : b], -\xi)$, and by the same arguments as before we find

$$T^*\mathbb{RP}^3//\mathrm{SO}(3) \cong \tilde{\Lambda}/\mathbb{Z}_2.$$

At a glance this seems to be the same quotient we reduced $T^*\mathrm{SU}(2)$ to, but there is a subtle difference; it is a quotient over antipodal equivalence, so it can be viewed as $T^*\mathbb{RP}^1$ with the fibre at one point replaced by a half-space. In any case, the base space *is* in fact $\mathbb{RP}^1 \cong S^1 \cong \mathrm{SO}(2)$.

9 Closing Remarks

In this report, we have gone to considerable depths in the mathematical discussion of Hamiltonian mechanics. We began by constructing a mathematical formulation of the Hamilton formalism, including a few illustrative examples along the way. We then discussed Lie theory in order to set up the technique of symplectic reduction. We only applied the totality of the contents to two problems, but the potential of this theory ranges far wider than we have been able to discuss within the constraints of this work.

We stumbled upon one avenue for further development of the theory in Example 5.5, where we found that our formulation of Hamiltonian mechanics fails in cases where special relativity is needed. One could take a look at relativistic extensions of Hamiltonian mechanics like the one described in Chapter 7 of [4], and work out a symplectic-geometric formulation of the theory accordingly. To go even further, one could also write down a symplectic theory of general relativity by working on Riemannian manifolds. Similarly, one could pursue a symplectic description of (relativistic) quantum mechanics.

Another type of problem which we cannot tackle yet but that came up during the research, were "continuous" physical systems. That is to say, systems which contain a continuum of moving parts, like a string or a surface. It turns out that these systems can perhaps still be treated with symplectic geometry, but require infinite-dimensional symplectic manifolds. One such problem caught our interest during this research, which is discussed in Appendix A. This is *not* a rigorous discussion, but lays out a possible strategy for tackling the problem using this extended theory.

Finally, perhaps the most obvious point of improvement is that of singular reduction. In the last few sections, we noted that this is often necessary in practical applications, though we have not worked out the details.

To conclude, what we have discussed here is only a glimpse of what symplectic geometry is capable of with regards to physics. There is still much more to be seen, like mathematical formulations of modern physics and extensions to a wider range of problems. In any case, it forms a fruitful and challenging pursuit to rigorously describe the mathematical fundaments on which modern physics is built.

Acknowledgements

I want to thank both my supervisors, Federica and Subodh, for guiding me through this research. It plainly wouldn't have turned out the way it did without Federica's critical questions about my reasoning and Subodh's insightful explanations of physical concepts related to my work. In addition, I want to thank my amazing partner, Hannah, for her moral support during the writing of this thesis.

Part IV
Appendix

A Elastic surface deformations

One problem we came across during this research involves an elastic surface. The problem is as follows: consider a fixed circular rim laid down horizontally. From the rim an elastic surface is hung. Under a uniform downward pointing gravitational field, what shape will the surface take on?

This problem has been solved for flexible but inelastic surfaces (see [6, p. 177]), but there does not seem to be a general solution for elastic surfaces. Since the problem has a rotational symmetry, it might be possible to tackle the problem using symplectic geometry. Since we have not covered the mathematics required to solve this problem rigorously, this section serves as nothing more than a roadmap for a possible solution. To that end, the problem is defined and a possible strategy for solving it is proposed. To begin, we need a definition of deformations of an object in $\mathbb{R}^3$, given as follows.

Definition A.1. *Let U be a smooth 2-manifold embedded into $\mathbb{R}^3$. A* ***deformation*** *of U is a smooth mapping $F : [0,1] \times U \to \mathbb{R}^3$ together with a subset of its topological closure $K \subseteq \overline{U}$ such that*

(1) $F_t := F(t, \cdot) : U \to \mathbb{R}^3$ is a diffeomorphism onto its image for all $t \in [0,1]$

(2) $F_0 = \mathrm{id}_U$

(3) If F_t is smoothly extended to $\overline{U}$, we have $F_t|_K = \mathrm{id}_K$ for all $t \in [0,1]$

(4) There is a $t \in [0,1]$ such that F_t is not an isometry, unless all F_t are id_U

We call a deformation ***trivial*** *if $F_t = \mathrm{id}_U$ for all $t \in [0,1]$.*

This gives a deformation as essentially a smooth homotopy with some additional constraints. Particularly, the last condition ensures that it truly does deform the surface if it is not trivial.

To be able to employ a Hamiltonian approach, we need to define a configuration space for our problem. Then the question becomes what path through the configuration space is the correct one. Given a surface U, we define the **deformation space** of U as the collection of all deformations of U with fixed points K:

$$\mathcal{D}_K(U) := \{F : [0,1] \times U \to \mathbb{R}^3 \mid (F, K) \text{ is a deformation of } U\}$$

For brevity, we denote $F_1[U]$ by $F(U)$ and $F_1(x)$ by $F(x)$ for $x \in U$.

Now we run into the issue of rigour. $\mathcal{D}_K(U)$ is our configuration space, but it is not immediately clear that it is even a manifold. That is because it is unlike any manifold we have encountered so far; if it is one, it is infinite-dimensional! Unfortunately, we do not have the machinery to deal with these objects, so we proceed with an approach from physics. We leave the nature of the deformation space as a conjecture.

Conjecture A.2. *$\mathcal{D}_K(U)$ is a smooth submanifold of $C^\infty([0,1] \times U, \mathbb{R}^3)$.*

In physics, there is theory which can be applied to problems like this with an infinite number of degrees of freedom. This is known as (classical) field theory, and more information on the subject can be found in Chapter 12 of [4]. This discipline deals with the problem by using densities instead of the usual quantities like the Hamiltonian, force and so forth. Such a density is chosen such that taking the integral results in the original quantity.

To get a sense of the physics involved in this problem, let us determine the force density. With the notation as above, U will be the open unit disk in the $z = 0$ plane of $\mathbb{R}^3$, and we let K be the bounding circle of the disk. The force density of a given deformation F of U is a smooth mapping $f^F : U \to T\mathbb{R}^3$. Here U acts as a sort of index set for the deformed surfaces. The force density has two components, namely the gravitational force density and the elastic force density:

$$f^F = f_g^F + f_e^F.$$

The gravitational force density is given by

$$f_g^F = \rho^F g,$$

where $g \in \mathcal{X}(\mathbb{R}^3)$ is a constant downward pointing vector field representing the gravitational acceleration and $\rho^F : U \to \mathbb{R}$ is the mass density of $F(U)$, indexed by U. This can be computed from the initial density by noting that the density is inversely proportional to the surface area. Taking a limit, we find

$$\rho_x^F = \lim_{\varepsilon \to 0} \rho_x \frac{\text{area}(B_\varepsilon(x))}{\text{area}(F(B_\varepsilon(x)))},$$

where ρ_x is the initial density and $B_\varepsilon(x)$ the disk in U of radius ε around x.

Now we want to define the elastic force density. We need to approximate this force as being conservative, since otherwise the Hamiltonian formalism does not work. We employ an idea of tension.

Given a point $x \in U$, a unit vector $v \in T_{F(x)}F(U)$ and a sufficiently small $a > 0$ such that $x + av$ lies within U, let L_x^{av} be the line segment in U between x and $x + av$. Then we define the **tension** at x in the direction v as

$$t_v(x) := \lim_{a \to 0} Yv \frac{\text{length}(F(L_x^{av})) - \text{length}(L_x^{av})}{a} = \lim_{a \to 0} Yv \left(\frac{\text{length}(F(L_x^{av}))}{a} - 1 \right),$$

where $Y \in \mathbb{R}$ is a fixed elasticity constant, and the length is defined by the Euclidean metric on $\mathbb{R}^3$. So the tension measures the elongation at a point on the deformed surface in a certain direction. Note the resemblance to Hooke's law. The total elastic force density is obtained by integrating the tension over all directions v:

$$f_e(x) = \oint_{\substack{v \in T_{F(x)}F(U) \\ \|v\|=1}} t_v(x)\,\mathrm{d}v.$$

If we denote the unit circle in $T_{F(x)}F(U)$ by C_x^F, then the total force density can be written as

$$f_x = \rho_x^F g_x + \oint_{C_x^F} t_v \,\mathrm{d}v.$$

To summarise, we defined the following quantities:

- mass density ρ^F of deformation F
- gravitational acceleration g
- tension t_v in direction v

Before we move on, we need to prove that this force density is indeed conservative. We know that the gravitational part is, but this is not so clear for the elastic part. So if we fix $x, y \in U$, then for any path γ in U from x to y, the integral

$$I_\gamma = \int_\gamma \oint_{C^F_{\gamma(s)}} t_v(\gamma(s))\,\mathrm{d}v \cdot \dot{\gamma}(s)\,\mathrm{d}s$$

should evaluate to the same value.

Conjecture A.3. *In the above, the integral I_γ depends only on the endpoints of γ.*

Using the force density, we can now compute the potential density of the system. We integrate the force over an arbitrary path from a fixed point x_0 on the boundary K of U to any point $x \in U$. Let γ_x denote one of those paths, so the potential density becomes

$$V_x = \rho_x^F \|g\|\, h_x + \int_{\gamma_x} \oint_{C^F_{\gamma_x(s)}} t_v(\gamma_x(s))\,\mathrm{d}v \cdot \dot{\gamma}_x(s)\,\mathrm{d}s,$$

in which h_x is the z-coordinate of x. To write down the Hamiltonian density we also need a kinetic energy. To that end, we need an additional momentum coordinate $\pi \in \mathbb{R}^3$, so that the Hamiltonian density becomes a mapping $U \times \mathbb{R}^3 \to \mathbb{R}$ given by

$$\mathcal{H}^F(x,\pi) = \frac{\|\pi\|^2}{2\rho_x^F} + \rho_x^F \|g\|\, h_x + \int_{\gamma_x} \oint_{C^F_{\gamma_x(s)}} t_v(\gamma_x(s))\,\mathrm{d}v \cdot \dot{\gamma}_x(s)\,\mathrm{d}s.$$

We can now integrate this to find a Hamiltonian of the system. Let $\Pi : U \to \mathbb{R}^3$ be a vector field, then we can define this integrated Hamiltonian as

$$H(F, \Pi) = \int_U \mathcal{H}^F(x, \Pi_x)\, \mathrm{d}x.$$

In a way, H can be seen as a smooth mapping on the "cotangent bundle" of $\mathcal{D}_K(U)$, if this is indeed a manifold.

This concludes the definition of the problem in mathematical terms. Now the question is to find stable critical points of H. This may be accomplished through symplectically reducing the phase space under the obvious S^1-symmetry. We do need to make the relatively weak assumption that the initial density ρ_x is invariant under this symmetry. In addition, one might want to show that the symmetry of the initial configuration is preserved when it is acted upon by the uniform gravitation, although it is quite intuitive that this must be true. Once this is done, one could symplectically reduce the phase space. This should give something resembling the deformation space of a one-dimensional string.

References

[1] Ralph Abraham and Jerrold E. Marsden. *Foundations of Mechanics.* Addison-Wesley Publishing Company, 1978.

[2] A. Y. Alekseev and A. Z. Malkin. Symplectic geometry of the chern-simons theory. In *Integrable Models and Strings*, pages 59–97, Berlin, Heidelberg, 1994. Springer Berlin Heidelberg.

[3] Ana Cannas da Silva. *Lectures on Symplectic Geometry*, volume 1764. Springer Berlin Heidelberg, 2008.

[4] Herbert Goldstein. *Classical Mechanics.* Addison-Wesley, second edition, January 1980.

[5] Gert Heckman and Peter Hochs. Geometry of the momentum map, June 2012.

[6] Stefan Hildebrandt and Hermann Karcher. *Geometric analysis and nonlinear partial differential equations.* Springer, 2003.

[7] Alexander Kirillov. *An Introduction to Lie Groups and Lie Algebras.* Cambridge University Press, July 2008.

[8] Eleanor Knox and Alastair Wilson. *The Routledge companion to philosophy of physics.* Routledge, 2022.

[9] John M. Lee. *Introduction to Smooth Manifolds*, volume 218. Springer New York, 2012.

[10] Eugene Lerman, Richard Montgomery, and Reyer Sjamaar. Examples of singular reduction. *London Mathematical Society Lecture Note Series*, 192:127–155, 1993.