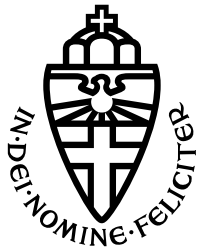
Radboud University Nijmegen



Faculty of Science

Symplectic geometry

Darboux's theorem and Hamiltonian systems

Bachelor thesis Mathematics

Author:
Matthijs Lau

Supervisor:
dr. Ioan Mărcuț

Second reader:
prof. Klaas Landsman

April 5, 2023

Abstract

In this thesis, we will discuss the local structure of symplectic manifolds in the form of Darboux's theorem, which tells us all symplectic forms are locally trivial. We introduce all the linear algebra and differential geometry needed for the proof. Furthermore, we will discuss classical mechanics both from a physics and a mathematical perspective. We will see that a symplectic manifold combined with a smooth function lets us build a mathematical model of classical systems. We will show how this mathematical formalisation can give results on the dynamics of classical systems in terms of local and global behaviour.

Contents

1 Introduction **2**

2 Linear Symplectic Geometry **4**
2.1 Basics and Definitions . . . 4
2.2 Symplectomorphisms . . . 6
2.3 Standard Form . . . 7

3 Symplectic Geometry on Manifolds **9**
3.1 Basics and Definitions . . . 9
3.2 Symplectomorphisms . . . 11
3.3 Normal Coordinates . . . 12
3.3.1 Operators on Differential Forms . . . 13
3.3.2 Darboux's Theorem . . . 25

4 Classical Mechanics as a Physicist **28**
4.1 Newtonian Formalism . . . 28
4.1.1 Space, Time and Kinematics . . . 28
4.1.2 Newton's Laws of Motion . . . 30
4.2 Lagrangian Formalism . . . 33
4.2.1 Energy . . . 33
4.2.2 Euler-Lagrange Equations . . . 36
4.3 Hamiltonian Formalism . . . 39
4.3.1 Cartesian Hamiltonian Mechanics . . . 39
4.3.2 The Hamiltonian in Generalised Coordinates . . . 40

5 Mathematical Methods of Classical Mechanics **43**
5.1 Hamiltonian Vector Fields . . . 43
5.1.1 Riemannian Gradients . . . 43
5.1.2 Symplectic Analogue . . . 44
5.1.3 Bracket Algebras . . . 46
5.2 Hamiltonian Systems . . . 49
5.3 Conserved Quantities . . . 51
5.3.1 Symmetries . . . 52
5.3.2 Integrable Systems . . . 54

A Vector Fields **61**
A.1 Vector fields . . . 61
A.1.1 Algebraic Structures . . . 62
A.1.2 Integral Curves and Flow . . . 65
A.2 Time-dependent Vector Fields . . . 70
A.2.1 Integral Curves and Flow . . . 70

Chapter 1

Introduction

Any avid student of both mathematics and physics has come in contact with a manifold at some point. The apt description of a space like the earth, a sphere which looks flat for us standing on it, immediately paints the usefulness of an abstract object like a manifold in the description of the real physical world. It is in this marriage of mathematics and physics that symplectic geometry is born and where it was first found as a manifold with a geometric structure much like a Riemannian manifold. Lagrange is known as the first who came upon these structures in his research on constants of motion in 1808. It was first known as complex groups due to their association with line complexes, but the inverter of this name, H. Weyl, noted that this name seemed to imply a connection to complex numbers. He proposed to use the Greek calque: symplectic, which is what we use nowadays. Even then, symplectic geometry only played a role in the research of mathematical physics and it was not until V.I. Arnol'd posed some interesting conjectures surrounding symplectic geometry in 1960 that it became a separate field in mathematics.

In this thesis, we will focus our attention on the local structure of symplectic manifolds and Hamiltonian systems. Therefore, it is of utmost importance that the reader is familiar with the concepts surrounding smooth manifolds. There is a heavy focus on tensor fields and differential forms, hence, one should at least be comfortable with their construction and the definition of the exterior derivative. Furthermore, as many different authors have their own notations for vector fields and flows, we have included an appendix discussing vector fields, flows and time-dependent vector fields where all the notation is also defined. Good sources on all the prerequisites are Chapters 1, 2, 3, 4, 8, 9, 12 and 14 [13] or the first 13 lectures of [18] excluding Lecture 11. These sources are also used for most proofs in manifold theory. As these sources do not contain adequate chapters on symplectic geometry, we will follow [4] for symplectic geometry and Hamiltonian systems, and [1] for Hamiltonian systems. We will not assume any prior knowledge on the topic of symplectic geometry, be it on vector spaces or manifolds, and all necessary definitions and references are included in this thesis. At some points, we do assume the reader is familiar with Riemannian geometry. However, these parts are never mandatory to understand the results of symplectic geometry as they serve the purpose of adding context to our results. Lastly, we work out a couple of examples of both mechanical and Hamiltonian systems, including some figures of mechanical motions made with TikZ and Python.

Motivation

As mentioned in the introduction, any student of both the mathematics track and the physics track comes in contact with manifolds at some point, as did I. In the previous year, I followed a triplet of courses introducing the concepts of smooth manifolds, Riemannian geometry and mathematical physics. It was in this last course that I was introduced to symplectic geometry as a method in mathematical physics as a geometric method to classical mechanics. Meanwhile, in the course on Riemannian geometry, we were studying a close relative to symplectic geometry. Together this piqued my interest in geometry, more specifically symplectic geometry. Together with my supervisor, Ioan Mărcuț, I proposed to investigate the basic geometry of a symplectic

manifold and then showcase how this geometry comes together with physics. This marriage of physics and mathematics is what inspired me to choose this topic.

Besides this emotional note of motivation for this thesis, it of course also has a lot of mathematical merits. The flatness of all symplectic manifolds paints a stark contrast when put next to its not-so-flat cousin Riemannian geometry. Furthermore, the study of Hamiltonian systems, a simple variation on the structure of symplectic manifolds, let us find many useful results about classical mechanics using a geometric point of view. Furthermore, this thesis gives an introduction to the topic of symplectic geometry and can be used as a starting point for further study.

Conventions

All notations and conventions on manifold theory are taken from [13], but as we already mentioned any vector field-related quantities are defined in Appendix A. We will follow the conventions and notation of [4] for symplectic geometry as it forms the main source on this topic. Furthermore, we assume that every manifold and vector field in this thesis is smooth, we usually only mention the smoothness in the context of functions.

Acknowledgements

I would like to thank my supervisor, Ioan Mărcuț, for all the help and patience even when on sabbatical. Thank you for all the explanations and the interesting working environment, allowing me to find a topic which interested me. Furthermore, I would like to thank Klaas Landsman for being my second reader. Secondly, I want to thank Rozemarijn for her feedback but more importantly her support and ever-lasting willingness to be a listening ear. Furthermore, many thanks to my friends and roommates for their joining me in my writing sessions. Lastly, I want to thank my family for always motivating me and giving me the confidence to go on.

Chapter 2

Linear Symplectic Geometry

Let us start with a chapter on the study of symplectic geometry on vector spaces, also called linear symplectic geometry. The motivation for such a structure may not be apparent at first, but it forms the foundation of symplectic geometry in general, which finds its application in classical mechanics, see Chapter 5. In this chapter, we will introduce the basic definitions surrounding linear symplectic geometry, i.e. symplectic vector spaces and linear symplectomorphisms. Along the way, we will also give some noteworthy examples. We will then show that the study of linear symplectic geometry comes down to the study of a simple model. In other words, there are no special invariants in this theory. In this chapter, we will follow the structure of [4, Section 1.2].

2.1 Basics and Definitions

As mentioned, we will start with the basic building block of linear symplectic geometry, namely symplectic vector spaces. Akin to any other geometric space, like inner product spaces and normed spaces, this is a pair consisting of a vector space with a structural function, which is called a linear symplectic form. In other words, a skew-symmetric and non-degenerate bilinear form.

Definition 2.1. Let V be an n-dimensional real vector space over $\mathbb{R}$ and let $\Omega : V \times V \to \mathbb{R}$ be a bilinear map. The map Ω is called a **linear symplectic form** if the following conditions hold:

- **Skew-symmetric:** For all $u, v \in V$ it holds that $\Omega(u, v) = -\Omega(v, u)$.
- **Non-degenerate:** For all non-zero $u \in V$ there exists a $v \in V$ such that $\Omega(u, v) \neq 0$.

A pair of a vector space V and linear symplectic form $\Omega : V \times V \to \mathbb{R}$, denoted by (V, Ω), is called a **symplectic vector space**. //

Notice the similarities between an inner product, symmetric and positive-definitive bilinear form, and a linear symplectic form, skew-symmetric and non-degenerate bilinear form. We see that we simply switched the symmetry and imposed a different degeneracy condition. Due to this resemblance, we will often compare symplectic vector spaces with inner product spaces. The definitions of inner product spaces and symplectic vector spaces are not that far apart. Both introduce a structural function that is a bilinear mapping with some symmetry constraint. Furthermore, the non-degeneracy of the linear symplectic forms and the positive definiteness of the inner products induce a mapping from the vector space to the dual space by $\widehat{S} : V \to V^* : v \mapsto S(v, \cdot)$, where S is either a symplectic form or an inner product. As the inner product and symplectic form are non-degenerate, we can deduce that the induced mapping is invertible. Before we proceed with the geometrical aspects, we will go over some examples. The first of these will be the main object of this chapter.

Example 2.2. Let us define the **trivial symplectic vector space of dimension** $2n$, denoted by $\left(\mathbb{R}^{2n}, \Omega_{\text{st},2n}\right)$. For this we will use the identification of $\mathbb{R}^{2n}$ as $\mathbb{R}^n \oplus \mathbb{R}^n$, such that a vector

$u \in \mathbb{R}^{2n}$ can be written as $u = (u_1, u_2) \in \mathbb{R}^n \oplus \mathbb{R}^n$. If we then denote the standard inner product on $\mathbb{R}^n$ as $\langle \cdot, \cdot \rangle_n$, we can define the mapping $\Omega_{\text{st},2n} : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}$ as

$$\Omega_{\text{st},2n}\left((u_1, u_2), (v_1, v_2)\right) = \langle u_1, v_2 \rangle_n - \langle u_2, v_1 \rangle_n .$$

It should be clear from the bilinearity of the inner product that this mapping is bilinear as well and the symmetry of the inner product directly implies the skew-symmetry of $\Omega_{\text{st},2n}$. Furthermore, from the positive definiteness of the inner product, the bilinear form inherits its non-degeneracy. Therefore, $\Omega_{\text{st},2n}$ is a linear symplectic form, called the **linear trivial symplectic form**, and $\left(\mathbb{R}^{2n}, \Omega_{\text{st},2n}\right)$ is a symplectic vector space called the **trivial vector space**. //

Example 2.3. Let us define another symplectic vector space, $(V \oplus V^*, \Omega_{\text{can}})$, where V is some finite dimensional real vector space. Define the mapping $\Omega_{\text{can}} : (V \oplus V^*) \times (V \oplus V^*) \to \mathbb{R}$ as

$$\Omega_{\text{can}}\left((u, \phi), (v, \psi)\right) = \psi(u) - \phi(v) .$$

It should be clear that this is a bilinear mapping as the elements of V^* are linear maps. Furthermore, it should be clear from the definition that this is a skew-symmetric map. The non-degeneracy can then be shown in a basis $\beta = \{u_1, \ldots, u_n\}$ of V, which gives a dual basis $\beta^* = \{\xi_1, \ldots, \xi_n\}$ of V^*, such that $\xi_i(u_j) = \delta_{ij}$. We can then define an isomorphism $\iota : V \to V^*$ by its action on β as $\iota(u_i) = \xi_i$. Now suppose that $(u, \phi) \in V \times V^*$ is non-zero, in other words $u \neq 0$ or $\phi \neq 0$. Consider the case where $u \neq 0$, such that $\iota(u)$ is also non-zero. We can then consider the action of the bilinear form on (u, ϕ) and $(0, \iota(u))$, which results in

$$\Omega_{\text{can}}\left((u, \phi), (0, \iota(u))\right) = \iota(u)(u) \neq 0.$$

Using a similar argument we can deal with the case where $\phi \neq 0$. We then consider $\left(\iota^{-1}(\phi), 0\right)$, such that

$$\Omega_{\text{can}}\left((u, \phi), \left(\iota^{-1}(\phi), 0\right)\right) = \phi\left(\iota^{-1}(\phi)\right) \neq 0.$$

This proves that Ω_{can} is a linear symplectic form on $V \oplus V^*$, we call this form the **linear canonical form of** V. //

Even though these examples are important and show that linear symplectic forms pop up anywhere, we will now give a similar example that is not a symplectic vector space.

Example 2.4. Take the vector space $\mathbb{R}^{2n-1}$ which we will identify with the $\mathbb{R}^{2n-1} \oplus \{0\} \subset \mathbb{R}^{2n}$. Then define the bilinear form $\Omega = \Omega_{\text{st},2n}\big|_{\mathbb{R}^{2n-1}}$. We now want to show that this bilinear form is degenerate, implying that it is not a linear symplectic form.

To find a vector u for which $\Omega(u, \cdot) = 0$, we use the standard basis for $\mathbb{R}^{2n-1}$. Then by choosing the nth basis vector as u it should be clear that $\Omega(u, v) = 0$ for all $v \in \mathbb{R}^{2n-1}$. Therefore, Ω is not a linear symplectic form and $\left(\mathbb{R}^{2n-1}, \Omega\right)$ is not a symplectic vector space. //

The definition of a symplectic vector space is not only dependent on the bilinear form but also on the vector space. As we might expect this is due to the non-degeneracy condition, which we also saw fail in Example 2.4.

Furthermore, notice that Examples 2.2 and 2.3 look quite similar. If we were to adopt the notation of $\langle u, \phi \rangle_e = \phi(u)$ for some $u \in V$ and $\phi \in V^*$, we notice that the canonical form can be written as

$$\Omega\left((u, \phi), (v, \psi)\right) = \langle u, \psi \rangle_e - \langle v, \phi \rangle_e .$$

This is the same expression as that of $\Omega_{\text{st},2n}$. This can be explained as $\mathbb{R}^{2n} \cong \mathbb{R}^n \times \mathbb{R}^n \cong \mathbb{R}^n \times (\mathbb{R}^{n*})$. However, it is not really clear how $\langle \cdot, \cdot \rangle_e$ and $\langle \cdot, \cdot \rangle_n$ are related. This relation can be made more clear by the use of certain isomorphisms, which we will discuss in the next section.

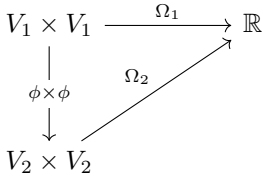


Figure 2.1: A commutative diagram where (V_1, Ω_1) and (V_2, Ω_2) are symplectic vector spaces and $\phi : V_1 \to V_2$ is an isomorphism. Here $\phi \times \phi$ denotes the isomorphism that maps $(u, v) \in V_1 \times V_1$ to $(\phi(u), \phi(v)) \in V_2 \times V_2$.

2.2 Symplectomorphisms

The symplectic vector spaces form a new category of mathematical structures. Naturally, we want to look for the functions that preserve this structure. In this section, we will introduce such functions as linear symplectomorphisms. For these functions to be natural with respect to the structure, they should preserve both the vector space structure and the symplectic structure. Hence, it is clear that linear symplectomorphisms should be isomorphisms to preserve the vector space structure. To preserve the linear symplectic form, we would want them to satisfy the commutative diagram of Figure 2.1. This commutative diagram can be written formally using the pullback of bilinear forms.

Definition 2.5. Let $\Omega_2 : V_2 \times V_2 \to \mathbb{R}$ be a bilinear form and $\phi : V_1 \to V_2$ a linear map, then the **pullback of Ω_2 by ϕ**, denoted by $\phi^*\Omega_2$, is defined as

$$\phi^*\Omega_2 : V_1 \times V_1 \to \mathbb{R} : (u, v) \mapsto \Omega_2(\phi(u), \phi(v)).$$

It follows that $\phi^*\Omega_2$ is a bilinear form on V_1, by using that Ω_2 is bilinear and ϕ is linear. //

The pullback of a bilinear form is comparable to the pre-composition in each argument. If we compare this operation to the commutative diagram in Figure 2.1, we notice that this is the exact operation for a linear symplectomorphism to make the diagram commute. This formalizes the definition of linear symplectomorphisms.

Definition 2.6. Let (V_1, Ω_1) and (V_2, Ω_2) be symplectic vector spaces. A **linear symplectomorphism** between (V_1, Ω_1) and (V_2, Ω_2) is a linear isomorphism $\phi : V_1 \to V_2$ such that $\phi^*\Omega_2 = \Omega_1$. If such a linear symplectomorphism exists, then (V_1, Ω_1) and (V_2, Ω_2) are called **symplectomorphic**. //

As isomorphisms are invertible and $\left(\phi^{-1}\right)^* \circ \phi^* = \mathrm{Id}$, it is clear that being symplectomorphic is an equivalence relation. The study of linear symplectic geometry then concerns itself with finding invariants of such linear symplectomorphisms. In the next section, we will look for such an invariant. Firstly, let us consider the similarity we saw in the last section using the context of symplectomorphism.

Example 2.7. Take the symplectic vector spaces $\left(\mathbb{R}^{2n}, \Omega_{\mathrm{st},2n}\right)$ and $(\mathbb{R}^n \times (\mathbb{R}^n)^*, \Omega_{\mathrm{can}})$. Remark that there is a natural isomorphism $\iota : \mathbb{R}^n \to (\mathbb{R}^n)^*$ generated by the standard inner product, i.e. $\iota(u)(v) = \langle u, v \rangle_n$. By again using the identification of $\mathbb{R}^{2n}$ as $\mathbb{R}^n \times \mathbb{R}^n$, we define the mapping ϕ as

$$\phi : \mathbb{R}^{2n} \to \mathbb{R}^n \times (\mathbb{R}^n)^* : (u, v) \mapsto (u, \iota(v)).$$

We can see that this is an isomorphism as ι is an isomorphism such that ϕ is linear and the inverse is given by $\phi^{-1} : \mathbb{R}^n \times (\mathbb{R}^n)^* \to \mathbb{R}^{2n} : (u, \phi) \mapsto \left(u, \iota^{-1}(\phi)\right)$. Furthermore, it follows that

$$\begin{aligned}
\phi^*\Omega_{\mathrm{can}}((u_1, u_2), (v_1, v_2)) &= \Omega_{\mathrm{can}}(\phi(u_1, u_2), \phi(v_1, v_2)) = \Omega_{\mathrm{can}}((u_1, \iota(u_2)), (v_1, \iota(v_2))) \\
&= \iota(v_2)(u_1) - \iota(u_2)(v_1) = \langle v_2, u_1 \rangle_n - \langle u_2, v_1 \rangle_n \\
&= \langle u_1, v_2 \rangle_n - \langle u_2, v_1 \rangle_n = \Omega_{\mathrm{st},2n}((u_1, u_2), (v_1, v_2)).
\end{aligned}$$

Hence, $\left(\mathbb{R}^{2n}, \Omega_{\mathrm{st},2n}\right)$ and $(\mathbb{R}^n \times (\mathbb{R}^n)^*, \Omega_{\mathrm{can}})$ are symplectomorphic and ϕ is a linear symplectomorphism. //

2.3 Standard Form

In this section, we will search for an invariant of linear symplectomorphisms. The primary way of doing this in linear algebra is by finding a suitable basis such that the linear symplectic form is of some standard form. For example, the inner product of a real product space is symmetric and can thus be represented by a diagonal matrix, see [8, Theorem 6.35]. We would also like to find such a simple representation for an arbitrary symplectic form. The corollary of [8, Theorem 6.34] already shows that this can not be a diagonal matrix. Luckily, we can recognise that the symplectic form of Example 2.2, $\left(\mathbb{R}^{2n}, \Omega_{\mathrm{st},2n}\right)$ has a very simple matrix representation in the standard basis.

$$\Omega_{\mathrm{st},2n}\left(e_i, e_j\right) = \begin{cases} 1, & j - i = n \text{ with } 1 \leq i \leq n, \\ -1, & i - j = n \text{ with } 1 \leq j \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

$$\implies \Omega_{\mathrm{st},2n}\left(x, y\right) = x^T \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} y.$$

This entices us to look for a basis of a symplectic vector space, such that its linear symplectic form is represented in this form.

Theorem 2.8. *Let (V, Ω) be a symplectic vector space of dimension m, then there exists an $n \in \mathbb{N}$ such that $m = 2n$ and a basis $\beta = \{u_1, \ldots, u_n, v_1, \ldots, v_n\}$ such that*

$$\Omega\left(u_i, v_j\right) = \delta_{ij}, \quad \Omega\left(u_i, u_j\right) = 0 = \Omega\left(v_i, v_j\right).$$

In this basis, the action of the symplectic form on some $u, v \in V$ is given by

$$\Omega\left(u, v\right) = \left[u\right]_\beta^T \begin{bmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{bmatrix} \left[v\right]_\beta.$$

*Such a basis is called the **symplectic basis** of (V, Ω).*

Proof. Let (V, Ω) be a symplectic vector space of dimension m. Take $V_1 = V$. We know by the non-degeneracy of the symplectic form that for any non-zero $u_1 \in V_1$, there must exist a $v_1 \in V_1$ such that $\Omega\left(u_1, v_1\right) \neq 0$. As Ω is bilinear, we may assume that $\Omega\left(u_1, v_1\right) = 1$. Then define the following sets:

$$W_1 := \operatorname{span}\left\{u_1, v_1\right\}, \quad W_1^\Omega := \left\{u \in V_1 | \Omega\left(u, v\right) = 0 \ \forall v \in W_1\right\}.$$

As Ω is bilinear, we can deduce that W_1^Ω is a linear subspace of V. We will show that $V = W_1 \oplus W_1^\Omega$. Take an arbitrary $w \in W_1 \cap W_1^\Omega$. Because $w \in W_1$ we can find a and b such that $w = au_1 + bv_1$. However, as $w \in W_1^\Omega$, we can also deduce that

$$0 = \Omega\left(w, u_1\right) = -b, \quad 0 = \Omega\left(w, v_1\right) = a.$$

Thus $w = 0$ and therefore $W_1 \cap W_1^\Omega = \{0\}$. Secondly, suppose that $w \in V$ with $\Omega\left(w, u_1\right) = \alpha$ and $\Omega\left(w, v_1\right) = \beta$, then

$$w = \left(-\alpha v_1 + \beta u_1\right) + \left(w + \alpha v_1 - \beta u_1\right).$$

Remark that $-\alpha v_1 + \beta u_1 \in \operatorname{span}\left\{v_1, u_1\right\} = W$. Using the bilinearity and skew-symmetry of the symplectic form it follows that

$$\Omega\left(w + \alpha v_1 - \beta u_1, u_1\right) = \Omega\left(w, u_1\right) + \alpha\omega\left(v_1, u_1\right) - \beta\Omega\left(u_1, u_1\right) = \alpha - \alpha = 0,$$
$$\Omega\left(w + \alpha v_1 - \beta u_1, v_1\right) = \Omega\left(w, v_1\right) + \alpha\omega\left(v_1, v_1\right) - \beta\Omega\left(u_1, v_1\right) = \beta - \beta = 0.$$

Hence, it follows that $w + \alpha v_1 - \beta u_1 \in W^\Omega$ and thus $w \in W_1 + W_1^\Omega$. This shows that $V_1 = W_1 \oplus W_1^\Omega$.

Now by choosing $V_2 = W_1^\Omega$, we want to prove that $(V_2, \Omega|_{V_2})$ is a symplectic vector space, such that we can finish the proof using induction. It should be clear that Ω_{V_2} is still skew-symmetric. The non-degeneracy can be checked for an arbitrary element $u \in V_2$. As $V_2 \subset V$ and the fact that Ω is non-degenerate, it follows that there exists a $v \in V$ such that $\Omega(u, v) \neq 0$. We will now show that we can choose $v \in V_2$. We have already shown that $v = \tilde{v} + \overline{v}$, where $\tilde{v} \in W_1$ and $\overline{v} \in V_2 = W_1^\Omega$. We can then deduce that

$$0 \neq \omega(u, v) = \Omega(u, \tilde{v} + \overline{v}) = \Omega(u, \tilde{v}) + \Omega(u, \overline{v}) = \Omega(u, \overline{v})$$

Hence, there is some vector $\overline{v}$ such that $\Omega|_{V_2}(u, \overline{v}) = 0$. This proves that $\Omega|_{V_2}$ is non-degenerate and a symplectic form.

As we mentioned, we can then iterate this process and create a chain of $\{W_i\}_{i=1}^k$ such that

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k \oplus W_k^\Omega.$$

As the dimension of W_i is two for all i and the dimension of V is finite, it follows that there is some n such that $W_n^\Omega = \{0\}$, implying that $V = W_1 \oplus \cdots \oplus W_n$. If we join all the bases for each W_i, we then get a basis for V, which is then given by $\{u_1, \ldots, u_n, v_1, \ldots, v_n\}$. The action of the bilinear form on this basis is given by

$$\Omega(u_i, v_j) = \delta_{ij}, \quad \Omega(u_i, u_j) = 0 = \Omega(v_i, v_j), \quad V = W_1 \oplus \cdots \oplus W_n.$$

Therefore, the bilinear form takes the following matrix form in this basis:

$$\Omega(u, v) = [u]_\beta^T \begin{bmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{bmatrix} [v]_\beta .$$

From this basis, it is also clear that $\dim V = 2n$. □

As we saw above, the symplectic basis for the trivial symplectic vector space of dimension $2n$ is the standard basis on $\mathbb{R}^{2n}$. However, the symplectic basis tells us much more.

Theorem 2.9. *A $2n$-dimensional symplectic vector space is symplectomorphic to $\left(\mathbb{R}^{2n}, \Omega_{st,2n}\right)$.*

Proof. Let (V, Ω) be a $2n$-dimensional symplectic vector space. By Theorem 2.8 there exists a symplectic basis of $\mathbb{R}^{2n}$, call this basis β. We now construct the function, which is our linear symplectomorphism

$$\phi_\beta : V \to \mathbb{R}^{2n} : v \mapsto [v]_\beta .$$

As ϕ_β is the standard representation of V with respect to β, in other words, it constitutes the choice of a basis, it is an isomorphism. Furthermore, from the definition of a symplectic basis, it follows that $\Omega = {\phi_\beta}^* \Omega_{\mathrm{st},2n}$. Thus ϕ_β is a linear symplectomorphism and therefore (V, Ω) is symplectomorphic to $\left(\mathbb{R}^{2n}, \Omega_{\mathrm{st},2n}\right)$. □

Theorem 2.9 means that we can classify each symplectic vector space by its dimension. The existence of such a simple invariant makes linear symplectic geometry quite straightforward as it can be reduced to the study of the spaces $\left(\mathbb{R}^{2n}, \omega_{\mathrm{st}}\right)$. Furthermore, we can use this theorem to easily prove that some vector spaces do not admit a linear symplectic form, as their dimension should be even. This then immediately implies that $\mathbb{R}^{2n-1}$ can not be a symplectic vector space and that the bilinear form in Example 2.4 is not a linear symplectic form.

Chapter 3

Symplectic Geometry on Manifolds

In the previous chapter, we introduced linear symplectic geometry in terms of symplectic vector spaces. Now we will evolve this to a more general geometric theory by extending it to manifolds. This is done naturally by using the locally linear structure of the tangent space, hence, imposing a symplectic form as a non-degenerate 2-form which we will also require to be closed. We will see how this theory differs from its linear counterpart in the main focus of this chapter: Darboux's theorem. This is the non-linear version of Theorem 2.9, but it requires a much more involved theory on differential forms. In this chapter, we follow the structure and definitions of chapters 1, 7 and 8 in [4]. Furthermore, we will use differential forms and the exterior derivative in this chapter, the definitions of which are recalled in Section 3.3.1. If one is not yet familiar with these concepts, one should refer to more detailed sources like [13, 18, 22].

3.1 Basics and Definitions

We will start this section with the definition of a symplectic manifold, followed by some examples. As mentioned, we want to generalise a linear algebra structure to differential geometry. Hence, we want to make use of the linear structure a manifold has in the form of its tangent space. This method is comparable to that of Riemannian geometry, which generalises an inner product space using a symmetric covariant 2-tensor field that is positive definite everywhere, such that the tensor field is an inner product at each point. Similarly, we generalise the linear symplectic forms by looking at a tensor that is a linear symplectic form at each point.

Definition 3.1. Let $\mathcal{M}$ be a manifold, a 2-form ω is **symplectic** if it is closed, i.e. $d\omega = 0$ and $\omega_p : T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$ is symplectic for all $p \in \mathcal{M}$ in the sense of vector spaces. The pair $(\mathcal{M}, \omega)$ is then called a **symplectic manifold**. //

Once again we will go over some basic examples of symplectic vector spaces before we proceed. The first two of these are the equivalents to Examples 2.2 and 2.3.

Example 3.2. Let us define the **trivial symplectic manifold of dimension** $2n$. Take the manifold $\mathbb{R}^{2n}$ and take the global coordinates $\left(x^1, \dots, x^n, y^1, \dots, y^n\right)$. Define the 2-form $\omega_{\mathrm{st},2n}$ as

$$\omega_{\mathrm{st},2n} = \sum_{i=1}^{n} dx^i \wedge dy^i.$$

This form is symplectic as for each $p \in \mathcal{M}$ the set

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

forms a symplectic basis for $T_p\mathcal{M}$. Furthermore, one can easily check that this form is closed with a calculation.

$$d\omega_{\mathrm{st},2n} = d\sum_{i=1}^{n} dx^i \wedge dy^i = \sum_{i=1}^{n} d\left(dx^i \wedge dy^i\right) = \sum_{i=1}^{n} \left(d^2\right) x^i \wedge dy^i - dx^i \wedge \left(d^2\right) y^i = 0.$$

This proves that $\left(\mathbb{R}^{2n}, \omega_{\mathrm{st},2n}\right)$ is a symplectic manifold. We call $\omega_{\mathrm{st},2n}$ a **trivial symplectic form** of which we sometimes mention the dimension. //

Example 3.3. Let $\mathcal{M}$ be an n-dimensional manifold and let $\mathrm{T}^*\mathcal{M}$ be its **cotangent bundle**, which is defined as

$$\mathrm{T}^*\mathcal{M} = \coprod_{p\in\mathcal{M}} \left(\mathrm{T}_p\mathcal{M}\right)^* = \coprod_{p\in\mathcal{M}} \mathrm{T}_p^*\mathcal{M}.$$

Firstly, we define a **tautological 1-form on** $\mathrm{T}^*\mathcal{M}$, denoted by α_{tau}, pointwise as

$$\alpha_{\mathrm{tau}(p,\xi)} = d\pi^*_{(p,\xi)}\left(\xi\right),$$

We then define the **canonical symplectic form on** $\mathrm{T}^*\mathcal{M}$, denoted as ω_{can}, as

$$\omega_{\mathrm{can}} = -d\alpha_{\mathrm{tau}}.$$

It should be clear that this is a closed and alternating 2-tensor. To show that it is non-degenerate we will consider it in coordinates. Let us first look at the structure of $\mathrm{T}^*\mathcal{M}$, which has some natural coordinates induced by a chart on $\mathcal{M}$. Suppose that $\left(U, \left(x^1, \ldots, x^n\right)\right)$ is a chart around $p \in \mathcal{M}$. Via the differentials, $\left(dx^1_p, \ldots, dx^n_p\right)$, we induce the following coordinates on $\mathrm{T}^*\mathcal{U}$:

$$\mathrm{T}^*\mathcal{U} \to \mathbb{R}^{2n} : \xi_i dx^i|_p \mapsto \left(x^1\left(p\right), \ldots, x^n\left(p\right), \xi_1, \ldots, \xi_n\right).$$

In these coordinates the projection is trivial and the differential is of the form $d\pi^*_{(x,\xi)}\left(dx^i\right) = dx^i$. Hence, in coordinates the tautological 1-form is given by

$$\alpha_{\mathrm{tau}(x,\xi)} = \xi_i dx^i.$$

As the charts are smooth, the tautological form is also smooth. Furthermore, the coordinate form of α_{tau} implies that ω_{can} is given by

$$\omega_{\mathrm{can}} = \sum_{i=1}^{n} dx^i \wedge d\xi_i,$$

in the coordinates $\left(x^1, \ldots, x^n, \xi_1, \ldots, \xi_n\right)$. This form is locally equivalent to that of Example 3.2, therefore it is symplectic by the same reasoning. It follows that $\left(\mathrm{T}^*\mathcal{M}, \omega_{\mathrm{can}}\right)$ is a symplectic manifold. //

Proposition 3.4. *Every manifold admits a symplectic form on its cotangent bundle.*

Proof. This follows directly from the construction of the canonical symplectic form in Example 3.3. □

There are of course many different symplectic manifolds and even different symplectic forms on a single manifold.

Example 3.5. Take the 2-sphere, denoted by S^2. For the construction of the symplectic form, we identify S^2 as a subset of $\mathbb{R}^3$, which we endow with the dot product, denoted by $\langle\cdot,\cdot\rangle$, and the cross product, denoted by $\times$. We then define the $\mathring{\omega}$ at a point $p \in S^2$ as

$$\mathring{\omega}_p : \mathrm{T}_pS^2 \times \mathrm{T}_pS^2 \to \mathbb{R} : (u, v) \mapsto \langle p, u \times v\rangle.$$

By the skew-symmetry and bilinearity of the cross product and the linearity of the inner product, it follows that $\mathring{\omega}_p$ is a symplectic form on T_pS^2. As the inner product and cross product are linear in both components, it is also smooth. Due to it being top-degree, it is closed. It follows that $\mathring{\omega}$ is a symplectic form and $\left(S^2, \mathring{\omega}\right)$ is a symplectic manifold. //

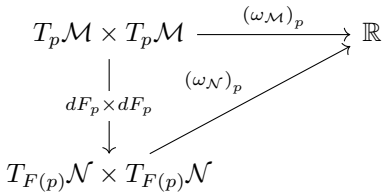


Figure 3.1: A commutative diagram showing the way the symplectomorphism $F : \mathcal{M} \to \mathcal{N}$ should commute with the symplectic forms. Here, dF_p denotes the pointwise pushforward of F.

Example 3.6. Let us define a symplectic form on $\mathbb{R}^2$ that is not the trivial symplectic form. Take the global coordinates (x, y) of $\mathbb{R}^2$ and define the 2-form ω_{exp} as

$$\omega_{\text{exp}} = e^{-\left(x^2+y^2\right)} dx \wedge dy.$$

We can quite easily recognise that this is a skew-symmetric non-degenerate form. Furthermore, as it is of top-degree it is also closed. Therefore, $\left(\mathbb{R}^2, \omega_{\text{exp}}\right)$ is a symplectic vector space. //

With these examples, we will again go on to the geometric side of symplectic manifolds.

3.2 Symplectomorphisms

With this new category of structured objects, we look for a class of functions preserving our structure, which we will call symplectomorphisms. As they should be functions preserving manifolds they ought to be diffeomorphisms. Furthermore, they should preserve the structure added to the tangent spaces as well. Hence, we want the function to satisfy the commutative diagram of Figure 3.1. This is done using the pullback of differential forms, see [13, p. 360]. This then leads to the following definition.

Definition 3.7. Let $(\mathcal{M}_1, \omega_1)$ and $(\mathcal{M}_2, \omega_2)$ be symplectic manifolds. A **symplectomorphism** between $(\mathcal{M}_1, \omega_1)$ and $(\mathcal{M}_2, \omega_2)$ is a diffeomorphism $\phi : \mathcal{M}_1 \to \mathcal{M}_2$ such that $\phi^* \omega_2 = \omega_1$. If such a symplectomorphism exists, then $(\mathcal{M}_1, \omega_1)$ and $(\mathcal{M}_2, \omega_2)$ are **symplectomorphic**. //

For vector spaces, the next important result was Theorem 2.8, which described how we can find a basis in which the symplectic form is written as

$$[\Omega]_\beta = \begin{pmatrix} 0 & \text{Id} \\ -\text{Id} & 0 \end{pmatrix}.$$

Implying that any symplectic vector space is isomorphic to $\left(\mathbb{R}^{2n}, \Omega_{\text{st},2n}\right)$ for some n. However, this runs into a problem when we try it with manifolds. For example, consider $\left(\mathbb{R}^2, \omega_{\text{st},2}\right)$ and $\left(S^2, \mathring{\omega}\right)$ as in Example 3.5. If we could construct a symplectomorphism between these spaces, it would in particular be a homeomorphism between S^2 and $\mathbb{R}^2$. However, as S^2 is bounded and closed in $\mathbb{R}^3$ it is compact and homeomorphisms preserve compactness. This would then imply that $\mathbb{R}^2$ is compact, which leads to a contradiction. Thus, a symplectomorphism between $\left(\mathbb{R}^2, \omega_{\text{st},2}\right)$ and $\left(S^2, \mathring{\omega}\right)$ can not exist.

We might then expect that we should at least be able to find a symplectomorphism between symplectic forms on the same manifold, but Example 3.6 showcases that this is also not true. If a symplectomorphism ϕ between $\left(\mathbb{R}^2, \omega_{\text{st},2}\right)$ and $\left(\mathbb{R}^2, \omega_{\text{exp}}\right)$ as in Example 3.6, would exist, it should preserve integrals, see [13, Proposition 16.6]. Hence the following equation should hold.

$$\int_{\mathbb{R}^2} \omega_{\text{exp}} = \int_{\mathbb{R}^2} \phi^* \omega_{\text{exp}} = \int_{\mathbb{R}^2} \omega_{\text{st},2}. \tag{3.1}$$

Now we can calculate both sides of the equation. The integral of $\omega_{\text{st},2}$ diverges as $\omega_{\text{st},2}$ is a positive constant on the whole of $\mathbb{R}^2$. Meanwhile, the integral of ω_{exp} can be calculated using a

simple change of coordinates and a substitution as in the proof of the Gauss integral.

$$\int_{\mathbb{R}^2} \omega_{\exp} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\left(x^2+y^2\right)} dxdy = \int_0^{2\pi}\int_0^{\infty} e^{-r^2} r dr d\phi = \pi \int_0^{\infty} e^{-u} du = \pi.$$

As this integral does not diverge, we can conclude from Equation 3.1 that a symplectomorphism between $\left(\mathbb{R}^2, \omega_{\mathrm{st},2}\right)$ and $\left(\mathbb{R}^2, \omega_{\exp}\right)$ can not exist. Furthermore, this implies that these two symplectic manifolds are not symplectomorphic.

From these examples, we can conclude that we can not simply study the global behaviour of $\left(\mathbb{R}^2, \omega_{\mathrm{st},2}\right)$ in symplectic geometry. The construction of a symplectomorphism is not even limited to the topological properties of the manifolds and may even fail for two symplectic forms on the same manifold.

3.3 Normal Coordinates

We saw in the previous section that there are many symplectic manifolds which are not symplectomorphic to $\left(\mathbb{R}^{2n}, \omega_{\mathrm{st},2n}\right)$. This means we cannot encapsulate the global behaviour of symplectic manifolds by just studying this simple model space. Luckily, the structure of a manifold lends itself perfectly to studying local behaviour instead. To this extent, we will look for certain coordinates in which the symplectic form is of a 'nice' form. Such coordinates are generally called normal coordinates.

To get an idea of what is possible, we will once again compare symplectic manifolds to Riemannian manifolds. We already remarked that they are constructed similarly; take a structure on a vector space and implement it at each point to a tangent space. For the symplectic manifold, we use symplectic forms and for the Riemannian case, we consider inner products. For both linear structures, there are no invariants that determine the structures, i.e. the inner product is diagonal and the linear symplectic form is anti-diagonal and skew-symmetric, see Theorem 2.8. We would hope that this simplistic structure would naturally be inherited by the manifold structure locally. Riemannian manifolds already show that this is not necessarily true. At a point, we can find a chart in which the metric becomes diagonal, but when extended to a neighbourhood this fails. Instead of a diagonal form, it becomes a function of the Riemannian curvature. We will not prove the following theorem, but one can refer to [9, Corollary 9.8] for a thorough exposition and proof.

Theorem 3.8. *Let $(\mathcal{M}, g)$ denote an Riemannian manifold with curvature tensor R_{ijkl}. Around any point $p \in \mathcal{M}$ there is a coordinate chart $\left(U, \left(x^i\right)\right)$ centred at p, such that the metric is given by*

$$g_{ij}|_U = \delta_{ij} - \frac{1}{3}\sum_{kl=1}^{n} R_{kilj}(p)\, x^k x^l - \frac{1}{6}\sum_{klm=1}^{n} \nabla_k R_{limj}(p)\, x^k x^l x^m + h \circ x,$$

where $\lim_{x\to 0} h(x) \Big/ \|x\|^5 = 0$.

This theorem tells us that a Riemannian manifold has some local invariant, namely the curvature, that determines whether two Riemannian manifolds are locally isometric. We will now prove this theorem as it is not in our current domain of study. However, we can give an informal argument which explains why we expect to find such an invariant. This is called the counting argument and was already given by Riemann in [19], see [20, Chapter 4] for a translation. In this 'proof', he tries to count the number of degrees of freedom in the metric and the amount we can determine under a transformation of coordinates.

Suppose that $(\mathcal{M}, g)$ is an n dimensional Riemannian manifold. Around a point $p \in \mathcal{M}$ we can write g is some coordinate chart $\left(U, \left(x^i\right)\right)$ as $g|_U = \sum_{i=1}^n g_{ij} dx^i dx^j$, such that g is fully determined in the neighbourhood U by n^2 functions. This implies that we have n^2 degrees of freedom, but we lose $n(n-1)/2$ degrees of freedom to the symmetry of the metric. Hence, there are just $n^2 - n(n-1)/2 = n(n+1)/2$ degrees of freedom left. If we then take into account the ambiguity of choosing coordinates, we lose another n degrees of freedom. This leaves us with

just $n(n-1)/2$ degrees of freedom we can not determine further. Hence, we have shown that there are some degrees of freedom in our choice of metric, which we conveniently capture in the Riemannian curvature.

This goes to show that we can not ensure the existence of a chart in which the metric is given by $g|_U = \sum_{i=1}^n dx^i dx^i$. A Riemannian manifold does therefore not directly inherit the properties of the inner product locally. If there exists a map from the Riemannian manifold to the Euclidean space which locally restricts to an isometry, we call this Riemannian manifold flat. It should be clear that there exist some coordinates $\left(U, \left(x^i\right)\right)$ on a flat Riemannian manifold such that the metric is given by $g|_U = \sum_{i=1}^n dx^i dx^i$. We can then find a direct connection between the flatness of a Riemannian manifold and its Riemannian curvature tensor, something we will again not prove here, but a proof can be found in [14, Theorem 7.10].

Proposition 3.9. *A Riemannian manifold is flat if and only if its Riemannian curvature tensor vanishes identically.*

This is a rather simple condition for a Riemannian manifold to be flat. However, we will show that this holds in general for symplectic forms in a theorem that is called Darboux's theorem. A proof of this was first given by Jean Gaston Darboux in 1882 [6], but later it was simplified by Weinstein [23] based on a method developed by Moser [15]. This second proof is the one we will follow in Section 3.3.2. This proof requires some more knowledge of the geometry of differential forms and their associated operator. Moreover, even to make the counting argument work for symplectic forms, we would need this extra knowledge. For now, we will simply state the theorem and prove it after going into the prerequisites of differential geometry.

Theorem 3.10. *Around every point on a symplectic manifold $(\mathcal{M}, \omega)$, there exists a neighbourhood $\left(U, \left(x^1, \ldots, x^n, y^1, \ldots, y^n\right)\right)$ such that on U the symplectic form is given by*

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

Such coordinates are called Darboux coordinates.

3.3.1 Operators on Differential Forms

Before we can prove Darboux's theorem, we will delve a bit deeper into the structure of differential forms and operators on them. In this section we will prove general results for k-forms, which are of course all applicable to symplectic forms as they are a special case with $k = 2$. We will apply much of this theory in the proof of Darboux's theorem. Specifically, we will first discuss three operators: interior multiplication, exterior derivative and Lie derivative. Afterwards, we look into their connection in the form of Cartan's magic formula and prove Poincaré's lemma, which gives information on the local structure of closed differential forms. The definition and proofs will mainly be taken from [13], but the proof of Cartan's magic formula takes some insights from [18] and the proof of Poincaré's lemma is an adaptation of the proof of Corollary 13.2.14 in [18].

Firstly, we remind ourselves of the definition of k-forms. These stem from linear algebra, namely the exterior algebra. A k-form on a manifold can be interpreted as putting an alternating covariant k-tensor at each point in a smooth manner.

Definition 3.11. For a real vector space V, we call a map $\omega : \underbrace{V \times \cdots \times V}_{k\text{-times}} \to \mathbb{R}$ an **alternating k-form** if it is multilinear,

$$\omega\left(v_1, \cdots, a v_i + \tilde{a}\tilde{v}_i, \ldots, v_k\right) = a\omega\left(v_1, \cdots, v_i, \ldots, v_k\right) + \tilde{a}\omega\left(v_1, \cdots, \tilde{v}_i, \ldots, v_k\right)$$

and it is alternating

$$\omega\left(v_{\sigma(1)}, \ldots, v_{\sigma(k)}\right) = \operatorname{sign}(\sigma)\, \omega\left(v_1, \ldots, v_k\right).$$

We denote the space of alternating k-forms on V as $\bigwedge^k V^*$.

On this space we define the **wedge product** $\wedge : \bigwedge^k V^* \times \bigwedge^l V^* \to \bigwedge^{k+l} V^*$ as

$$(\alpha \wedge \beta)(v_1, \ldots, v_{k+l}) = \frac{1}{k!l!} \sum_{\sigma \in S(k+l)} \operatorname{sign}(\sigma)\, \alpha\left(v_{\sigma(1)} \ldots, v_{\sigma(k)}\right) \beta\left(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}\right),$$

where $S(n)$ is the permutation group on n elements. //

Example 3.12. Suppose that V is a real vector space and $\alpha, \beta \in \bigwedge^1 V^*$. The action of $\alpha \wedge \beta$ on $v_1, v_2 \in V$ is then given by

$$\begin{aligned}(\alpha \wedge \beta)(v_1, v_2) &= \frac{1}{1!1!} \sum_{\sigma \in S(2)} \operatorname{sign}(\sigma)\, \alpha\left(v_{\sigma(1)}\right) \beta\left(v_{\sigma(2)}\right) \\ &= \alpha(v_1)\beta(v_2) - \alpha(v_2)\beta(v_1) = (\alpha \otimes \beta - \beta \otimes \alpha)(v_1, v_2).\end{aligned}$$

Hence, the wedge product of two 1-forms is given by $\alpha \wedge \beta = \alpha \otimes \beta - \beta \otimes \alpha$. //

As we can generate $(k+l)$-forms from a k-form and an l-form, we would want this operator to be one of a single space. To do this we define a more general algebraic structure.

Definition 3.13. The **exterior algebra of** V^* is defined as

$$\bigwedge V^* = \bigoplus_{k=0}^{n} \bigwedge^k V^*.$$

Furthermore, we extend the wedge product to a binary operation $\wedge : \bigwedge V^* \times \bigwedge V^* \to \bigwedge V^*$ using a bilinear extension of $\wedge : \bigwedge^k V^* \times \bigwedge^l V^* \to \bigwedge^{k+l} V^*$. In other words, we decompose $\alpha, \beta \in \bigwedge V^*$ as

$$\alpha = \alpha_0 + \cdots + \alpha_k, \qquad \alpha_i \in \bigwedge^i V^*, \qquad \beta = \beta_0 + \cdots + \beta_l, \qquad \beta_i \in \bigwedge^i V^*.$$

Such that $\alpha \wedge \beta$ is then given by

$$\alpha \wedge \beta = \sum_{n=0}^{k+l} \sum_{i+j=n} \alpha_i \wedge \beta_j. \tag{3.2}$$

This defines the exterior algebra $(\bigwedge V^*, +, \wedge)$. //

Proposition 3.14. *The triplet $(\bigwedge V^*, +, \wedge)$ is an graded-commutative graded associative unital algebra. Meaning that $\bigwedge^k V^* \wedge \bigwedge^l V^* \subset \bigwedge^{k+l} V^*$ and for $\alpha \in \bigwedge^k V^*, \beta \in \bigwedge^l V^*$ the wedge product satisfies $\alpha \wedge \beta = (-1)^{kl}\, \beta \wedge \alpha$.*

Proof. The fact that it is graded-commutative, and associative is a direct consequence of the definition of the wedge product as in Equation 3.2. The unit is $1 \in \mathbb{R}$, hence, it is unital. □

We can extend these alternating k-forms on real vector spaces to differential forms on manifolds by defining them pointwise.

Definition 3.15. An **alternating k-tensor** is an element of the set

$$\bigwedge^k \mathrm{T}^*\mathcal{M} = \coprod_{p \in \mathcal{M}} \bigwedge^k \mathrm{T}_p^*\mathcal{M}.$$

A **differential k-form**, or simply **k-form**, is then a smooth section of $\bigwedge^k \mathrm{T}^*\mathcal{M}$, where we denote the set of k-forms by

$$\Omega^k(\mathcal{M}) = \Gamma\left(\bigwedge^k \mathrm{T}^*\mathcal{M}\right).$$

This naturally forms a real vector space with pointwise addition and scalar multiplication. We define the **wedge product** $\wedge : \Omega^k(\mathcal{M}) \times \Omega^l(\mathcal{M}) \to \Omega^{k+l}(\mathcal{M})$ pointwise by $(\omega \wedge \eta)_p = \omega_p \wedge \eta_p$. Leading to the definition of the algebra

$$\Omega(\mathcal{M}) = \bigoplus_{k=0}^{n} \Omega^k(\mathcal{M}).$$

This gives us the algebra $(\Omega(\mathcal{M}), +, \wedge)$, where $\wedge$ is the bilinear extension of the wedge product above as in Definition 3.13 //

Proposition 3.16. *The algebra $(\Omega(\mathcal{M}), +, \wedge)$ is a graded-commutative graded associative unital algebra.*

Proof. This is a direct consequence of Proposition 3.14. In this case, the unit is given by the function$f : \mathcal{M} \to 1 \in C^\infty(\mathcal{M})$, hence, it is a unital algebra. □

In any smooth chart $\left(U, (x^i)\right)$ we can write an $\omega \in \Omega^k(\mathcal{M})$ as

$$\omega = \sum_{i_1 < \cdots < i_k} \omega_{i_1 \ldots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

where $\omega_{i_1,\ldots i_k} \in C^\infty(\mathcal{M})$. This notation is called ordered indices, but there are more possible notations. Most importantly

$$\omega = \sum_{i_1 \ldots i_k} \frac{1}{k!} \omega_{i_1 \ldots i_k} dx^i \wedge \cdots \wedge dx^{i_k}, \qquad \omega_{\sigma(i_1)\ldots\sigma(i_k)} = \operatorname{sign}(\sigma)\, \omega_{i_1 \ldots i_k} \in C^\infty(\mathcal{M}).$$

This is with unordered indices, which can be useful sometimes. Lastly, we may abbreviate our notation for conciseness to

$$\omega = \sum_I \omega_I dx^I.$$

Here $\omega_I \in C^\infty(\mathcal{M})$ and dx^I is an abbreviation for $dx^{i_1} \wedge \cdots \wedge dx^{i_k}$. All of these forms are equivalent, but some may be more convenient to use. They can always be identified by how the indices in the sum are written.

With the construction of the differential forms and structure of $\Omega(\mathcal{M})$ more clear, we will now look at some different operators.

Interior multiplication

The first operation we are interested in is interior multiplication. We will define this on the exterior algebra of a real vector space, which is then naturally extended to differential forms. The interior multiplication is defined such that it fixes the first argument of a k-form.

Definition 3.17. For an $\omega \in \bigwedge^k V^*$ and $v \in V$ we define the **interior multiplication of ω by** v, denoted as $\iota_v \omega$, by fixing the first argument in ω. This is more clear by considering its action on some $v_1, \ldots, v_{k-1} \in V$, then

$$\iota_v \omega(v_1, \ldots, v_{k-1}) = \omega(v, v_1, \ldots, v_{k-1}),$$

and $\iota_v \omega = 0$ if $k = 0$. The associated operator $\iota_v : \bigwedge^\star V^* \to \bigwedge^{\star-1} V^*$ is called the **interior multiplication by** v or **contraction by** v. //

The most important property of this operator is that it respects the structure of the exterior algebra and vector space.

Proposition 3.18. *The interior multiplication is $\mathbb{R}$-linear in v and ω, i.e. $\iota : V \times \bigwedge^k V^* \to \bigwedge^{k-1} V^* : (v, \omega) \mapsto \iota_v \omega$ is bilinear.*

Proof. The linearity in v follows from the definition, combined with the multilinearity of k-forms. Suppose that $\omega \in \bigwedge^k V^*$, $v, \tilde{v}, v_1, \ldots, v_{k-1} \in V$ and $a, \tilde{a} \in \mathbb{R}$, then

$$\begin{aligned}\iota_{av+\tilde{a}\tilde{v}}(\omega)(v_1, \ldots, v_{k-1}) &= \omega(av + \tilde{a}\tilde{v}, v_1, \ldots, v_{k-1}) \\ &= a\omega(v, v_1, \ldots, v_{k-1}) + \tilde{a}\omega(\tilde{v}, v_1, \ldots, v_{k-1}) \\ &= a\iota_v(\omega)(v_1, \ldots, v_{k-1}) + \tilde{a}\iota_{\tilde{v}}(\omega)(v_1, \ldots, v_{k-1}).\end{aligned}$$

Thus, ι is linear in its first component. For the second component, we use the definition of addition and scalar multiplication of tensor fields. Hence, for $\omega, \tilde{\omega} \in \bigwedge^k V^*$, $v, v_1, \ldots, v_{k-1} \in V$ and $a, \tilde{a} \in \mathbb{R}$ we have

$$\begin{aligned}\iota_v(a\omega + \tilde{a}\tilde{\omega})(v_1, \ldots, v_{k-1}) &= (a\omega + \tilde{a}\tilde{\omega})(v, v_1, \ldots, v_{k-1}) \\ &= a\omega(v, v_1, \ldots, v_{k-1}) + \tilde{a}\tilde{\omega}(v, v_1, \ldots, v_{k-1}) \\ &= a\iota_v(\omega)(v_1, \ldots, v_{k-1}) + \tilde{a}\iota_v(\tilde{\omega})(v_1, \ldots, v_{k-1}).\end{aligned}$$

Hence, ι is also linear in its second component. □

Besides respecting the vector space structure, it also works with the graded-commutative graded algebra structure of $\bigwedge V^*$.

Proposition 3.19. *The interior multiplication is a graded derivation of degree -1. In other words, $\iota_v : \bigwedge^\star V^* \to \bigwedge^{\star-1} V^*$ satisfies*

$$\iota_v(\omega \wedge \eta) = \iota_v\omega \wedge \eta + (-1)^k \omega \wedge \iota_v\eta, \tag{3.3}$$

where $\omega \in \bigwedge^k V^*$ and $\eta \in \bigwedge^l V^*$.

Proof. Take some v from a vector space V. As ι_v is $\mathbb{R}$-linear as a mapping $\iota_v : \bigwedge^\star V^* \to \bigwedge^{\star-1} V^*$, see Proposition 3.18, we may assume $\omega = \alpha^1 \wedge \cdots \wedge \alpha^k$ and $\eta = \alpha^{k+1} \wedge \cdots \wedge \alpha^{k+l}$, where $\alpha^i \in \bigwedge^1 V^*$. We can then prove Equation 3.3 by evaluating it in some vectors $v_1 \ldots, v_{k+l} \in V$ and using the inductive definition of the determinant, see [13, Proposition 14.11 (e)], and splitting the sum,

$$\begin{aligned}\iota_v(\omega \wedge \eta) = \iota_v\left(\alpha^1 \wedge \cdots \wedge \alpha^{k+l}\right) &= \sum_{i=1}^{k+l} (-1)^{i-1} \alpha^i(v)\, \alpha^1 \wedge \cdots \wedge \widehat{\alpha}^i \wedge \cdots \wedge \alpha^{k+l} \\ &= \left(\sum_{i=1}^{k} (-1)^{i-1} \alpha^i(v)\, \alpha^1 \wedge \cdots \wedge \widehat{\alpha}^i \wedge \cdots \wedge \alpha^k\right) \wedge \alpha^{k+1} \wedge \cdots \wedge \alpha^{k+l} \\ &\quad + (-1)^k \alpha^1 \wedge \cdots \wedge \alpha^k \left(\sum_{i=1}^{l} (-1)^{i-1} \alpha^{k+i}(v)\, \alpha^{k+1} \wedge \cdots \wedge \widehat{\alpha}^{k+i} \wedge \cdots \wedge \alpha^{k+l}\right) \\ &= \iota_v\omega \wedge \eta + (-1)^k \omega \wedge \iota_v\eta.\end{aligned}$$

As we mentioned, this is enough to prove that ι_v is an anti-derivation of degree -1. □

We can quite naturally extend the interior multiplication by recognising that a differential form is an element of the exterior algebra of the tangent spaces at each point. Hence, the interior multiplication can be implemented locally. To do this we should remark that we will need a vector at each point to contract with, enticing us to make use of a vector field instead of a single vector.

Definition 3.20. Let ω be a k-form and X a vector field, both on a manifold $\mathcal{M}$. We denote the **interior multiplication of ω by X** as $\iota_X\omega$ and define it for a point $p \in \mathcal{M}$ as

$$(\iota_X\omega)_p = \iota_{X_p}\omega_p.$$

Here the right-hand side is the interior multiplication on the exterior algebra as in Definition 3.17. Similar to the linear case, we call $\iota_X : \Omega^\star(\mathcal{M}) \to \Omega^{\star-1}(\mathcal{M})$ the **interior multiplication with X** or **contraction by X**. //

As this extension is done pointwise, it still holds the same properties as the interior multiplication on the exterior algebra.

Proposition 3.21. *The mapping* $\iota : \mathfrak{X}(\mathcal{M}) \times \Omega^k(\mathcal{M}) \to \Omega^{k-1}(\mathcal{M}) : (X, \omega) \mapsto \iota_X \omega$ *is* $\mathbb{R}$*-linear in both components.*

Proof. This is a direct consequence of the definition of the interior multiplication, see Definition 3.20, and Proposition 3.18. We should check that for a k-form ω the form $\iota_X \omega$ is indeed a smooth section.

Suppose that $\mathcal{M}$ is a manifold and $X \in \mathfrak{X}(\mathcal{M})$. Take some $\omega \in \Omega^k(\mathcal{M})$ and a chart $\left(U, \left(x^i\right)\right)$. We can then write ω locally as

$$\omega|_U = \sum_{i_1 \dots i_k} \frac{1}{k!} \omega_{i_1 \dots i_k} dx^i \wedge \dots \wedge dx^{i_k}, \quad \omega_{\sigma(i_1) \dots \sigma(i_k)} = \operatorname{sign}(\sigma)\, \omega_{i_1 \dots i_k}.$$

As the interior multiplication is defined pointwise, it follows that $(\iota_X \omega)|_U = \iota_X(\omega|_U)$. Hence, we can express $(\iota_X \omega)|_U$ as

$$(\iota_X \omega)|_U = \sum_{i_1 \dots i_k} \frac{1}{(k-1)!} X^{i_1} \omega_{i_1 \dots i_k} dx^{i_2} \wedge \dots \wedge dx^{i_k}, \quad \omega_{\sigma(i_1) \dots \sigma(i_k)} = \operatorname{sign}(\sigma)\, \omega_{i_1 \dots i_k}. \tag{3.4}$$

As $X \in \mathfrak{X}(\mathcal{M})$ and $\omega \in \Omega^k(\mathcal{M})$, it follows that X^{i_1} and $\omega_{i_1 \dots i_k}$ are both smooth functions. Therefore $\iota_X \omega$ has smooth coordinate functions, implying it is a smooth section of $\bigwedge^{k-1} \mathrm{T}^* \mathcal{M}$, i.e. it is a $(k-1)$-form. □

Proposition 3.22. *The interior multiplication is a graded-derivation of degree* -1*. In other words, it is a mapping* $\iota_X : \Omega^\star(\mathcal{M}) \to \Omega^{\star-1}(\mathcal{M})$ *such that*

$$\iota_X(\omega \wedge \eta) = \iota_X \omega \wedge \eta + (-1)^k \omega \wedge \iota_X \eta,$$

where $\omega \in \Omega^k(\mathcal{M})$ *and* $\eta \in \omega^l(\mathcal{M})$*.*

Proof. This all follows from the fact that the interior multiplication is defined pointwise in combination with Proposition 3.19. □

Exterior derivative

The second operator we introduce is the exterior derivative. This forms the extension of the differential on a function. The definition follows from the fact that a 1-form that is exact, meaning there exists an $f \in C^\infty(\mathcal{M})$ such that $\omega = df$, necessarily is closed, meaning $\partial \omega_j / \partial x^i - \partial \omega_i / \partial x^j = 0$ in any chart. We will first introduce the exterior derivative on a Euclidean space, after which it can easily be generalised to a manifold.

Definition 3.23. For a k-form $\omega = \sum_I \omega_I dx^I$ on $\mathbb{R}^n$ we define **the exterior derivative of** ω, denoted as $d\omega$, using the formula

$$d\omega = d\left(\sum_I \omega_I dx^I\right) = \sum_I d\omega_I \wedge dx^I,$$

where $d\omega_I$ is defined as the differential of a function. The associated operator $d : \Omega^\star(\mathcal{M}) \to \Omega^{\star+1}(\mathcal{M})$ is called the **exterior derivative on** $\mathbb{R}^n$. //

In the definition, we used the shorthand notation as it gives a clean formula. For calculations, the ordered notation will prove more useful. Translating the exterior derivative to this notation, we get the expression

$$d\omega = \sum_{i_1 < \dots < i_k} \sum_i \frac{\partial \omega_{i_1 \dots i_k}}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k}.$$

Again this operator acts as a form of derivation on $\Omega^\star(\mathbb{R}^n)$.

Proposition 3.24. *The exterior derivative on $\mathbb{R}^n$ has the following properties:*

(1) d is linear over $\mathbb{R}$.

(2) For an $f \in C^\infty(\mathbb{R}^n)$ and $X \in \mathfrak{X}(\mathbb{R}^n)$ is satisfies $df(X) = Xf$.

(3) It is a graded derivation, i.e. for any $\omega \in \Omega^k(\mathbb{R}^n)$ and $\eta \in \Omega^l(\mathbb{R}^n)$ we have

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta.$$

(4) Its square is zero, i.e. $d \circ d(\omega) = 0$ for any $\omega \in \Omega^k(\mathbb{R}^n)$.

More concisely, we might call it a graded derivation of the graded algebra $\Omega(\mathbb{R}^n)$ of degree $+1$ with a vanishing square, which coincides with the differential on functions.

Proof. The linearity of d follows from the definition, the fact that d is linear on smooth functions and that $\wedge$ is distributive over addition.

The second property follows directly from the definition. Take an arbitrary $f \in C^\infty(\mathbb{R}^n)$ and $X \in \mathfrak{X}(\mathbb{R}^n)$, a straightforward calculation of the action of df on X shows the second property.

$$df(X) = \frac{\partial f}{\partial x^i} dx^i(X) = X^i \frac{\partial f}{\partial x^i} = Xf.$$

To prove the third property, we will use the first property much like we did in the proof of Proposition 3.21. Therefore, we only consider terms of the form $\omega = \omega_I dx^I \in \Omega^k(\mathbb{R}^n)$ and $\eta = \eta_J dx^J \in \Omega^l(\mathbb{R}^n)$. By the fact that the exterior derivative is the differential on functions and therefore satisfies the Leibniz rule, it follows that,

$$\begin{aligned} d(\omega \wedge \eta) &= d\left(\omega_I \eta_J dx^{IJ}\right) = d(\omega_I \eta_J) dx^{IJ} = ((d\omega_I)\eta_J + \omega_I(d\eta_J)) \wedge dx^{IJ}, \\ &= \left(d\omega_I \wedge dx^I\right) \wedge \eta_J dx^J - (-1)^k \omega_I dx^I \wedge \left(d\eta_J \wedge dx^J\right) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta. \end{aligned}$$

As we mentioned, by the $\mathbb{R}$-linearity of d this implies that d satisfies the product rule with respect to $\wedge$.

To prove the last property, we will first consider the action of the exterior derivative on a function. For an arbitrary $f \in \Omega^0(\mathbb{R}^n) = C^\infty(\mathbb{R}^n)$ we use the fact that the second order partial differential is symmetric such that

$$\begin{aligned} d(df) = d\left(\frac{\partial f}{\partial x^i} dx^i\right) &= d\left(\frac{\partial f}{\partial x^i}\right) \wedge dx^i = \frac{\partial^2 f}{\partial x^i \partial x^j} dx^j \wedge dx^i, \\ &= \sum_{i<j} \left(\frac{\partial^2 f}{\partial x^i \partial x^j} - \frac{\partial^2 f}{\partial x^j \partial x^i}\right) dx^i \wedge dx^j = 0. \end{aligned}$$

Hence $d \circ d = 0$ on functions, luckily we can reduce the case on arbitrary k-forms to that of functions as follows using

$$\begin{aligned} d(d\omega) &= d\left(d\omega_{i_1 \ldots i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right), \\ &= d(d\omega_{i_1 \ldots i_k}) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} + d\omega_{i_1 \ldots i_k} \wedge d\left(dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right). \end{aligned}$$

The first term is zero as we showed that $d \circ d$ is zero on functions. The second term can also be shown to be zero by using the following inductive argument

$$d\left(dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right) = d\left(dx^{i_1}\right) \wedge \cdots \wedge dx^{i_k} - dx^{i_1} \wedge d\left(dx^{i_2} \wedge \cdots \wedge dx^{i_k}\right).$$

The first term is zero by the same argument as before, the second term is zero due to the inductive argument. Hence, we have shown that $d(d\omega) = 0$ for an arbitrary ω. □

Besides these algebraic properties, the exterior derivative also works naturally with smooth functions between open subsets.

Proposition 3.25. *For any $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ both open and function $F : V \to U$ we have*

$$F^* (d\omega) = d (F^* \omega),$$

where ω is an arbitrary k-form on V.

Proof. We can use the linearity of the exterior derivative, see Proposition 3.24(1), such that it suffices to show this for $\omega = \omega_{i_1 \dots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$. Suppose that $F : V \to U$, where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ and $\omega \in \Omega^k (U)$, then

$$\begin{aligned} F^* (d\omega) &= d (\omega_{i_1 \dots i_k} \circ F) \wedge d \left(x^{i_1} \circ F\right) \wedge \cdots \wedge d \left(x^{i_k} \circ F\right), \\ &= d \left((\omega_{i_1 \dots i_k} \circ F) \wedge d \left(x^{i_1} \circ F\right) \wedge \cdots \wedge d \left(x^{i_k} \circ F\right)\right) = d (F^* \omega). \end{aligned}$$

This proves our result. □

We can use the properties of the exterior derivative on $\mathbb{R}^n$, see Propositions 3.24 and 3.25, to extend our definition to an arbitrary manifold.

Proposition 3.26. *For a manifold $\mathcal{M}$, there exists a unique $\mathbb{R}$-linear operator $d : \Omega^k (\mathcal{M}) \to \Omega^{k+1} (\mathcal{M})$ for any k which satisfies the following:*

(1) For a function $f \in C^\infty (\mathcal{M})$ it is defined as the differential.

(2) It is a graded derivation, i.e. for any $\omega \in \Omega^k (\mathcal{M})$ and $\eta \in \omega^l (\mathcal{M})$ we have

$$d (\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta.$$

(3) Its square is zero, i.e. $d \circ d (\omega) = 0$ for any $\omega \in \Omega^k (\mathcal{M})$.

Proof. Much like any unique existence theorem, we will first prove the existence and then prove that this is also unique. The existence can be derived locally from the definition of the exterior derivative on $\mathbb{R}^n$. The uniqueness is a consequence of the product rule.

Suppose that $\mathcal{M}$ is a manifold and $\omega \in \Omega^k (\mathcal{M})$ for some arbitrary k, then let (U, ϕ) be a chart on $\mathcal{M}$. We define $d\omega$ locally on U as

$$(d\omega) |_U = \left(\phi^* \circ d \circ \left(\phi^{-1}\right)^*\right) (\omega|_U).$$

Remark that the d on the right-hand side is the exterior derivative on $\mathbb{R}^n$.

As this defines the exterior derivative in a specific chart, we should now check the value does not depend on our choice of chart. Let (V, ψ) be another chart on $\mathcal{M}$, we will then calculate $d\omega|_{U \cap V}$ and show that it is invariant under coordinate transformations by using Proposition 3.25 on $\phi \circ \psi^{-1} : \psi (U \cap V) \to \phi (U \cap V)$,

$$\begin{aligned} \left(\phi^* \circ d \circ \left(\phi^{-1}\right)^*\right) (\omega|_{U \cap V}) &= \left(\psi^* \circ \left(\psi^{-1}\right)^* \circ \phi^* \circ d \circ \left(\phi^{-1}\right)^*\right) (\omega|_{U \cap V}), \\ &= \left(\psi^* \circ d \circ \left(\psi^{-1}\right)^* \circ \phi^* \circ \left(\phi^{-1}\right)^*\right) (\omega|_{U \cap V}), \\ &= \left(\psi^* \circ d \circ \left(\psi^{-1}\right)^*\right) (\omega|_{U \cap V}). \end{aligned}$$

Thus we have shown that the exterior derivative is independent of the choice of a coordinate chart, which implies that it is well-defined globally. It is then clear from Proposition 3.24 that it satisfies the properties mentioned locally and by construction everywhere.

For uniqueness, we will first show that any operator which is a graded derivation is a local operator, i.e. if $\eta \in \Omega^k (\mathcal{M})$ satisfy $\eta|_U = 0$ for some open $U \subset \mathcal{M}$, then $d\eta|_U = 0$. This then lets us prove the uniqueness locally.

Suppose that D is a graded derivation on $\Omega^k (\mathcal{M})$, and let η be some k-form for which there exists an open $U \subset \mathcal{M}$ such that $\eta|_U = 0$. Take a $p \in U$ and take a bump function ψ such that there exists an open neighbourhood V of p for which $\psi|_V = 1$ and $\operatorname{supp} (\psi) \subset U$. We can

then conclude from the fact that $\eta|_{\text{supp}(\psi)} = 0$ that $\psi\eta = 0$. Hence, it follows with $\eta_p = 0$ and $\psi(p) = 1$ that

$$0 = D(\psi\eta)_p = (D\psi)_p \wedge \eta_p + \psi(p) \wedge (D\eta) = (D\eta)_p .$$

This shows that $D\eta|_U = 0$ if $\eta|_U = 0$, and hence D is a local operator. Remark that this specifically applies to the exterior derivative.

Now suppose that $D : \Omega^k(\mathcal{M}) \to \Omega^{k+1}(\mathcal{M})$ is an operator which satisfies the properties of the proposition. We will then show that D coincides with d by first proving this for smooth functions, and then showing that D vanishes on products of differentials. It should be clear that for any $f \in C^\infty(\mathcal{M})$ we have $Df = df$ as they are both equal to the differential of f. Suppose that $f^1, \ldots, f^k \in C^\infty(\mathcal{M})$, it follows from the properties of the proposition that

$$D\left(df^1 \wedge \cdots \wedge df^k\right) = D\left(Df^1 \wedge \cdots \wedge Df^k\right) = \sum_{i=1}^{k} (-1)^{i-1} Df^1 \wedge \cdots \wedge D^2 f^i \wedge \cdots \wedge Df^k = 0.$$

Take some arbitrary $\eta \in \Omega^k(\mathcal{M})$ and $p \in \mathcal{M}$, and let $\left(U, \left(x^i\right)\right)$ be some coordinates around p. Construct some smooth bump function ψ such that there exists an open neighbourhood V of p such that $\psi|_V = 1$ and $\text{supp}(\psi) \subset U$. Furthermore, in these coordinates we can write η as $\eta|_U = \sum_I \eta_I dx^I$. Extend the functions x^i and η_I smoothly using the smooth bump function to functions $\tilde{x}^i$ and $\tilde{\eta}_I$ such that $\tilde{x}^i|_V = x^i|_V$ and $\tilde{\eta}_I|_V = \eta_I|_V$. We can then define an extension of η as $\tilde{\eta} = \sum_I \tilde{\eta}_I d\tilde{x}^I$ such that it is defined on the whole of $\mathcal{M}$ and satisfies $\tilde{\eta}|_V = \eta|_V$. By the locality of D and d it follows that $(D\eta)|_V = (D\tilde{\eta})|_V$ and $(d\eta)|_V = (d\tilde{\eta})|_V$. As $p \in V$, we can conclude from the discussion above that

$$\begin{aligned}(D\eta)_p = (D\tilde{\eta})_p &= \left(D\sum_I \tilde{\eta}_I d\tilde{x}^I\right)_p = \left(\sum_I D\tilde{\eta}_I \wedge d\tilde{x}^I + \tilde{\omega}_I \wedge D d\tilde{x}^I\right)_p \\ &= \left(\sum_i d\tilde{\eta}_I \wedge d\tilde{x}^I\right)_p = (d\tilde{\eta})_p = (d\eta)_p .\end{aligned}$$

As our choice of p is arbitrary, it follows that $D = d$ on the whole manifold and thus that d is uniquely defined. □

Definition 3.27. The unique operator d defined in Proposition 3.26 is called the **exterior derivative on** $\mathcal{M}$. It is the unique linear operator that is a graded derivation on $\Omega(\mathcal{M})$ of degree $+1$ with a vanishing square that coincides with the differential on smooth functions. //

The exterior derivative lets us define special classes of differential forms, namely closed and exact differential forms. These play a much bigger role in the study of the de Rham cohomology groups.

Definition 3.28. A differential form $\omega \in \Omega^k(\mathcal{M})$ is called **closed** if $d\omega = 0$, and **exact** if there exists an $\alpha \in \Omega^{k-1}(\mathcal{M})$ such that $\eta = d\alpha$. Let $Z^k(\mathcal{M})$ denoted the set of closed k-forms on $\mathcal{M}$ and $B^k(\mathcal{M})$ the set of exact k-forms, remark that $Z^k(\mathcal{M}) = \ker\left(d : \Omega^k(\mathcal{M}) \to \Omega^{k+1}(\mathcal{M})\right)$ and $B^k(\mathcal{M}) = \text{Im}\left(d : \Omega^{k-1}(\mathcal{M}) \to \Omega^k(\mathcal{M})\right)$. //

We will now show a similar statement to Proposition 3.25, proving that the exterior derivative on an arbitrary manifold is natural in the sense that it commutes with the pullback.

Proposition 3.29. *For a smooth function $F : \mathcal{M} \to \mathcal{N}$ and an arbitrary $\omega \in \Omega^k(\mathcal{N})$, we have that*

$$F^*(d\omega) = d(F^*\omega) .$$

In other words, the exterior derivative commutes with the pullback.

Proof. This follows directly from applying Proposition 3.25 to the coordinate representations of F and ω. □

A last important property of the differential is the fact that it can commute with integrals in a special manner that is reminiscent of the Leibniz integral rule.

Proposition 3.30. *Given a smooth family of differential forms $\omega_t \in \Omega^k(\mathcal{M})$, with $t \in [0,1]$, it satisfies the following*

$$\int_0^1 (d\omega_t)\, dt = d\left(\int_0^1 \omega_t dt\right).$$

where d is the exterior derivative on $\mathcal{M}$.

Proof. We can assume that $\mathcal{M}$ has some global coordinates $\left(x^i\right)$. Next up, suppose that $\omega_t \in \Omega^k(\mathcal{M})$ is a smooth family of k-forms with $t \in [0,1]$. We can then use the fact that partial derivatives commute with integrals of constant boundaries.

$$\begin{aligned}
\int_0^1 (d\omega_t)\, dt &= \int_0^1 \left(d\left((\omega_t)_I\, dx^I\right)\right) dt = \int_0^1 \left(\frac{\partial\,(\omega_t)_I}{\partial x^i} dx^i \wedge dx^I\right) dt, \\
&= \left(\int_0^1 \frac{\partial\,(\omega_t)_I}{\partial x^i} dt\right) dx^i \wedge dx^I = \frac{\partial}{\partial x^i}\left(\int_0^1 (\omega_t)_I\, dt\right) dx^i \wedge dx^I, \\
&= d\left(\int_0^1 (\omega_t)_I\, dt \wedge dx^I\right) = d\left(\int_0^1 \omega_t dt\right).
\end{aligned}$$

Which was what we wanted. □

Lie derivative

The last operator we will introduce in this section is the Lie derivative of a differential form. This forms a natural extension of the Lie derivative on functions and vector fields, see Definitions A.24 and A.26.

Definition 3.31. The **Lie derivative** of a differential form $\omega \in \Omega^k(\mathcal{M})$ along some $X \in \mathfrak{X}(\mathcal{M})$ is defined as

$$\mathcal{L}_X \omega = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\phi_X^t\right)^* \omega.$$

The existence of the derivative is ensured by Theorem A.23. We call the operator $\mathcal{L}_X : \Omega^\star(\mathcal{M}) \to \Omega^\star(\mathcal{M})$ the **Lie derivative along** X. //

As the name implies, this operator is a derivation on $\Omega(\mathcal{M})$.

Proposition 3.32. *The Lie derivative is a derivation of degree 0. In other words, for any $X \in \mathfrak{X}(\mathcal{M})$ it is a mapping $\mathcal{L}_X : \Omega^\star(\mathcal{M}) \to \Omega^\star(\mathcal{M})$ such that*

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

where $\omega \in \Omega^k(\mathcal{M})$ and $\eta \in \Omega^l(\mathcal{M})$.

Proof. Let X be a vector field on $\mathcal{M}$, ω a k-form and η an l-form. Remark that the pullback distributes over the wedge product, see [13, Lemma 14.16 (b)], i.e.

$$\phi_X^{t\,*}(\omega \wedge \eta) = \left(\phi_X^{t\,*}\omega\right) \wedge \left(\phi_X^{t\,*}\eta\right).$$

Hence, we get

$$\begin{aligned}
\mathcal{L}_X(\omega \wedge \eta) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left[\left(\phi_X^t\right)^*(\omega \wedge \eta)\right] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left[\left(\phi_X^t\right)^*\omega \wedge \left(\phi_X^t\right)^*\eta\right], \\
&= \left[\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\phi_X^t\right)^*\omega\right] \wedge \eta + \omega \wedge \left[\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\phi_X^t\right)^*\eta\right] = \mathcal{L}_X\omega \wedge \eta + \omega \wedge \mathcal{L}_X\eta.
\end{aligned}$$

This proves the proposition. □

Now we can already make a simple connection between the Lie derivative and the exterior derivative.

Proposition 3.33. *The exterior derivative commutes with the Lie derivative.*

Proof. Suppose X is some vector field on $\mathcal{M}$ and $\omega \in \Omega^k(\mathcal{M})$. Using Proposition 3.29, we obtain our result

$$(\mathcal{L}_X \circ d)\,\omega = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\phi_X^t\right)^* (d\omega) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} d\left(\left(\phi_X^t\right)^* \omega\right) = (d \circ \mathcal{L}_X)\,\omega.$$

As ω is arbitrary, we get $\mathcal{L}_X \circ d = d \circ \mathcal{L}_X$. □

Besides this connection to the exterior derivative, the Lie derivative lets us have a natural connection with the flow along a vector field.

Proposition 3.34. *Let X be a vector field on a manifold $\mathcal{M}$. If ϕ_X^t denotes the flow of X at time t, then for any $\alpha \in \Omega(\mathcal{M})$*

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} \left(\phi_X^t\right)^* \alpha = \left(\phi_X^{t_0}\right)^* (\mathcal{L}_X \alpha).$$

Proof. Let X be a vector field on a manifold $\mathcal{M}$ and $\alpha \in \Omega(\mathcal{M})$, we can write out the definitions above for a point $p \in \mathcal{M}$, where we need to ensure that $(t_0, p) \in \mathcal{D}(X)$. It then follows by substituting $t = s + t_0$ and using Proposition A.22 and Proposition 12.25 (e) in [13],

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} \left(\phi_X^t\right)^* \alpha &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \left(\phi_X^{s+t_0}\right)^* \alpha = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \left(\phi_X^{t_0}\right)^* (\phi_X^s)^* \alpha, \\ &= \left(\left(\phi_X^{t_0}\right)^*\right)^* \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} (\phi_X^s)^* \alpha = \left(\phi_X^{t_0}\right)^* \mathcal{L}_X \alpha.\end{aligned}$$

Which was what we wanted. □

Cartan's magic formula

The combination of these three operators comes in the form of Cartan's magic formula. It tells us that the Lie derivative can be expressed in terms of the interior multiplication and exterior derivative. Concretely this is stated as follows.

Proposition 3.35. *The Lie derivative of along a vector field X is given by*

$$\mathcal{L}_X = d \circ \iota_X + \iota_X \circ d.$$

Before we prove this statement, we will go into some properties of the **commutator of d and ι_X**, given by $D_X = d \circ \iota_X + \iota_X \circ d$.

Lemma 3.36. *For a vector field $X \in \mathfrak{X}(\mathcal{M})$ and function $f \in C^\infty(\mathcal{M})$, the action of D_X is equal to the Lie derivative.*

Proof. Let X be a vector field on $\mathcal{M}$ and f a function. By using the definition of the interior multiplication, Proposition 3.26.(1), and Corollary A.25, it follows that

$$D_X f = (d \circ \iota_X + \iota_X \circ d) f = d(\iota_X f) + \iota_X (df) = df(X) = Xf = \mathcal{L}_X f.$$

Hence, this shows that $D_X f = \mathcal{L}_X f$. □

Lemma 3.37. *The operator D_X commutes with d, i.e. $D_X \circ d = d \circ D_X$.*

Proof. This proof simply follows from the fact that the exterior derivative's square is zero, see Proposition 3.26.(3). We can see that

$$D_X \circ d = (d \circ \iota_X + \iota_X \circ d) \circ d = d \circ \iota_X \circ d = d \circ (d \circ \iota_X + \iota_X \circ d) = d \circ D_X.$$

Therefore, d and D_X commute. □

Lemma 3.38. *For a vector field $X \in \mathfrak{X}(\mathcal{M})$, the operator D_X is a derivation on $\Omega(\mathcal{M})$ of degree 0.*

Proof. Suppose that $\mathcal{M}$ is a manifold, $X \in \mathfrak{X}(\mathcal{M})$. We can easily verify that $D_X : \Omega^k(\mathcal{M}) \to \Omega^k(\mathcal{M})$ which implies it has degree 0.

To prove that it is a derivation, we will make use of Propositions 3.19 and 3.26.(2). Suppose that $\omega \in \Omega^k(\mathcal{M})$ and $\eta \in \Omega^l(\mathcal{M})$, then

$$\begin{aligned} D_X(\omega \wedge \eta) &= (d \circ \iota_X + \iota_X \circ d)(\omega \wedge \eta) \\ &= d\left(\iota_X\omega \wedge \eta + (-1)^k \omega \wedge \iota_X\eta\right) + \iota_X\left(d\omega \wedge \eta + (-1)^k \omega \wedge d\eta\right) \\ &= d\iota_X\omega \wedge \eta + (-1)^{k-1}\iota_X\omega \wedge d\eta + (-1)^k d\omega \wedge \iota_X\eta + (-1)^{2k}\omega \wedge d\iota_X\eta \\ &\quad + \iota_X d\omega \wedge \eta + (-1)^{k+1} d\omega \wedge \iota_X\eta + (-1)^k \iota_X\omega \wedge d\eta + (-1)^{2k}\omega \wedge \iota_X d\eta \\ &= (d \circ \iota_X + \iota_X \circ d)\omega \wedge \eta + \omega \wedge (d \circ \iota_X + \iota_X \circ d)\eta = D_X\omega \wedge \eta + \omega \wedge D_X\eta. \end{aligned}$$

This shows that, $D_X(\omega \wedge \eta) = D_X\omega \wedge \eta + \omega \wedge D_X\eta$ and it is therefore a derivation of degree 0. □

We will now apply Lemmas 3.36 and 3.38 to prove Proposition 3.35.

Proof of Proposition 3.35. As all the operators satisfy the product rule, we can deduce that they are local, see the proof of Proposition 3.26. If we combine this with the linearity of the operators, we remark that we only need to consider a k-form ω which we can write in a coordinate chart $\left(U, \left(x^i\right)\right)$ as $\omega|_U = \omega_{i_1,\ldots,i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$. Then we can split this k-form into two parts $\omega|_U = \alpha \wedge \eta$ with $\alpha = dx^{i_1}$ and $\beta = \omega_{i_1\ldots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$. Then remark that ω is some wedge product of an exact 1-form and a $k-1$-form. By using the derivation properties of the commutator and the Lie derivative, see Lemma 3.38 and Proposition 3.32, we can show the equivalence on a k-form by proving the equivalence on exact 1-forms and using induction. To prove the equivalence on exact 1-form, we will use Lemmas 3.36 and 3.37 and Proposition 3.33.

$$D_X(df) = d(D_X f) = d(\mathcal{L}_X f) = \mathcal{L}_X(df).$$

This is then enough to prove Cartan's magic formula as we know that they coincide on 0-forms, see Lemma 3.36. □

Poincaré's lemma

Let us now apply Cartan's magic formula to prove a basic statement on differential forms called Poincaré's lemma. This tells us that any closed differential form is locally also exact. We will first prove this for differential forms on open subsets of $\mathbb{R}^n$, which naturally extends to the mentioned statement. Our proof consists of defining an operator that almost inverts the exterior derivative, called a homotopy operator.

Lemma 3.39. *Every closed k-form with $k \geq 1$ on a star-shaped domain of $\mathbb{R}^n$ is exact.*

Proof. Suppose that V is a star-shaped domain of $\mathbb{R}^n$ and let ω be a k-form on V with $k \geq 1$. Without loss of generality, we can assume that V is star-shaped around 0. We want to construct an operator $h : \Omega^k(V) \to \Omega^{k-1}(V)$ such that $d \circ h = \mathrm{Id}$. This is impossible in general as it would imply that every differential form is exact. Therefore we generalize our formula to $d \circ h + h \circ d = \mathrm{Id}$, such that it is equivalent to $d \circ h = \mathrm{Id}$ for closed forms.

This operator h is defined using the contraction mapping $m_t : V \to V : x \mapsto tx$, which is a well-defined function as long as $t \in [0,1]$ due to the star-shapedness of V. Remark that $m_0 = 0$ and $m_1 = \mathrm{Id}_V$, therefore ${m_0}^* = 0$ and ${m_1}^* = \mathrm{Id}_{\Omega(V)}$ as well. Define $h : \Omega^k(V) \to \Omega^{k-1}(V)$ as follows

$$h(\omega) = \int_0^1 \frac{1}{t} {m_t}^*(\iota_X\omega)\, dt.$$

where $X = x^i \, \partial/\partial x^i$ such that $m_t = \phi_X^{\ln(t)}$. Remark that this integral is well-defined even though we divide by t. This can be seen by writing the integrand out in coordinates, where we assume $\omega = \omega_{i_1 \dots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ in this calculation as every operation is linear. By noticing that m_t is simply the multiplication by t in each coordinate and the fact that d is $\mathbb{R}$ linear, we can calculate the action of the pullback in coordinates, see Lemma 14.16 (c) in [13], we see

$$\begin{aligned}
&\frac{1}{t} {m_t}^* \left(\iota_X \omega\right) \\
&= \frac{1}{t} {m_t}^* \left(\iota_X \omega_{i_1 \dots i_k} \, dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right) \\
&= \frac{1}{t} {m_t}^* \left(\sum_{j=1}^{k} (-1)^{j-1} \, dx^{i_j} \left(X\right) \omega_{i_1 \dots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_{j-1}} \wedge dx^{i_{j+}} \wedge \cdots \wedge dx^{i_k} \right) \\
&= \frac{1}{t} \sum_{j=1}^{k} (-1)^{j-1} \, X\left(tx^{i_j}\right) \left(\omega_{i_1 \dots i_k} \circ m_t\right) d\left(tx^{i_1}\right) \wedge \cdots \wedge d\left(tx^{i_{j-1}}\right) \wedge d\left(tx^{i_{j+1}}\right) \wedge \cdots \wedge d\left(tx^{i_k}\right) \\
&= t^{k-1} \sum_{j=1}^{k} (-1)^{j-1} \, X\left(x^{i_j}\right) \left(\omega_{i_1 \dots i_k} \circ m_t\right) dx^{i_1} \wedge \cdots \wedge dx^{i_{j-1}} \wedge dx^{i_{j+1}} \wedge \cdots \wedge dx^{i_k}.
\end{aligned}$$

This shows that the integrand is well-defined for $t \in [0, 1]$ if $k \geq 1$.

All that is left to show, is that this operator satisfies the equation

$$d \circ h + h \circ d = \mathrm{Id}\,.$$

We can prove this by working out the calculation for some ω. In this calculation, we will use the fact that the integral is a linear operator and that the exterior derivative commutes with the integral and the pullback, see Propositions 3.30 and 3.29

$$\begin{aligned}
\left(d \circ h + h \circ d\right) \omega &= d \int_0^1 \frac{1}{t} {m_t}^* \left(\iota_X \omega\right) dt + h\left(d\omega\right), \\
&= \int_0^1 \frac{1}{t} {m_t}^* \left(d\left(\iota_X \omega\right)\right) dt + \int_0^1 \frac{1}{t} {m_t}^* \left(\iota_X \left(d\omega\right)\right) dt, \\
&= \int_0^1 \frac{1}{t} {m_t}^* \left(\left(d \circ \iota_X + \iota_X \circ d\right) \omega\right) dt.
\end{aligned}$$

Now we use Cartan's magic formula to rewrite this in terms of the Lie derivative. Furthermore, we recognise m_t as the flow of X and use the commutation property of the Lie derivative. Then we obtain our results from the chain rule.

$$\begin{aligned}
\left(d \circ h + h \circ d\right) \omega &= \int_0^1 \frac{1}{t} {m_t}^* \left(\mathcal{L}_X \omega\right) dt = \int_0^1 \frac{1}{t} \left(\phi_X^{\ln(t)}\right)^* \left(\mathcal{L}_X \omega\right) dt, \\
&= \int_0^1 \frac{1}{t} \left. \frac{\mathrm{d}}{\mathrm{d}s} \right|_{s=\ln(t)} \left(\left(\phi_X^s\right)^* \omega\right) dt, \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \left(\left(\phi_X^{\ln(t)}\right)^* \omega\right) dt = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \left({m_t}^* \omega\right) dt = {m_1}^* \omega - {m_0}^* \omega = \omega.
\end{aligned}$$

Hence, the operator h as defined above is exactly the operator we were looking for. By using the fact that ω is closed, we can deduce that $d\left(h\left(\omega\right)\right) = \omega$ and thus ω is exact on V. □

Corollary 3.40. *For every closed differential form ω and point p on the manifold, there exists a neighbourhood of p on which ω is exact.*

Proof. Let $\mathcal{M}$ be a n-manifold and suppose that ω is a closed k-form on $\mathcal{M}$ and let p be a point in $\mathcal{M}$. Take a chart (U, ϕ), such that $\phi : U \to V \subset \mathbb{R}^n$ is a diffeomorphism. We can assume that

V is star-shaped because if it is not we can take an open sphere $\tilde{V}$ contained in V as V is open, the diffeomorphism $\phi|_{\phi^{-1}(\tilde{V})}$ then gives a new chart.

We would like to use Poincaré's lemma now, however, this lemma only considers differential forms on $\mathbb{R}^n$. Therefore, we want to use the diffeomorphism of the chart to pullback the differential form to a differential form on V.

To be able to do this use that the exterior derivative commutes with the pullback such that $\left(\phi^{-1}\right)^* \omega\big|_U$ a closed k-form on V. As we assumed that V is star-shaped we can use Poincaré's lemma implying the existence of a $(k-1)$-form η on V with the property that $\left(\phi^{-1}\right)^* \omega\big|_U = d\eta$. By again taking the pullback by ϕ, and using the fact that $\left(\phi^{-1} \circ \phi\right)^* = \mathrm{Id}_U$, it follows that

$$\omega\big|_U = \left(\phi^{-1} \circ \phi\right)^* \omega\big|_U = \phi^* \left(\left(\phi^{-1}\right)^* \omega\big|_U\right) = \phi^* \left(d\eta\right) = d\left(\phi^* \eta\right).$$

This implies that ω is exact on U. □

3.3.2 Darboux's Theorem

We are now able to prove Theorem 3.10. However, much like the Riemannian metric, we can also give a short intuitive proof of this statement akin to Riemann's counting argument given above. It now deviates as we can use Corollary 3.40. When given a symplectic form ω it can be written is coordinates $\left(x^i\right)$ as $\omega = d\theta = d\left(\theta_i dx^i\right)$. This then leaves just n function we need to determine, however with our n choices for coordinates we have these exactly covered and there are no local invariants for a symplectic form.

Now we will give a formal proof, for which will restate the theorem once more. Then we will go into the proof where we use linear symplectic geometry and extend its structure locally using Moser's trick, Poincaré's lemma and Cartan's magic formula.

Theorem 3.10. *Around every point on a symplectic manifold $(\mathcal{M}, \omega)$, there exists a neighbourhood $\left(U, \left(x^1, \ldots, x^n, y^1, \ldots, y^n\right)\right)$ such that on U the symplectic form is given by*

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

Such coordinates are called Darboux coordinates.

Proof. To prove this theorem for a point p on a symplectic manifold $(\mathcal{M}, \omega)$ we will do the following. Firstly, we will use Theorem 2.8 to generate a basis for the tangent space at this point. We will then extend this basis to a neighbourhood of this point, generating a chart. Subsequently, we will show that part of this neighbourhood can be pulled back to the standard symplectic form in these coordinates. This will then define coordinates that satisfy Darboux's theorem.

Suppose that $(\mathcal{M}, \omega)$ is a symplectic m-manifold and $p \in \mathcal{M}$. By theorem 2.8 there exists a symplectic basis of $\mathrm{T}_p\mathcal{M}$, denoted as $\{u_1, \ldots, u_n, v_1, \ldots, v_n\}$. This also implies that $m = 2n$ such that $\mathcal{M}$ is even dimensional. Furthermore, the symplectic form ω_p can be written in the associated dual basis $\left\{\mu^1, \ldots, \mu^n, \nu^1, \ldots, \nu^n\right\}$ as

$$\omega_p = \sum_{i=1}^{n} \mu^i \wedge \nu^i.$$

We will try to extend this basis to a coordinate chart $\left(x^1, \ldots, x^n, y^1, \ldots, y^n\right)$ such that $\mu^i = dx^i$ and $\nu^i = dy^i$. These can be found rather easily by taking an arbitrary coordinate chart around p, $\left(U, \left(\tilde{x}^1, \ldots, \tilde{x}^n, \tilde{y}^1, \ldots, \tilde{y}^n\right)\right)$, and then remarking that both $\left(\mu^1, \ldots, \mu^n, \nu^1, \ldots, \nu^n\right)$ and $\left(d\tilde{x}^1_p, \ldots, d\tilde{x}^n_p, d\tilde{y}^1_p, \ldots, d\tilde{y}^n_p\right)$ are both basis for $\mathrm{T}^*_p\mathcal{M}$. Hence, there exists a basis transformation, which is linear and non-singular, given by some matrices A, B, C and D

$$\mu^i = A^i_j d\tilde{x}^j_p + B^i_j d\tilde{y}^j_p, \quad \nu^i = C^i_j d\tilde{x}^j_p + D^i_j d\tilde{y}^j_p.$$

Then define now coordinates $\left(x^1,\ldots,x^n,y^1,\ldots,y^n\right)$ as follows

$$x^i = A^i_j\tilde{x}^j + B^i_j\tilde{y}^j, \quad y^i = C^i_j\tilde{x}^j + D^i_j\tilde{y}^j.$$

These form a coordinate chart, as they are the composition of a diffeomorphism and a non-singular linear map. Furthermore, notice that by definition $\mu^i = dx^i_p$ and $\nu^i = dy^i_p$. Thus we have found a coordinate chart $\left(U,\phi=\left(x^1,\ldots,x^n,y^1,\ldots,y^n\right)\right)$ that satisfies

$$\omega_p = \sum_{i=1}^{n} dx^i_p \wedge dy^i_p.$$

We now want to compare ω with the standard symplectic form $\omega_1 = \sum_{i=1}^n dx^i \wedge dy^i$. To do this, we will have to restrict to some smaller neighbourhood of p than U. We will construct a symplectomorphism ψ between (V,ω) and (V,ω_1) such that $\psi(p) = p$. Let us set $\omega_0 = \omega$ in this process.

To construct such a symplectomorphism we use Moser's trick. This comes down to constructing an isotopy, which is a map $\rho : U\times[0,1]\to U : (p,t)\mapsto \rho_t(p)$ such that ρ_t is a diffeomorphism on U and $\rho_0 = \mathrm{Id}_U$. We will give a short exposition to motivate this construction. Suppose that we are given an isotopy ρ_t and a family of symplectic forms ω_t such that ${\rho_t}^*\omega_t = \omega_0$. Define some vector field X_t generated by ρ_t as

$$X_t = \frac{\mathrm{d}\rho_t}{\mathrm{d}t}\circ\rho_t^{-1}.$$

Let us now consider the derivative of the pullback of ω_t by ρ_t. We can rewrite this using Proposition 3.34 and Cartan's magic formula.

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=t}({\rho_s}^*\omega_s) = {\rho_t}^*\left(\mathcal{L}_{X_t}\omega_t + \frac{\mathrm{d}\omega_t}{\mathrm{d}t}\right) = {\rho_t}^*\left(d\left(\iota_{X_t}\omega_t\right) + \frac{\mathrm{d}\omega_t}{\mathrm{d}t}\right).$$

As ρ_t is a diffeomorphism, we deduce the following equation called Moser's equation:

$$\left(d\circ\iota_{X_t}\right)\omega_t = -\frac{\mathrm{d}\omega_t}{\mathrm{d}t}.$$

Hence, an isotopy generates a vector field which satisfies Moser's equation. Remark that we could also recover the isotopy if we were given a vector field that satisfies Moser's equation. Hence, to generate an isotopy, we will try to solve Moser's equation.

Consider the difference $\eta = \omega_1 - \omega_0$ and remark that this is a closed 2-form. Therefore, by Corollary 3.40 we can find a neighbourhood U_0 of p such that η is exact, i.e. there exists a 1-form α such that $\eta = -d\alpha$ and we can assume that $\alpha_p = 0$ using the linearity of the exterior derivative. Furthermore, we can assume that $U_0\subset U$. We will then consider the family of symplectic forms ω_t with $t\in[0,1]$ given by

$$\omega_t = \omega_0 - td\alpha = (1-t)\,\omega_0 + t\omega_1.$$

This is again closed and as $\omega_t|_p$ is non-degenerate and ω_t is smooth it follows that ω_t is non-degenerate in a neighbourhood U_1 of p. Again, we can assume that $U_1\subset U_0$. Thus we have found a family of symplectic forms that we can enter into Moser's equations to get the following

$$\left(d\circ\iota_{X_t}\right)\omega_t = -\frac{\mathrm{d}\omega_t}{\mathrm{d}t} = \omega_0 - \omega_1 = -\eta = d\alpha.$$

To solve this equation, it is sufficient to solve for $\iota_{X_t}\omega_t = \alpha$. By the non-degeneracy of ω_t, this equation can be solved locally. As for an arbitrary $Y\in\mathfrak{X}(\mathcal{M})$ we can calculate this expression in the coordinates $\left(z^1,\ldots,z^n,z^{n+1},\ldots,z^{2n}\right) = \left(x^1,\ldots,x^n,y^1,\ldots,y^n\right)$. The left-hand side then becomes

$$\begin{aligned}\iota_{X_t}\omega_t(Y) = \omega_t(X_t,Y) &= \sum_{i,j}\frac{1}{2}(\omega_t)_{ij}\,dz^i\wedge dz^j(X_t,Y) = \sum_{i,j}\frac{1}{2}(\omega_t)_{ij}\left(X^i_tY^j - X^j_tY^i\right),\\ &= \sum_{i,j}\frac{1}{2}\left((\omega_t)_{ij} - (\omega_t)_{ji}\right)X^i_tY^j = \sum_{i,j}(\omega_t)_{ij}\,X^i_tY^j. \qquad (3.5)\end{aligned}$$

Meanwhile, the right-hand side can be expressed as

$$\alpha(Y) = \sum_j \alpha_j dz^j(Y) = \sum_j \alpha_j Y^j. \tag{3.6}$$

As Y is arbitrary, we can assume that there is some $1 \leq k \leq 2n$ such that $Y = \partial/\partial z^k$. Combining Equations 3.5 and 3.6 then gives us that $\sum_i (\omega_t)_{ik} X_t^i = \alpha_k$. As ω_t is non-degenerate, $(\omega_t)_{ik}$ has an inverse, which we denote as $(\omega_t)^{ik}$. We can then express X_t in terms of ω_t and α

$$X_t^i = \sum_k (\omega_t)^{ik} \alpha_k.$$

From the smoothness of ω_t and α, and the fact that $(\omega_t)^{ik}$ is a rational function of the coefficients of $(\omega_t)_{ik}$ it follows that X_t^i is smooth as well. Thus we can smoothly solve for X_t.

Remark that $\alpha_p = 0$ and therefore $X(t,p) = 0$ which implies that $\phi_X^{t,0}(p) = p$ for all $t \in [0,1]$. As the flow domain of is open, we can use the tube lemma, see [16, Lemma 26.8] to find a neighbourhood U_2 of p such that $\phi_X^{t,0}$ is defined on U_2 for all $t \in [0,1]$, where we can once again assume $U_2 \subset U_1$. Then by defining $\rho_t = \phi_X^{t,0}$, we get the isotopy we wanted as $\rho_0 = \phi_X^{0,0} = \mathrm{Id}_{U_2}$, and we have already shown that this satisfies Moser's equation. It follows that $\psi = \rho_1$ defines a symplectomorphism on U_2 between ω_0 and ω_1 that preserves p.

We can now define the coordinate chart on $V = U_2$ by $\hat{x}^i = x^i \circ \psi$ and $\hat{y}^i = y^i \circ \psi$, then it follows

$$\sum_{i=1}^n d\hat{x}^i \wedge d\hat{y}^i = \sum_{i=1}^n d\left(x^i \circ \psi\right) \wedge d\left(y^0 \circ \psi\right) = \psi^* \left(\sum_{i=1}^n dx^i \wedge dy^i \right) = \psi^* \omega_1 = \omega_0.$$

Thus $\left(\hat{x}^i, \hat{y}^i\right)$ are the coordinates we wanted. □

Remark. The closedness of the symplectic form is necessary as $\omega_{\mathrm{st},2n}$ is closed. If ω is symplectomorphic to $\omega_{\mathrm{st},2n}$ in some neighbourhood, this symplectomorphism would preserve the closedness. Hence, closedness is necessary for the symplectic form to be 'flat'. //

The local theory is therefore always that of the trivial theory, however, this does not mean that there are no interesting global phenomena to explore. We will not go in this direction now, instead, we will discuss one of the main applications and the birthplace of symplectic geometry: Classical mechanics.

Chapter 4

Classical Mechanics as a Physicist

In the previous chapters, we discussed abstract geometrical objects called symplectic manifolds. From this point forward, we will turn towards an application: classical physics. Using symplectic geometry we can build a formal theory of classical mechanics, namely Hamiltonian mechanics. As many mathematicians may not be familiar with the contemporary formulations of classical mechanics past Newton's formalism, we will first dedicate a chapter to introducing physics to showcase the heuristic approach and create some intuition behind the methods. To do this we will discuss three formalisms: Newtonian, Lagrangian and Hamiltonian. Each of these formalisms is based on different principles, resulting in different methods of solving mechanical systems. For anyone familiar with these descriptions of physics, this chapter can easily be skipped without any continuity problems. This chapter is a combination of [21] and [2]. The translation of classical mechanics to symplectic geometry will be dealt with in Chapter 5.

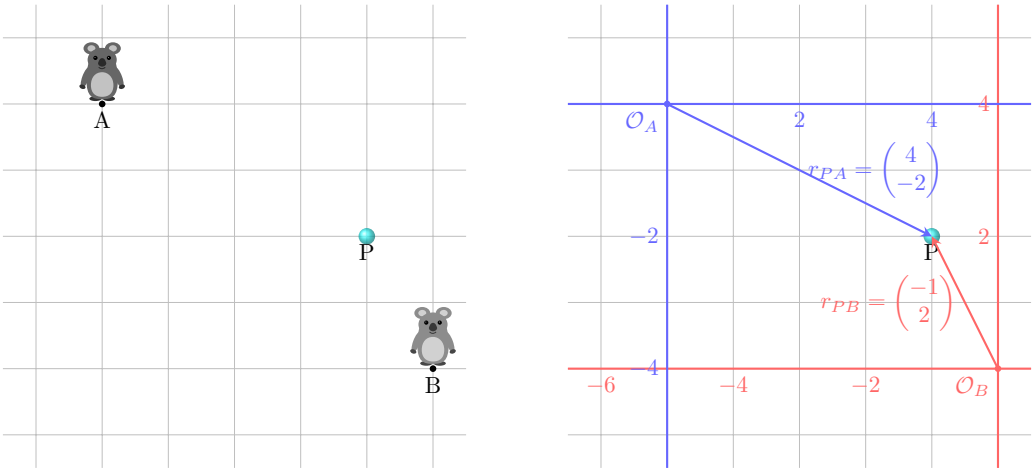
4.1 Newtonian Formalism

We will make our first step into describing classical mechanics by taking a look at Newton's formulation and basic principles in the form of his three laws of motion. Nowadays these still form the foundation of classical mechanics. Newton was one of the first persons who saw that we could describe physics using some general mathematical model, which can be seen as the goal of classical physics nowadays: to predict the motions of a physical space using a mathematical model. Before we go into the actual model Newton built, we will discuss how we can even translate a physical space to a mathematical one in the first place.

4.1.1 Space, Time and Kinematics

As the goal is to describe the motions of mechanical systems mathematically, we should first determine the types of systems we are studying and how we could translate these to mathematical spaces. The assumptions in classical mechanics are that the objects are relatively large such that there are no quantum mechanical effects and the speeds are relatively small to stay non-relativistic. We then follow our intuition and suppose that positions in physical space are points of a three-dimensional Euclidean space E^3. We would like to induce some vector space structure onto E^3, which can be done by fixing an origin $o \in E^3$, also called an observer, and we then identify a point $s \in E^3$ with the vector $\vec{os} \in \mathbb{R}^3$. When working with these vectors, we then also need to choose some basis vectors. We have some freedom of choice for what position acts as the origin and how we arrange the basis vectors, depending on this choice the position of an object may seemingly change, see Example 4.1. In a physical problem, we can often choose a reference frame that uses the symmetries of a system.

Example 4.1. Let us assume that the reference frames are chosen such that the positions in this example are constrained to $\mathbb{R}^2 \times \{0\} \subset \mathbb{R}^3$. Therefore, we will only consider the position on a plane instead of a three-dimensional space. Suppose that we have a ball at a point p on E^3



(a) The space E^3 given by some grid on a plane on which a ball is placed at position P and two observers are positioned at A and B.

(b) The reference frames and measurements of observer A and B. The reference frame of observer A is given in blue, and the measured position of P is denoted by r_{PA}. For observer B everything is in red and the position is given by r_{PB}.

Figure 4.1: Example of how to translate a physical situation, Figure 4.1a, to measurements in different reference frames, Figure 4.1b.

and some observers A and B, see Figure 4.1a. We can then view the position p from both the reference frame of A and B. We can then measure the position of P in both reference frames, see Figure 4.1b. Observer A measures the position of P as $r_{PA} = (4, -2)^T$, while Observer B measures P to be at $r_{PB} = (-1, 2)^T$. //

As we mentioned, we are not dealing with any relativistic effects in this theory, and hence, we can identify time as a separate axis, called the time axis identifiable with $\mathbb{R}$. This time axis is important when talking about motion, as this is all about the position over time. We define the motion of an object as some smooth function $\mathcal{C} : I \to E^3$, where I is an interval on the time axis. In our reference frame, we can obtain a function $r : I \to \mathbb{R}^3$ called the trajectory of the object. Notice that this is the composition of the actual motion with our choice of reference frame, hence, it is dependent on this choice which may differ over time. Using this mathematical trajectory, we can describe some physical quantities as vectors. Namely, the velocity and acceleration, v and a respectively, are defined as follows

$$v(t) = \frac{\mathrm{d}}{\mathrm{d}t} r(t) = \dot{r}(t) \qquad \text{and} \qquad a(t) = \frac{\mathrm{d}^2}{\mathrm{d}t^2} r(t) = \ddot{r}(t) = \dot{v}(t).$$

The importance of these quantities stems from Newton's laws of motion. However, before discussing these in detail, we should find a way to model more than just a single object in a system. We have thus far described a position of a single object while we are often dealing with a system that includes many bodies that are interacting. Intuitively we assign a single E^3 for each object in the system, which can be formalised in terms of a product space. In other words, in an n-body system a position of the system is described as a point in $E^{3n} = E^3 \times \cdots \times E^3$. The position of the bodies in the system is described using a single vector $r = (r_1, \ldots, r_n) \in \mathbb{R}^{3n}$, implying that the motion becomes a smooth map $r(t) = (r_1(t), \ldots, r_n(t)) \in \mathbb{R}^{3n}$. The definition of velocity and acceleration then still applies.

4.1.2 Newton's Laws of Motion

Newton's laws of motion determine the motions of a system through two important concepts: momentum and forces. Momentum is a quantity of an object and is given for some object of mass m by $p = mv$. This is often intuitively thought of as the amount of movement an object has. Forces on the other hand should be thought of as the interactions bodies have with each other or the system. These come in many shapes and forms, for example, the gravitational force or Coulomb force. All these forces are the product of some interaction between two physical bodies and are described as some vector $F \in \mathbb{R}^3$ acting on a body. As we will approximate motion to our best capability, it will often be useful to neglect some forces acting on larger bodies, this choice will make more sense in the light of Newton's second law. Let us now state these laws.

1. If the sum of all forces acting on a body is zero, there exists a reference frame in which its velocity is constant.
2. In any inertial frame, the time derivative of the momentum of a body is equal to the sum of forces acting on it.
3. For every action, there is an equal and opposite reaction.

While the first law seems to follow from the second law, we still need to ensure the existence of an inertial frame to use the second law. The second law then gives us the following equation, also called the equation of motion

$$\sum_i F_i\left(r, \dot{r}, t\right) = \dot{p}\left(t\right),$$

where F_i are all the forces acting on the body, $p\left(t\right)$ is its momentum at time t. We will often denote the sum of all forces acting on a body with F_{net}. If we also assume that the body has a constant mass this equation simplifies to

$$F_{\text{net}}\left(t, r, \dot{r}\right) = m\ddot{r}.$$

If we generalise this to an n-body system, we obtain an equation of motion for each of the bodies. In the case that the masses of all objects are constant, which is nearly always the case, this results in the following system of equations:

$$\begin{pmatrix} \dot{r} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} v \\ M^{-1}F\left(r, v, t\right) \end{pmatrix}, \qquad M = \operatorname{diag}\left(m_1 I_3, \ldots, m_n I_3\right). \tag{4.1}$$

Remark that we have transformed the differential equation by substituting $v = \dot{r}$. This system of equations is the one we solve most often in Newtonian mechanics. The process of solving for the motions of a system uses the following steps:

a) Choose a reference frame, and initial values then determine which forces are at play.

b) Describe the forces in terms of the reference frame.

c) Solve the equation of motion.

Let us showcase this with a couple of examples.

Example 4.2. Suppose we have set up a cannon atop an h meter high cliff and we want to determine the distance it can shoot a cannonball which weighs m kilograms, see Figure 4.2. We would then want to find a mathematical model to capture the motion of the cannonball.

First, we need to determine a reference frame which we have already chosen in Figure 4.2 to be at the foot of the cliff with the x corresponding to the horizontal direction and the y axis with the vertical one. Next up, we determine the initial values of the system. From Equation 4.1, it is clear that we need both the initial position and velocity of a system to solve for the motion. In this case, the initial position of the cannonball can be considered to be the position of the cannon, i.e. $r\left(0\right) = \left(0, h\right)^T$. The initial velocities are determined by some parameters, such that

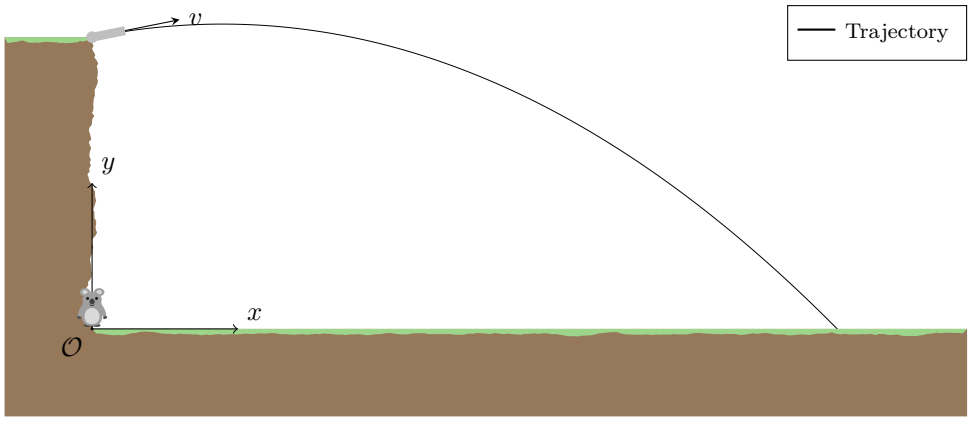


Figure 4.2: A sketch of a cannon atop a hill shooting a cannonball at some angle and speed. In the figure, we plotted a trajectory which we calculated by numerically solving Newton's equations of a projectile motion in a constant gravitational field with linear and quadratic air resistance using the SciPy package in Python. Our initial velocity for this trajectory was set to 15 ms^{-1} in the x-direction and 3 ms^{-s} in the y-direction and the height of the cliff was set at 10 m. We will see in our analysis that the mass of the cannonball is irrelevant to the problem. At the foot of the cliff, the reference frame for the analysis is drawn as well.

$v(0) = (v_{x0}, v_{y0})$. Lastly, we need to determine the forces which we simply have to guess and tune until we are satisfied with the model. In this case, we introduce just the force of gravity.

$$F_{\text{grav}}(t, r, \dot{r}) = F_{\text{grav}} = \begin{pmatrix} 0 \\ -mg \end{pmatrix}.$$

Here, $g \in \mathbb{R}$ is the acceleration due to gravity. As this is the only force we will be considering, we need to solve the following initial value problem:

$$\dot{u} = \begin{pmatrix} \dot{r}_x \\ \dot{r}_y \\ \dot{v}_x \\ \dot{v}_y \end{pmatrix} = \begin{pmatrix} v_x \\ v_y \\ 0 \\ -g \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} u + \begin{pmatrix} 0 \\ 0 \\ 0 \\ -g \end{pmatrix}, \qquad u(0) = \begin{pmatrix} 0 \\ h \\ v_{x0} \\ v_{y0} \end{pmatrix}.$$

Remark that this equation is of the form $\dot{u} = Au + b$, thus we can find the solution by integrating, see Theorem 2.4.1 in [17], such that

$$u(t) = e^{tA}u(0) + \int_0^t e^{(t-s)A} b\, ds = \begin{pmatrix} v_{x0}t \\ h + v_{y0}t - gt^2/2 \\ v_{x0} \\ v_{y0} - gt \end{pmatrix}.$$

Therefore, we get the motion for the cannonball with gravity

$$r(t) = \begin{pmatrix} v_{x0}t \\ h + v_{y0}t - gt^2/2 \end{pmatrix}. \tag{4.2}$$

We have plotted an example of such a motion in Figure 4.3. Here we see that our model is not quite perfect, but it does approximate the trajectory quite well. //

Example 4.3. Let us consider an Atwood machine, which is a system of two stationary masses, m_1 and m_2, hanging on a rope over a pulley, see Figure 4.4a. Assume that the rope and pulley are massless, there is no friction in the pulley, the rope does not slip on the pulley and the length of the rope is constant. Set the origin at the centre of the pulley such that the initial positions of the masses are given by $r_1 = (-R, -y_{10})^T$ and $r_2 = (R, -y_{20})^T$, where R is the radius of the pulley.

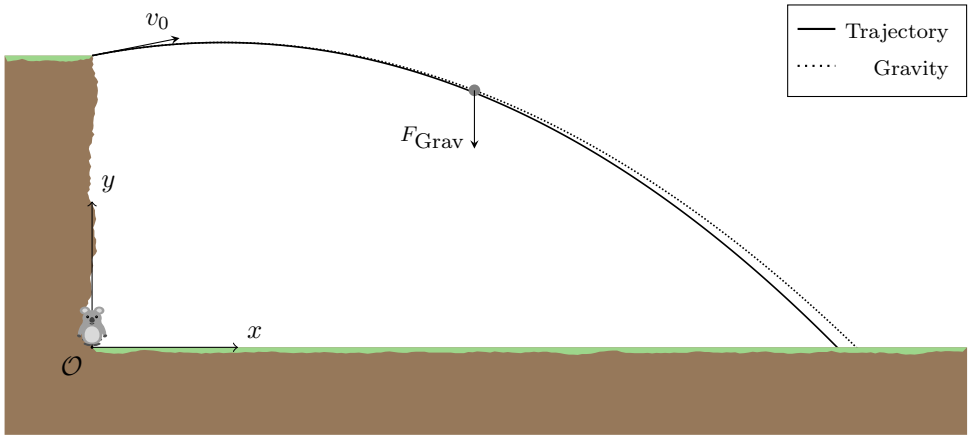
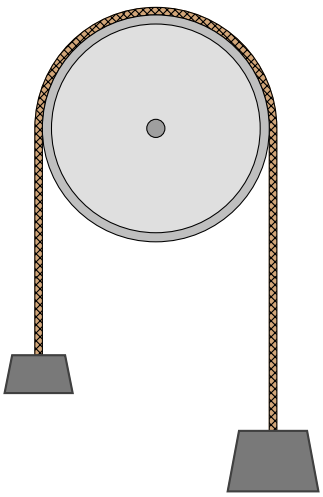
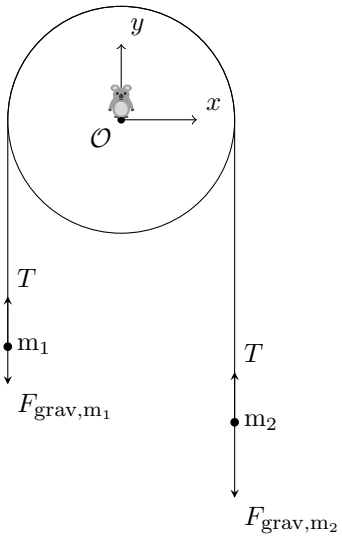


Figure 4.3: The trajectory of a cannonball as predicted by Equation 4.2 and the calculated trajectory of Figure 4.2. The same initial conditions were used in this figure. Remark that the trajectories are rather close to each other, but we see that the model predicts the range of the cannon to be further than the calculated range. Yet, our model is quite close to the numerical analysis.



(a) The configuration of a pulley system with two masses of different sizes hanging from it.



(b) The Force diagram associated with Figure 4.4a. The reference frame is drawn at the centre of the pulley.

Figure 4.4: Pulley system as described in Example 4.3 with both a sketch and a force diagram.

The forces in the system are then the force of gravity acting on both the objects, such that $F_{\text{grav,m}_i} = (0, -m_i g)^T$, and a tension force $T = (0, T)^T$ which is the same on both objects by the constraint on the length of the rope. These forces are drawn in Figure 4.4b. We can translate this to the following equation of motion.

$$\begin{pmatrix} m_1 a_{x1} \\ m_1 a_{y1} \\ m_2 a_{x2} \\ m_2 a_{y2} \end{pmatrix} = \begin{pmatrix} 0 \\ T + F_{\text{grav,m}_1} \\ 0 \\ T + F_{\text{grav,m}_2} \end{pmatrix} = \begin{pmatrix} 0 \\ T - m_1 g \\ 0 \\ T - m_2 g \end{pmatrix}.$$

We notice here that we can ignore the x-coordinates as there is no net force acting in this direction. However, when solving this equation, we run into the problem that T is an unknown. Luckily, we can recover it by remarking that $a_1 = -a_2$, which is a result of the fact that the length of the rope is constant. We can solve this equation for T.

$$\begin{aligned} \frac{T - m_1 g}{m_1} = a_1 = -a_2 = -\frac{T - m_2 g}{m_2} \\ m_2 T - m_1 m_2 g = -m_1 T + m_1 m_2 g \\ T = \frac{2 m_1 m_2}{m_1 + m_2} g. \end{aligned}$$

Substituting this into the equation of motion, we get

$$\begin{pmatrix} a_{y1} \\ a_{y2} \end{pmatrix} = \begin{pmatrix} 2m_2 g/(m_1 + m_2) - g \\ 2m_1 g/(m_1 + m_2) - g \end{pmatrix} = \begin{pmatrix} g\,(m_2 - m_1)/(m_1 + m_2) \\ g\,(m_1 - m_2)/(m_1 + m_2) \end{pmatrix}.$$

Solving this system is then rather simple and results in the motion of the masses.

$$r\,(t) = \begin{pmatrix} r_1\,(t) \\ r_2\,(t) \end{pmatrix} = \begin{pmatrix} y_{10} + \mu g t^2/2 \\ y_{20} - \mu g t^2/2 \end{pmatrix}, \qquad \mu = \frac{m_2 - m_1}{m_1 + m_2}. \tag{4.3}$$

Here, we can see that the mass difference is the main factor in the dynamics. If the difference is zero, the masses will stay stationary. If one of them is heavier, the greater mass will move downwards. //

4.2 Lagrangian Formalism

In the last section, we discussed Newton's formulation of classical mechanics. Here, we also went over its application to two rather simple examples, Examples 4.2 and 4.3. We saw that it can be quite cumbersome to work with the geometry of vectors and constraints. Hence, people started developing scalar theories of classical mechanics. Lagrangian formalism is an example of such a scalar approach to classical mechanics. Where Newton focussed on force and momentum, Lagrange only need to account for the energies in a system. He can connect the motion of the objects with the energies in the system through Hamilton's principle, more accurately called the principle of stationary action. From this principle, we are then able to extract the Euler-Lagrange equations which form the equations of motion in this formalism. These Euler-Lagrange equations are even stronger as they let us use generalised coordinates which cover the space of configurations of the system. Before we go into the methodology of Lagrangian formalism, we will shortly discuss the scalar quantity of energy.

4.2.1 Energy

Let us for now assume that we are still working on Euclidean space in Cartesian coordinates like in Newtonian mechanics. We distinguish two categories of energy: kinetic energy and potential energy. The first one is tied to the motion of an object while the second one results from the different interactions between objects and the system. We will introduce both of these objects through their interaction mechanism: work done by forces.

Definition 4.4. Let F be a force acting on an object which is moving along a curve $\mathcal{C}$, we define the **work done by F on the object along $\mathcal{C}$** as

$$W = \int_{\mathcal{C}} F \cdot ds = \int_{t_i}^{t_f} F\left(r\left(t\right), \dot{r}\left(t\right), t\right) \cdot \dot{r}\left(t\right) dt,$$

where $r : [t_i, t_f] \to \mathcal{C}$ is some parametrisation of the curve. //

Let us consider the work done by the net force acting on an object along the physical path it follows using Newton's equation of motion, we can then rewrite this integral

$$W = \int_{t_i}^{t_f} F\left(r\left(t\right), \dot{r}\left(t\right), t\right) \cdot \dot{r}\left(t\right) dt = \int_{t_i}^{t_f} m\ddot{r}\left(t\right) \cdot \dot{r}\left(t\right) dt = \int_{t_i}^{t_f} d\left(\frac{1}{2} m \|\dot{r}\left(t\right)\|^2\right).$$

If we then define the quantity $T\left(\dot{r}\right) = m\|\dot{r}\|^2/2$, called the **kinetic energy**, we can express the work done on the object to be the change in kinetic energy of the object. If we add up the kinetic energy of all objects in a system, we get the **total kinetic energy** T_{tot}, which we often denote with just T.

Meanwhile, we could also try to integrate the work done by a single force. In very few cases is this of a neat form, hence, we will consider forces that lend themselves to an interpretation much like the kinetic energy.

Definition 4.5. A force $F : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ is called a **general conservative force** if there exists some $U : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ such that given some path $r : \mathbb{R} \to \mathbb{R}^3$ we have

$$F_i\left(r\left(t\right), \dot{r}\left(t\right), t\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial U}{\partial \dot{r}_i}\left(r\left(t\right), \dot{r}\left(t\right), t\right)\right) - \frac{\partial U}{\partial r_i}\left(r\left(t\right), \dot{r}\left(t\right), t\right). \tag{4.4}$$

Here, the subscript i denotes the ith component of the force, position and velocity vector. The function U is called the **general potential energy** of F. //

Example 4.6. Take the Lorentz force acting on a particle with charge q moving along a path $r : \mathbb{R} \to \mathbb{R}^3$, i.e. the position is time-dependent, in an electric field E and magnetic field B. The Lorentz force is then given by

$$F_{\text{Lorentz}}\left(r, \dot{r}, t\right) = q\left(E\left(r, t\right) + \dot{r} \times B\left(r, t\right)\right).$$

Assume we have a scalar potential $V\left(r, t\right)$ and vector potential $A\left(r, t\right)$ that satisfy the following

$$B\left(r, t\right) = \nabla \times A\left(r, t\right), \qquad E\left(r, t\right) = -\nabla V\left(r, t\right) - \frac{\partial A}{\partial t}\left(r, t\right).$$

We can show that the potential of F in the sense of Equation 4.4 is given by

$$U\left(r, \dot{r}, t\right) = q\left(V\left(r, t\right) - \dot{r} \cdot A\left(r, t\right)\right).$$

We can check that this satisfies Equation 4.4 for the Lorentz force. The first component of the Lorentz force is given by

$$F_{\text{Lorentz},1} = q\left(-\frac{\partial V}{\partial r_1} - \frac{\partial A_1}{\partial t} + \dot{r}_2\left(\frac{\partial A_1}{\partial r_2} - \frac{\partial A_2}{\partial r_1}\right) - \dot{r}_3\left(\frac{\partial A_1}{\partial r_3} - \frac{\partial A_3}{\partial r_1}\right)\right).$$

Meanwhile, the first component of the right-hand side of Equation 4.4 can be expressed as

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial U}{\partial \dot{r}_1} - \frac{\partial U}{\partial r_1} &= q\left(-\frac{\mathrm{d}A_1}{\mathrm{d}t} - \frac{\partial V}{\partial r_1} + \sum_i \dot{r}_i \frac{\partial A_i}{\partial r_1}\right) \\
&= q\left(-\frac{\partial A_1}{\partial t} - \sum_i \dot{r}_i \frac{\partial A_1}{\partial r_i} - \frac{\partial V}{\partial r_1} + \sum_i \dot{r}_i \frac{\partial A_i}{\partial r_1}\right) \\
&= q\left(-\frac{\partial A_1}{\partial t} - \frac{\partial V}{\partial r_1} + \dot{r}_2\left(\frac{\partial A_2}{\partial r_1} - \frac{\partial A_1}{\partial r_2}\right) - \dot{r}_3\left(\frac{\partial A_1}{\partial r_3} - \frac{\partial A_3}{\partial r_1}\right)\right).
\end{aligned}$$

It follows from similar calculations that the other components are equal as well. Hence, U is indeed the general potential energy of the Lorentz force. //

In the Lagrangian formalism, we do allow for generalised conservative forces. However, in the case that the potential does not explicitly depend on time, even the work simplifies as follows.

$$\begin{aligned}
W &= \int_{t_i}^{t_f} F\left(r\left(t\right), \dot{r}\left(t\right), t\right) \cdot \dot{r}\left(t\right) dt \\
&= \sum_{i=1}^{3} \left[\int_{t_i}^{t_f} \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial U}{\partial \dot{r}_i} \left(r\left(t\right), \dot{r}\left(t\right)\right) \right) - \frac{\partial U}{\partial r_i} \left(r\left(t\right), \dot{r}\left(t\right)\right) \right] \cdot \dot{r}_i\left(t\right) dt \right] \\
&= \sum_{i=1}^{3} \left[\int_{t_i}^{t_f} \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial U}{\partial \dot{r}_i} \left(r\left(t\right), \dot{r}\left(t\right)\right) \cdot \dot{r}_i\left(t\right) \right) - \frac{\partial U}{\partial \dot{r}_i} \left(r\left(t\right), \dot{r}\left(t\right)\right) \cdot \ddot{r}_i\left(t\right) \right. \right. \\
&\qquad \left. \left. - \frac{\partial U}{\partial r_i} \left(r\left(t\right), \dot{r}\left(t\right)\right) \cdot \dot{r}_i\left(t\right) \right] dt \right] \\
&= \sum_{i=1}^{3} \left[\frac{\partial U}{\partial \dot{r}_i} \left(r\left(t\right), \dot{r}\left(t\right)\right) \cdot \dot{r}_i\left(t\right) \Big|_{t_i}^{t_f} - \int_{t_i}^{t_f} \frac{\mathrm{d}}{\mathrm{d}t} \left(U\left(r\left(t\right), \dot{r}\left(t\right)\right)\right) dt \right] \\
&= \sum_{i=1}^{3} \left[\frac{\partial U}{\partial \dot{r}_i} \left(r\left(t\right), \dot{r}\left(t\right) \cdot \dot{r}_i\left(t\right)\right) \Big|_{t_i}^{t_f} - U\left(r\left(t\right), \dot{r}\left(t\right)\right) \Big|_{t_i}^{t_f} \right].
\end{aligned}$$

Even though such a force does lead to a closed form for the work done by it, it can still be quite messy to work with. Therefore, we introduce the more simplistic **conservative force**, which is a force $F : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ such there exists a **potential energy** $U : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ which satisfies $F = -\nabla U$. In this case, the work simplifies to $W = -\Delta U$.

Proposition 4.7. *A force, $F : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$, is called conservative if it can be written as*

$$F\left(r, t\right) = f\left(\|r\|, t\right) \frac{r}{\|r\|}, \tag{4.5}$$

where $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$.

Proof. Suppose that $F\left(r, t\right)$ is as in Equation 4.5 and $r_0 \in \mathbb{R}^3$. Define $U : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ as

$$U\left(r, t\right) = \int_{\|r\|}^{\infty} f\left(s, t\right) ds.$$

We assume that this integral exists, however, as a potential is defined up to a constant the upper limit can be chosen arbitrarily such that the integral does exist. We can deduce that

$$\frac{\partial U}{\partial r_i}\left(r, t\right) = \frac{\mathrm{d}}{\mathrm{d}r_i} \int_{\|r\|}^{\infty} f\left(s, t\right) ds = -\frac{\partial \|r\|}{\partial r_i} f\left(\|r\|, t\right) = -f\left(\|r\|, t\right) \frac{r_i}{\|r\|}.$$

Hence, we can deduce that U is indeed the potential of F. □

Example 4.8. Consider the force of gravity acting on an object placed at r_1 and exerted by an object at r_2, the force can be expressed as

$$F\left(r_1\right) = \frac{G m_1 m_2}{\|r_1 - r_2\|^2} \frac{r_2 - r_1}{\|r_1 - r_2\|}.$$

We can define the potential energy $U\left(r_1\right)$ as

$$U\left(r_1\right) = \frac{G m_1 m_2}{\|r_2 - r_1\|}.$$

Let us check that this is indeed the correct potential,

$$-\nabla U\left(r_1\right) = \frac{G m_1 m_2}{\|r_1 - r_2\|^2} \frac{r_2 - r_1}{\|r_1 - r_2\|}.$$

Thus the force of gravity is conservative. See that it can indeed be written in terms of $r = r_2 - r_1$, i.e. $F(r) = Gm_1m_2r \Big/ \|r\|^3$. Furthermore, remark that the gravitational force on the second body exerted by the first can be obtained by taking the gradient with respect to r_2, i.e. if we fix r_1 and make r_2 a variable, we obtain

$$F(r_2) = -\nabla U(r_2) = \frac{Gm_1m_2}{\|r_1 - r_2\|^2} \frac{r_1 - r_2}{\|r_1 - r_2\|} = -F(r_1).$$

This is exactly Newton's third law. //

Example 4.8 shows us that a potential is much more of a measure of the interaction rather than a quantity tied to a body. We would like to define the total potential energy, U_{tot} in such a way that $\partial U_{\text{tot}}/\partial r_i$ is the net force on the i-th particle. Hence, the **total potential energy** of a system is given by $U_{\text{tot}} = \sum_i U_i$, where the sum runs over all the interactions.

4.2.2 Euler-Lagrange Equations

With the concepts of energy at hand, we can dive into Lagrangian formalism. This formalism lends itself to working with constrained systems in a more natural manner. In Newtonian mechanics, constraints led to imposing constraint forces, like the tension in Example 4.3. In Lagrangian formalism, we work around these constraints by choosing suitable coordinates which span all the possible configurations of the system. By choosing our coordinates wisely, we can often reduce the apparent dimensionality of the system. Such a system of coordinates is often denoted with $q = (q_1, \ldots, q_n)$ instead of $r = (r_1, \ldots, r_n)$. In such general coordinates, we can state the basic principle of Lagrangian formalism: Hamilton's principle.

Hamilton's principle. *The actual motion of a physical system, $q : [t_i, t_f] \to \mathbb{R}^3$, is a stationary point of the **action integral** defined as*

$$S(q) = \int_{t_i}^{t_f} \mathscr{L}(q(t), \dot{q}(t), t)\, dt.$$

*Where $\mathscr{L}(q, \dot{q}, t) = T(q, \dot{q}, t) - U(q, \dot{q}, t)$ is called the **Lagrangian** of the system.*

Remark. Remark that in the previous section, the kinetic energy was only a function of $\dot{r}$, but in general coordinates, we can have some dependence on the position. For example, if we use cylindrical coordinates, $(x, y, z) = (r\sin\theta, r\sin\theta, z)$ one can deduce that $T = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta} + \dot{z}^2\right)$. Hence, it is dependent on both the velocities and the position. Furthermore, U can be considered in the sense of a general potential and can therefore be dependent on the velocity. //

Remark. Remark that a Lagrangian only results in a well-posed mechanical situation if the stationary point of the action integral is uniquely for some boundary conditions. One can deduce that this implies that the Lagrangian must be a convex function in $\dot{q}$, see [3, p. 57] or [12, Section 1.4]. //

This principle is equivalent to the second law of motion posed by Newton. Finding the stationary points of an integral may seem like a convoluted way of finding the motions, and it is not even clear how this is equivalent to Newton's formalism. Luckily, both these problems are solved using the **Euler-Lagrange equations**.

Proposition 4.9. *For a Lagrangian $\mathscr{L}$, any stationary point $q : [t_i, t_f] \to \mathbb{R}$ of the action integral satisfies the following*

$$\frac{\partial \mathscr{L}}{\partial q} = \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathscr{L}}{\partial \dot{q}}. \tag{4.6}$$

Moreover, any path that satisfies this condition is a stationary point.

Proof. Suppose we are given a Lagrangian $\mathscr{L} = \mathscr{L}(q, \dot{q}, t)$. Let us define a stationary point of the action integral. First, we define the notion of a variation of a path $q : [t_i, t_f] \to \mathbb{R}$ as a path

$\eta : [t_i, t_f] \to \mathbb{R}$ with $\eta(t_i) = 0 = \eta(t_f)$. We can then vary the path q smoothly in the direction of η as $\rho_{\alpha,\eta} : [t_i, t_f] \to \mathbb{R} : t \mapsto q(t) + \alpha\eta(t)$. For each variation η, we can express S as a function of α,

$$S_\eta(\alpha) = S(\rho_{\alpha,\eta}) = \int_{t_i}^{t_f} \mathscr{L}(\rho_{\alpha,\eta}(t), \dot{\rho}_{\alpha,\eta}(t), t)\, dt.$$

We then call q a stationary point of S if for every variation η we have $\mathrm{d}S_\eta/\mathrm{d}\alpha\,|_{\alpha=0} = 0$. Now suppose that $q : [t_i, t_f] \to \mathbb{R}$ is a stationary point of S and η is some variation. Using the Leibniz and product rule, we can deduce that

$$\begin{aligned} 0 = \left.\frac{\mathrm{d}S_\eta}{\mathrm{d}\alpha}\right|_{\alpha=0} &= \int_{t_i}^{t_f} \left.\frac{\mathrm{d}}{\mathrm{d}\alpha}\right|_{\alpha=0} (\mathscr{L}(\rho_{\alpha,\eta}(t), \dot{\rho}_{\alpha,\eta}(t), (t)))\, dt \\ &= \int_{t_i}^{t_f} \left(\frac{\partial \mathscr{L}}{\partial q}(\rho_{0,\eta}(t), \dot{\rho}_{0,\eta}, t)\, \eta(t) + \frac{\partial \mathscr{L}}{\partial \dot{q}}(\rho_{0,\eta}(t), \dot{\rho}_{0,\eta}(t), t)\, \dot{\eta}(t) \right) dt. \end{aligned}$$

By rewriting $\rho_{0,\eta}(t)$ as $q(t)$ and using partial integration on the second term in combination that $n(t_i) = 0 = n(t_f)$ we get the following.

$$0 = \int_{t_i}^{t_f} \eta(t) \left(\frac{\partial \mathscr{L}}{\partial q}(q(t), \dot{q}(t), t) - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathscr{L}}{\partial \dot{q}}(q(t), \dot{q}(t), t) \right) dt.$$

As we have chosen our variation η arbitrarily, this is enough to conclude that

$$\frac{\partial \mathscr{L}}{\partial q}(q(t), \dot{q}(t), t) - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathscr{L}}{\partial \dot{q}}(q(t), \dot{q}(t), t) = 0.$$

Hence, any path q that is a stationary point of the action integral satisfies Equation 4.6. Moreover, any path that satisfies this equation is a stationary point of the action integral by the same logic. □

If we generalise this to a higher dimensional system with some general coordinates $q = (q_1, \ldots, q_n)$, we conclude that the physical path satisfies

$$\frac{\partial \mathscr{L}}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathscr{L}}{\partial \dot{q}_i},\ \forall 1 \leq i \leq n.$$

See [12, Proposition 1.4.1] for a formal proof of this statement. The Euler-Lagrange equations show that given n generalised coordinates, we end up with n second-order differential equations we need to solve in order to recover the physical motions of the bodies.

Remark. In Cartesian coordinates, the Euler-Lagrange equations are equivalent to Newton's second law.

$$\frac{\partial \mathscr{L}}{\partial q} = \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathscr{L}}{\partial \dot{q}} \iff -\frac{\partial U}{\partial q} = \frac{\mathrm{d}}{\mathrm{d}t} \left(m\dot{q} - \frac{\partial U}{\partial \dot{q}} \right) \iff \dot{p} = \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial U}{\partial \dot{q}} - \frac{\partial U}{\partial q} = F.$$

Here we used that U is the general potential of the force acting on the body. //

Solving problems with this formalism is often seen as more straightforward and less error-prone. In practice we need to go through the following steps:

a) Determine the kinetic and potential energies in an inertial frame.

b) Determine the Lagrangian and translate it to some general coordinates for the system.

c) Solve the Euler-Lagrange equation.

We will now showcase the power of Lagrangian formalism using two examples.

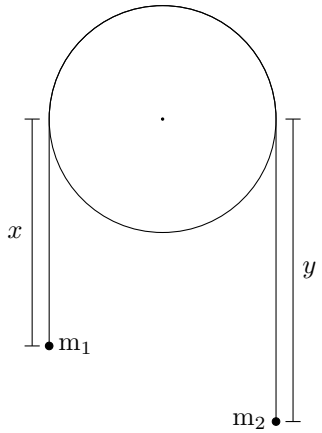


Figure 4.5: Generalised coordinates of an Atwood machine.

Example 4.10. Let us consider the Atwood machine of Example 4.3 again, i.e. we consider two stationary masses hanging from a rope which is suspended over a pulley. We assume that the rope and pulley are massless, the rope does not slip on the pulley, the pulley can rotate freely and the length of the rope is constant. We will now solve this using Lagrangian formalism. Define the coordinates x and y as in Figure 4.5. As the rope has a constant length, say L, we get the relation $x+y+R\pi = L$, implying that $y = -x+\tilde{C}$. Hence, the system can be described using a single general coordinate x. The kinetic energy of this system is given by

$$T(x) = \frac{1}{2}m_1\dot{x}^2 + \frac{1}{2}m_2\dot{y}^2 = \frac{1}{2}(m_1+m_2)\dot{x}^2.$$

The potential energy is the sum of the gravitational potentials of both masses.

$$U = -m_1gx - m_2gy = -(m_1-m_2)gx + C.$$

Here, we can set the constant to zero as a potential is determined up to a constant. This leads to the following Lagrangian.

$$\mathscr{L}(x,\dot{x}) = \frac{1}{2}(m_1+m_2)\dot{x}^2 + (m_1-m_2)gx. \tag{4.7}$$

If we enter this into the Euler-Lagrange equations we get the following differential equation:

$$(m_1-m_2)g = (m_1+m_2)\ddot{x}.$$

This is solvable for x, which results in

$$x(t) = \frac{1}{2}\frac{m_1-m_2}{m_1+m_2}gt^2 + x_0. \tag{4.8}$$

This solves our system and we can see this is equivalent to the Newtonian case by comparing Equation 4.8 to Equation 4.3. //

Example 4.11. Consider the case of a block of mass m_1 sliding on a wedge of mass m_2 that can move horizontally as sketched in Figure 4.6, where we assume everything to be frictionless. If we were to solve this problem in Newtonian mechanics, we would have to deal with the awkward constraint force of the wedge acting on the mass and work out a lot of geometry. Luckily, we can circumvent this problem by using the energy methods of Lagrange and imposing the constraints through the coordinates we choose, as depicted with q_1 and q_2 in Figure 4.6.

We can set up the Lagrangian in the inertial frame, depicted by the x and y axes in Figure 4.6. As these are Cartesian coordinates, this is quite simple.

$$\mathscr{L}(x_1,y_1,x_2,y_2,\dot{x}_1,\dot{y}_1,\dot{x}_2,\dot{y}_2) = \frac{1}{2}m_1\left(\dot{x}_1^2+\dot{y}_1^2\right) + \frac{1}{2}m_2\left(\dot{x}_2^2+\dot{y}_2^2\right) + m_1gy_1.$$

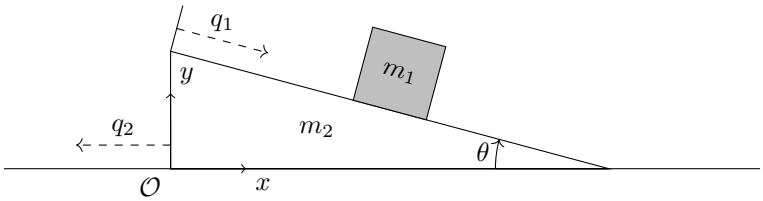


Figure 4.6: Sketch of a mass m_1 on a wedge of mass m_2 at an angle θ. Here, we allow the mass to slide along the wedge, and the wedge to slide horizontally. The origin is placed such that the right angle of the wedge coincides with it at $t = 0$, and the associated x and y-axis are also given. The coordinates q_1 and q_2 are also drawn.

Next, we need to translate the Lagrangian to the general coordinates q_1 and q_2, this is given by the following, notice that these are defined up to some constant.

$$\begin{aligned} x_1 &= -q_2 + q_1 \cos\alpha, & y_1 &= q_1 \sin\alpha, \\ x_2 &= -q_2, & y_2 &= 0. \end{aligned}$$

With these coordinate transformations, we can translate our Lagrangian to the general coordinates, resulting in

$$\mathscr{L}\left(q_1, q_2, \dot{q}_1, \dot{q}_2\right) = \frac{1}{2}\left(m_1 + m_2\right)\dot{q}_2^2 + \frac{1}{2}m_1\left(\dot{q}_1^2 - 2\dot{q}_1\dot{q}_2\cos\alpha\right) + m_1 g q_1 \sin\alpha.$$

The equations of motion are then given by the Euler-Lagrange equations,

$$m_1 g \sin\alpha = m_1\ddot{q}_1 - m_1\ddot{q}_2\cos\alpha, \tag{4.9}$$

$$0 = \left(m_1 + m_2\right)\ddot{q}_2 - m_1\ddot{q}_1\cos\alpha. \tag{4.10}$$

We can express $\ddot{q}_2$ as an equation of $\ddot{q}_1$ using Equation 4.10 and entering this into Equation 4.9 we can recover a closed form for both $\ddot{q}_1$ and $\ddot{q}_2$

$$\ddot{q}_1 = \frac{g\sin\alpha}{1 - \dfrac{m_1\cos^2\alpha}{m_1 + m_2}} \quad \text{and} \quad \ddot{q}_2 = \frac{m_1 g \sin\alpha\cos\alpha}{m_1\sin^2\alpha + m_2}.$$

Notice that both of these are constant, and one can thus easily solve these equations. We can check that these satisfy our intuition in the cases that $m_2 \to \infty$, $m_2 = 0$, $\alpha = 0$ or $\alpha = \pi/2$. //

4.3 Hamiltonian Formalism

Up until now, we have developed methods to solve an n-body problem using at most $3n$ differential equations, either Newton's second law or the Euler-Lagrange equations. In the Lagrangian formalism, we could lower the dimensionality of the problem by choosing generalised coordinates, which use the symmetries of our problem. However, both Newtonian and Lagrangian formalism end up giving us second-order differential equations, which are not very insightful. Hence, we would like to reduce the order of our system naturally. We will show that we can do this by introducing the general momenta of a Lagrangian system as the new coordinates, which results in the Hamiltonian through the Legendre transform. Before we go this route, we will introduce the Hamiltonian in Cartesian coordinates.

4.3.1 Cartesian Hamiltonian Mechanics

Let us consider an n-body system for which we have the total energy functional given by a function $\mathscr{H}\left(r_1, \ldots, r_n, p_1, \ldots, p_n, t\right)$, where r_i is the position of the ith body and p_i the momentum

of the ith body. We can write this in terms of the kinetic en potential energy, where we assume that the kinetic energy is only a function of p_i and the potential energy a function of r_i and t

$$\mathscr{H}(r_1,\ldots,r_n,p_1,\ldots,p_n,t) = T(p_1,\ldots,p_n) + U(r_1,\ldots,r_n,t) = \sum_i \frac{p_i^2}{2m_i} + U(r_1,\ldots,r_n,t).$$

Remark that the potential is chosen such that $F_i = \partial U/\partial r_i$, it then follows from the definition of momentum and Newton's second law that

$$\frac{\partial \mathscr{H}}{\partial p_i} = \frac{p_i}{m_i} = \dot{r}_i, \quad \frac{\partial \mathscr{H}}{\partial r_i} = \frac{\partial U}{\partial r_i} = -F_i = -\dot{p}_i. \tag{4.11}$$

This gives us $6n$ first-order differential equations to solve to obtain the motion of the n-bodies described by the energy functional. Furthermore, using these relations we obtain

$$\frac{\mathrm{d}\mathscr{H}}{\mathrm{d}t} = \dot{r}_i\frac{\partial \mathscr{H}}{\partial r_i} + \dot{p}_i\frac{\partial \mathscr{H}}{\partial p_i} + \frac{\partial \mathscr{H}}{\partial t} = \frac{\partial \mathscr{H}}{\partial t}.$$

Hence, the Hamiltonian is conserved as long as it does not depend on time directly.

4.3.2 The Hamiltonian in Generalised Coordinates

We will now try to extend the discussion of the previous section to work with general coordinates. Remark that the energy functional in the previous section was dependent on the position, momentum and time. We would like to replicate this in general coordinates to obtain a similar equation to Equation 4.11. To do this, we will first have to determine what our quantity of momentum is in generalised coordinates. We will define this in relation to a Lagrangian. Given a Lagrangian $\mathscr{L}$ in some generalised coordinates $(q_1,\ldots,q_n)$, we define the generalised momentum associated with a coordinate q_i as

$$p_i = \frac{\partial \mathscr{L}}{\partial \dot{q}_i}.$$

This definition might seem odd, but it works correctly in Cartesian coordinates.

Example 4.12. Consider the Lagrangian of n non-interacting objects. In this case, the Lagrangian in Cartesian coordinates is given by

$$\mathscr{L}(r_1,\ldots,r_n,\dot{r}_1,\ldots,\dot{r}_n,t) = \frac{1}{2}\sum_i m_i\|\dot{r}_i\|^2.$$

Hence, the generalised momentum associates with r_i is simply the momentum of the object

$$p_i = \frac{\partial \mathscr{L}}{\partial \dot{r}_i} = m_i\dot{r}_i.$$

The definition of the generalised momenta coincides with the usual definition when working in Cartesian coordinates when working with non-interacting particles. //

We would then like to naturally transform the Lagrangian $\mathscr{L}(q_1,\ldots,q_n,\dot{q}_1,\ldots,\dot{q}_n,t)$ to a function $\mathscr{H}\left(q_1,\ldots,q_n,\frac{\partial \mathscr{L}}{\partial \dot{q}_1},\ldots,\frac{\partial \mathscr{L}}{\partial \dot{q}_n},t\right) = \mathscr{H}(q_1,\ldots,q_n,p_1,\ldots,p_n,t)$, such that there is a change of dependent variable. This transformation can be formalised using the Legendre transform.

Legendre Transform

Consider a function $f: V \to \mathbb{R}$, most often we have $V = \mathbb{R}^n$, we define the **Legendre transform** of f as the function $f^*: S \subset V^* \to \mathbb{R}$ for which

$$f^*(\alpha) = \sup_{v\in V}\left(\alpha(v) - f(v)\right).$$

Remark that this function is defined for all $\alpha \in V^*$ for which the supremum is finite.

Example 4.13. Consider the function $f : \mathbb{R} \to \mathbb{R} : x \mapsto e^x$, we can recognise that $R^* \cong \mathbb{R}$ such that we can define the Legendre transform as

$$f^*(x^*) = \sup_{x \in \mathbb{R}} (x^* x - e^x),$$

where $x^* \in I^*$ which is some domain that we still have to determine. Let us first figure out what f^* is by calculating the supremum using the derivative with respect to x and setting it equal to zero.

$$0 = \frac{\mathrm{d}}{\mathrm{d}x}(x^* x - e^x) = x^* - e^x.$$

Hence we see that $x^* = e^x$ gives a critical point, and as $-e^x < 0$ for all x it follows that we achieve a maximum at $x = \ln(x^*)$. Thus we retrieve our Legendre transform

$$f^*(x^*) = (x^* - 1)\ln(x^*).$$

Remark that this function only exists for $x^* \in (0, \infty) = I^*$. //

Example 4.14. Take a function $g : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto g(x, y)$ which is convex in y and for a $x \in \mathbb{R}$ defined $f_x : \mathbb{R} \to \mathbb{R} : y \mapsto g(x, y)$. Remark that $\mathrm{d}f_x/\mathrm{d}y\,(y) = \partial g/\partial y\,(x, y)$. We can then determine the Legendre transform of f_x, which was defined as

$${f_x}^*(y^*) = \sup_{y \in \mathbb{R}} (y^* y - f_x(y)).$$

We can again determine this supremum by differentiating with respect to y. It then follows that

$$0 = \frac{\mathrm{d}}{\mathrm{d}y}(y^* y - f_x(y)) = y^* - \frac{\mathrm{d}f_x}{\mathrm{d}y}(y) \implies y^* = \frac{\mathrm{d}f_x}{\mathrm{d}y}(y).$$

As g is convex in y, it follows that $y = (\mathrm{d}f_x/\mathrm{d}y)^{-1}(y^*)$ is the maximum. We can now transform $g(x, y)$ to a function $h(x, \partial g/\partial y)$ by defining

$$h\left(x, \frac{\partial g}{\partial y}\right) = h(x, y^*) = {f_x}^*(y^*) = y^* \left(\frac{\mathrm{d}f_x}{\mathrm{d}y}\right)^{-1}(y^*) - g\left(x, \left(\frac{\mathrm{d}f_x}{\mathrm{d}y}\right)^{-1}(y^*)\right).$$

This leads to a natural transformation from $g(x, y)$ to $h(x, \partial g/\partial y)$. Furthermore, the differentials of g and h are given as

$$dg = u\,dx + v\,dy \implies dh = x\,du - v\,dy.$$

Remark that this is not merely a coordinate transformation, but also a transformation of the space on which the functions act. //

In the section on Lagrangian formalism, we remarked that a Lagrangian had to be convex in $\dot{q}$ if it were to result in a well-posed problem. Hence, we can use Example 4.14 to transform the Lagrangian in the manner discussed before. We define the **Hamiltonian** of a Lagrangian $\mathscr{L}$ as

$$\mathscr{H}(q_1, \ldots, q_n, p_1, \ldots, p_n, t) = \sum_{i=1}^{n} p_i \dot{q}_i - \mathscr{L}(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t),\ p_i = \frac{\partial \mathscr{L}}{\partial \dot{q}_i}. \tag{4.12}$$

Hence, the Hamiltonian can be seen as the Legendre transform of the Lagrangian. We can now determine the differential along a physical motion in two manners: directly and using equation 4.12. If we calculate it directly, we find

$$d\mathscr{H} = \sum_{i=1}^{n} \frac{\partial \mathscr{H}}{\partial q_i} + \sum_{i=1}^{n} \frac{\partial \mathscr{H}}{\partial p_i} dp_i + \frac{\partial \mathscr{H}}{\partial t} dt. \tag{4.13}$$

However, if we calculate it using Equation 4.12, we find that

$$d\mathscr{H} = \sum_{i=1}^{n} p_i d\dot{q}_i + \sum_{i=1}^{n} \dot{q}_i dp_i - d\mathscr{L}.$$

We can further expand this expression by first determining the differential of the Lagrangian.

$$d\mathscr{L} = \sum_{i=1}^{n} \frac{\partial \mathscr{L}}{\partial q_i} dq_i + \sum_{i=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial \mathscr{L}}{\partial t} dt.$$

Combining these two equations results in the differential of the Hamiltonian:

$$d\mathscr{H} = \sum_{i=1}^{n} p_i d\dot{q}_i + \sum_{i=1}^{n} \dot{q}_i dp_i - \sum_{i=1}^{n} \frac{\partial \mathscr{L}}{\partial q_i} dq_i - \sum_{i=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial \mathscr{L}}{\partial t} dt.$$

Again remark that $p_i = \partial \mathscr{L} / \partial \dot{q}_i$ and that it follows from the Euler-Lagrange equations that the physical path the motion follows satisfies $\dot{p}_i = \mathrm{d}/\mathrm{d}t \left(\partial \mathscr{L} / \partial \dot{q}_i \right) = \partial \mathscr{L} / \partial q_i$.

$$\begin{aligned} d\mathscr{H} &= \sum_{i=1}^{n} p_i d\dot{q}_i + \sum_{i=1}^{n} \dot{q}_i dp_i - \sum_{i=1}^{n} \dot{p}_i dq_i - \sum_{i=1}^{n} p_i d\dot{q}_i - \frac{\partial \mathscr{L}}{\partial t} dt \\ &= \sum_{i=1}^{n} \dot{q}_i dp_i - \sum_{i=1}^{n} \dot{p}_i dq_i - \frac{\partial \mathscr{L}}{\partial t} dt. \end{aligned} \tag{4.14}$$

Comparing Equation 4.13 and 4.14 gives us the $2n+1$ equations called **Hamilton's equations**:

$$-\frac{\partial \mathscr{L}}{\partial t} = \frac{\partial \mathscr{H}}{\partial t}, \quad -\dot{p}_i = \frac{\partial \mathscr{H}}{\partial q_i}, \quad \dot{q}_i = \frac{\partial \mathscr{H}}{\partial p_i}, \qquad \forall 1 \leq i \leq n.$$

Solving a mechanical problem comes down to solving this system of equations, most importantly the last $2n$, to obtain the motion of the objects. In practice solving a problem goes as follows:

a) Determine the kinetic and potential energies in an inertial frame.

b) Determine the Lagrangian and translate it to some general coordinates for the system.

c) Derive the generalised momenta from the Lagrangian and solve for the $\dot{q}$'s as functions of p's and q's.

d) Determine the Hamiltonian using Equation 4.12.

e) Solve Hamilton's equations.

Let us showcase this method using an example.

Example 4.15. Let us again consider the Atwood machine of Examples 4.3 and 4.10. We already showed how to derive the Lagrangian in the coordinate described in Figure 4.5, see Equation 4.7. To recover the Hamiltonian, we'll have to determine the generalised momentum

$$p = \frac{\partial \mathscr{L}}{\partial \dot{x}} = (m_1 + m_2)\, \dot{x}.$$

Using the Legendre transform, we then obtain the Hamiltonian

$$\mathscr{H}(x, p, t) = p\dot{x} - \mathscr{L}(x, \dot{x}, t) = \frac{p^2}{2(m_1 + m_2)} - \frac{m_1 - m_2}{g} x.$$

This results in the following equations for $\dot{x}$ and $\dot{p}$

$$\dot{x} = \frac{\partial \mathscr{H}}{\partial p} = \frac{p}{m_1 + m_2}, \quad \dot{p} = -\frac{\partial \mathscr{H}}{\partial x} = (m_1 - m_2)\, g.$$

We can see that these are equivalent to Equations 4.3 and 4.8. //

Chapter 5

Mathematical Methods of Classical Mechanics

In this chapter, we shall finally see symplectic geometry and classical mechanics come together and we will show that interpreting Hamiltonian mechanics as a geometric theory leads to some powerful results. To do this, we will first discuss how we let a symplectic form act on a smooth function, where we are again heavily inspired by Riemannian geometry. We will then shortly discuss Hamiltonian systems before diving deeper into the theory in terms of first integrals and integrable systems. We will see that these first integrals, which are the constants of the Hamiltonian flow, generate symmetries of the system and allow us to find Darboux charts that make sense in terms of mechanics. This chapter is mainly based on the ideas from [4, Chapter 18] and [1, Chapter 10]. Remark that we will only discuss Hamiltonians which do not explicitly depend on time.

5.1 Hamiltonian Vector Fields

We would like to determine some form of movement on a symplectic manifold with just a single smooth function. Normally one encodes movement on a manifold with a vector field, hence, we are looking for a manner of transforming a smooth function into a vector field. In calculus, this is done by using the gradient, hence, we will search for a symplectic equivalent of this, which we will call the Hamiltonian vector field of a function. Let us first consider how one does this on a Riemannian manifold and then replicate this method on symplectic manifolds.

5.1.1 Riemannian Gradients

In Euclidean geometry, we may define the gradient as follows.

Definition 5.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and $x \in \mathbb{R}^n$, if the partial derivates of f exist at a point $x \in \mathbb{R}^n$, define the gradient of f in $x = \left(x^1, \ldots, x^n\right)$ as

$$\operatorname{grad} f\left(x\right) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}\left(x\right) \hat{e}_i,$$

where $\hat{e}_i$ denotes the i-th basis vector of $\mathbb{R}^n$. //

It should be clear that the gradient takes a smooth function on $\mathbb{R}^n$ to a vector field, in this case, a function $\operatorname{grad} f : \mathbb{R}^n \to \mathbb{R}^n$. To generalise this construction to a Riemannian manifold, we consider the naturally defined differential of a function and map it to the tangent bundle using the musical isomorphisms generated by the Riemannian metric.

Definition 5.2. Let $(\mathcal{M}, g)$ be a Riemannian manifold and $f \in C^\infty\left(\mathcal{M}\right)$ then define

$$\operatorname{grad} f = \left(df\right)^\sharp.$$

Here, $\eta^\sharp = \hat{g}^{-1}(\eta)$ with $\hat{g} : \mathrm{T}\mathcal{M} \to \mathrm{T}^*\mathcal{M}$ is the induced map by the metric g. //

Remark. The isomorphism $\hat{g} : \mathrm{T}\mathcal{M} \to \mathrm{T}^*\mathcal{M}$ exists on the premise that g is non-degenerate. //

We should check that this is equivalent to the gradient in calculus when working on Euclidean space. In this case we have $\left(dx^i\right)^\sharp = \partial/\partial x^i$, such that

$$\operatorname{grad} f = (df)^\sharp = \left(\frac{\partial f}{\partial x^i} dx^i\right)^\sharp = \sum_{i=1}^n \frac{\partial f}{\partial x^i} \left(dx^i\right)^\sharp = \sum_{i=1}^n \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^i}.$$

We can then identify $\mathrm{T}\mathbb{R}^n \cong \mathbb{R}^n$ with $\partial/\partial x^i \mapsto \hat{e}_i$. This shows that this is indeed the vector field we wanted.

5.1.2 Symplectic Analogue

To create a symplectic analogue to the gradient, we should take extra notice of the non-degeneracy of the symplectic form and the Riemannian metric. The existence of the Riemannian gradient was dependent on this non-degeneracy, and hence, we can replicate this construction on symplectic manifolds. We will define the analogue, called a Hamiltonian vector field of a function, implicitly.

Definition 5.3. Let $(\mathcal{M}, \omega)$ be a symplectic manifold and $f \in C^\infty(\mathcal{M})$. We then define the **Hamiltonian vector field of** f as the vector fields X_f that satisfies $\iota_{X_f}\omega = df$. //

To see that this is the actual symplectic analogue of the gradient, remark that $\iota_X \omega = \hat{\omega}(X)$ and thus $\hat{\omega}^{-1}(df)$ satisfies this condition

$$\iota_{\hat{\omega}^{-1}(df)}\omega = df.$$

We can ensure the existence and uniqueness through the fact that ω is non-degenerate. However, this is not very insightful on the form of the Hamiltonian vector field. Therefore, we will prove the existence and uniqueness in coordinates.

Proposition 5.4. *For every function $f \in C^\infty(\mathcal{M})$, where $(\mathcal{M}, \omega)$ is a symplectic manifold, there exists a unique Hamiltonian vector field, X_f, and in a Darboux chart $\left(U, \left(x^1, \ldots, x^n, y^1, \ldots, y^n\right)\right)$ it is given by*

$$X_f|_U = \sum_{i=1}^n \left(\frac{\partial f}{\partial y^i}\frac{\partial}{\partial x^i} - \frac{\partial f}{\partial x^i}\frac{\partial}{\partial y^i}\right).$$

Proof. Suppose that $(\mathcal{M}, \omega)$ is a symplectic $2n$-manifold and $f \in C^\infty(\mathcal{M})$. We can then solve the equation $\iota_{X_f}\omega = df$ in a Darboux chart $\left(U, \left(x^1, \ldots, x^n, y^1, \ldots, y^n\right)\right)$. To this extent we write $X_f|_U = \sum_{i=1}^n \left(X_x^i\, \partial/\partial x^i + X_y^i\, \partial/\partial y^i\right)$ and the differential of f is then given by $df|_U = \sum_{i=1}^n \left(\frac{\partial f}{\partial x^i} dx^i + \frac{\partial f}{\partial y^i} dy^i\right)$. The contraction of ω by X_f can be calculated using Equation 3.4,

$$\iota_X \omega|_U = \sum_{i=1}^n \left(X_x^i dx^i - X_y^i dy^i\right).$$

Setting this equal to the differential $df|_U$ results in $2n$ equations of the form

$$X_x^i = \frac{\partial f}{\partial y^i} \quad \text{and} \quad X_y^i = -\frac{\partial f}{\partial x^i}.$$

Hence, the explicit expression of the Hamiltonian vector field of a function f in Darboux coordinates is

$$X_f|_U = \sum_{i=1}^n \left(\frac{\partial f}{\partial y^i}\frac{\partial}{\partial x^i} - \frac{\partial f}{\partial x^i}\frac{\partial}{\partial y^i}\right). \tag{5.1}$$

Thus we verified that the Hamiltonian vector field always exists. Furthermore, any Hamiltonian vector field of the function needs to satisfy this expression, implying they are also unique. □

Example 5.5. Take the manifold $\mathbb{R}^2 \setminus \{(0,p) : p \in \mathbb{R}\}$ with the global chart (q,p) and symplectic form $\omega = dq \wedge dp$. Consider the smooth function $f(q,p) = \frac{1}{2}p^2 - q^{-1}$, notice that q is never zero hence it is smooth. We can then obtain the Hamiltonian vector field of f from Equation 5.1.

$$X_f = \frac{\partial f}{\partial p}\frac{\partial}{\partial q} - \frac{\partial f}{\partial q}\frac{\partial}{\partial p} = p\frac{\partial}{\partial q} - q^{-2}\frac{\partial}{\partial p}.$$

We can check whether this is correct by entering it into the formula $\iota_{X_f}\omega(v) = df(v)$ for $v = v^q\,\partial/\partial q + v^p\,\partial/\partial p$.

$$\begin{aligned}\iota_{X_f}\omega(v) &= \omega(X_f, v) = dq \wedge dp\left(p\frac{\partial}{\partial q} - q^{-2}\frac{\partial}{\partial p}, v^q\frac{\partial}{\partial q} + v^p\frac{\partial}{\partial p}\right)\\ &= pv^p + q^{-2}v^q = v^q\frac{\partial f}{\partial q} + v^p\frac{\partial f}{\partial p} = df(v).\end{aligned}$$

Thus, X_f indeed satisfies $\iota_{X_f}\omega = df$, which implies it is the Hamiltonian vector field of f. //

We have shown that every function has a unique Hamiltonian vector field, giving us a linear mapping $C^\infty(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}) : f \to X_f$, where the linearity is a direct consequence of Proposition 3.21. The question now arises whether this mapping is injective or surjective. The answer to both is no.

Corollary 5.6. *Two smooth functions on a symplectic manifold have the same Hamiltonian vector field if and only if $f - g$ is a constant on each connected component of the manifold.*

Proof. Take some $f, g \in C^\infty(\mathcal{M})$, with $(\mathcal{M}, \omega)$ being a symplectic manifold. Suppose that their Hamiltonian vector fields be equal. We can then conclude that

$$df = \iota_{X_f}\omega = \iota_{X_g}\omega = dg.$$

Now suppose that their differentials are equal. Then we can conclude that

$$\iota_{X_f}\omega = df = dg = \iota_{X_g}\omega.$$

This implies, together with the uniqueness of Hamiltonian vector fields of functions, that $X_f = X_g$. □

Corollary 5.6 shows that the mapping $C^\infty(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}) : f \mapsto X_f$ induces a well-defined injective mapping $C^\infty(\mathcal{M})/Z^0(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}) : [f] \mapsto X_f$. However, this is still not necessarily a surjection, as we showcase with Example 5.7.

Example 5.7. Consider the manifold $\left(S^1 \setminus \{1\}\right)^2$ and take the global coordinates, (θ, ϕ). We can then define the symplectic form $\omega_0 = d\theta \wedge d\phi$ and the vector field $X = \partial/\partial\theta$. We can then calculate that $\iota_X\omega = d\phi$, implying that X is the Hamiltonian vector field of ϕ.

Suppose that we were able to smoothly extend X to a Hamiltonian vector field of some function $\mathbb{T}^2$. Notice that this function can not be ϕ as this is not defined on the whole of $\mathbb{T}^2$, hence suppose it is f. If follows that $df|_{\mathbb{T}^2\setminus\{(1,1)\}} = d\phi$. If we then integrate df over a closed curve γ, Stokes's theorem tells us the integral is 0. Meanwhile, take the curve $\gamma : [0,1] \to \mathbb{T}^2 : t \mapsto (0, \exp(2\pi t))$. We then calculate

$$\int_{[0,1]} \gamma^* df = \int_{[0,1]} \gamma^* d\phi = \int_{[0,1]} 2\pi dt = 2\pi \neq 0.$$

Hence, we cannot extend X to a Hamiltonian vector field. //

Luckily, we can quite easily determine the range of this mapping, i.e. all the vector fields that are the Hamiltonian vector field of some function.

Definition 5.8. A vector field X on a symplectic manifold $(\mathcal{M}, \omega)$ is called a **Hamiltonian vector field** if $\iota_X\omega$ is exact. The set of all Hamiltonian vector fields is denoted by $\mathfrak{X}_{\text{Ham}}(\mathcal{M})$. The function $h \in C^\infty(\mathcal{M})$ such that $\iota_X\omega = dh$ is called the **Hamiltonian function of** X. //

It should be clear that these Hamiltonian vector fields are exactly the vector fields that are the Hamiltonian vector field of some function. Furthermore, their flow preserves the structure of our symplectic manifold.

Proposition 5.9. *The symplectic form of a symplectic manifold is invariant under the flow of a Hamiltonian vector field on that same symplectic manifold.*

Proof. Suppose that $(\mathcal{M}, \omega)$ is a symplectic manifold and X is a Hamiltonian vector field with a Hamiltonian function h. If we rewrite the Lie derivative using Cartan's magic formula, we can use that ω is closed and $\iota_X \omega$ is exact. It then follows that

$$\mathcal{L}_X \omega = (\iota_X \circ d + d \circ \iota_X)\, \omega = \iota_X\, (d\omega) + d\,(\iota_X \omega) = \iota_X\, (0) + d \circ d\,(h) = 0.$$

This proves that the symplectic form is invariant under the flow of a Hamiltonian vector field. □

A well-behaved connection between the Hamiltonian vector fields and smooth functions can be made through their respective bracket algebras, the Lie algebra and the Poisson algebra.

5.1.3 Bracket Algebras

This section will focus on the algebraic properties of $\mathfrak{X}_{\text{Ham}}(\mathcal{M})$ and $C^\infty(\mathcal{M})$ and how we connect these through the view of bracket algebras. We will define abstract Lie algebras and Poisson algebras and see their connection with symplectic forms. We can get more grip on the behaviour of the mapping $C^\infty(\mathcal{M}) \to \mathfrak{X}_{\text{Ham}}(\mathcal{M}) : f \to X_f$ through these algebras.

Lie Algebras

It should be of no surprise that the algebraic structure of the Hamiltonian vector fields is familiar to that of vector fields in general, in other words, it has a Lie algebra structure. Let us first define what such a structure entails and then prove that the Hamiltonian vector fields are actually of this form.

Definition 5.10. A **Lie algebra** is a pair $(\mathfrak{L}, [\cdot,\cdot])$, such that $\mathfrak{L}$ is a real vector space and $[\cdot,\cdot] : \mathfrak{L} \times \mathfrak{L} \to \mathfrak{L}$ is a bilinear map such that the following hold for all $x, y, z \in \mathfrak{L}$

$$\begin{aligned} &[x, y] = -[y, x], \\ &[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0. \end{aligned}$$

These identities are respectively called skew-symmetry and the Jacobi identity. We call $[\cdot,\cdot]$ the **Lie bracket** of the Lie algebra. //

Proposition 5.11. *The bracket $[X, Y] = XY - YX$ defines a Lie bracket on the vector fields, and hence a Lie algebra $(\mathfrak{X}(\mathcal{M}), [\cdot,\cdot])$.*

Proof. See Corollary A.13. □

As the Hamiltonian vector fields are a subset of the vector fields on a symplectic manifold, we would like to show that it inherits a Lie algebraic structure. In other words, we would like to show that $(\mathfrak{X}_{\text{Ham}}(\mathcal{M}), [\cdot,\cdot])$ is a Lie subalgebra of $(\mathfrak{X}(\mathcal{M}), [\cdot,\cdot])$. In this proof, we will use the following lemma.

Lemma 5.12. *Let X, Y be vector fields on $\mathcal{M}$, then the following holds*

$$\iota_{[X,Y]} = \mathcal{L}_X \iota_Y - \iota_Y \mathcal{L}_X$$

Proof. The proof is similar to that of Cartan's magic formula. Once again we can notice that both sides are derivations of $\Omega(\mathcal{M})$ with degree -1. Hence, if we check the equality on smooth functions and exact 1-forms, we can extend it using induction to an arbitrary k-form.

Suppose that $f \in C^\infty(\mathcal{M})$ and $X, Y \in \mathfrak{X}(\mathcal{M})$. By then remarking that f and $\mathcal{L}_X f$ are both 0-forms, it follows from the definition of ι that both sides vanish.

Suppose that η is an exact 1-form, hence, there exists a smooth function $f \in C^\infty(\mathcal{M})$ such that $\eta = df$. By using the definitions of the interior multiplication and Lie derivative, and Proposition 3.33, it follows that

$$\begin{aligned}\iota_{[X,Y]}\eta &= \iota_{[X,Y]}df = df([X,Y]) = [X,Y]f = XYf - YXf = X(df(Y)) - d(Xf)(Y) \\ &= X(\iota_Y(df)) - \iota_Y(d(Xf)) = \mathcal{L}_X\iota_Y(df) - \iota_Y(d(\mathcal{L}_X f)) \\ &= \mathcal{L}_X\iota_Y(df) - \iota_Y\mathcal{L}_X(df) = (\mathcal{L}_X\iota_Y - \iota_Y\mathcal{L}_X)\eta.\end{aligned}$$

Hence, by induction, this shows that $\iota_{[X,Y]} = \mathcal{L}_X\iota_Y - \iota_Y\mathcal{L}_X$. □

Proposition 5.13. *The pair $(\mathfrak{X}_{Ham}(\mathcal{M}), [\cdot,\cdot])$, where $[\cdot,\cdot]$ is the Lie bracket on vector fields, defines a Lie algebra.*

Proof. Suppose that $(\mathcal{M},\omega)$ is a symplectic manifold. As $[\cdot,\cdot]$ is the Lie bracket on $\mathfrak{X}(\mathcal{M})$, it is clear that it is bilinear, skew-symmetric and satisfies the Jacobi identity. Furthermore, we can easily show that $\mathfrak{X}_{\text{Ham}}(\mathcal{M})$ is a real vector space as it is closed under addition and scalar multiplication in $\mathfrak{X}(\mathcal{M})$. Take arbitrary $X, Y \in \mathfrak{X}_{\text{Ham}}(\mathcal{M})$ and $a, b \in \mathbb{R}$, we can then write $\iota_X\omega = df$ and $\iota_Y\omega = dg$, it then follows using Propositions 3.21 and 3.26 that

$$\iota_{aX+bY}\omega = a\iota_X\omega + b\iota_Y\omega = adf + bdg = d(af + bg).$$

Thus $aX + bY \in \mathfrak{X}_{\text{Ham}}(\mathcal{M})$ implying that $\mathfrak{X}_{\text{Ham}}(\mathcal{M})$ is a real vector space.

We then only need to show that $\mathfrak{X}_{\text{Ham}}(\mathcal{M})$ is closed under the action of $[\cdot,\cdot]$. Take some arbitrary $X, Y \in \mathfrak{X}_{\text{Ham}}(\mathcal{M})$ and apply Lemma 5.12 and Cartan's magic formula.

$$\iota_{[X,Y]}\omega = \mathcal{L}_X\iota_Y\omega - \iota_Y\mathcal{L}_X\omega = d\iota_X\iota_Y\omega + \iota_X d\iota_Y\omega - \iota_Y d\iota_X\omega + \iota_Y\iota_X d\omega. \tag{5.2}$$

The fact that $\iota_X\omega$, $\iota_Y\omega$ and ω are closed, implies that the last three terms vanish, leaving just the first. Thus, Equation 5.2 simplifies to

$$\iota_{[X,Y]}\omega = d(\iota_X\iota_Y\omega) = -d(\omega(X,Y)). \tag{5.3}$$

Thus $\iota_{[X,Y]}\omega$ is exact and $[X,Y]$ is a Hamiltonian vector field. Together this implies that $(\mathfrak{X}_{\text{Ham}}(\mathcal{M}), [\cdot,\cdot])$ is indeed a Lie algebra. □

Poisson Algebras

Next up, we will introduce Poisson algebras. These are special cases of Lie algebras where we also assume that the space is a commutative associative algebra and compatibility of the algebraic structure with the Lie algebra structure. We will then show that the smooth functions on a symplectic manifold have such a structure.

Definition 5.14. A **Poisson algebra** is a Lie algebra $(\mathcal{P}, \{\cdot,\cdot\})$ such that $\mathcal{P}$ is a commutative associative algebra over $\mathbb{R}$ and the bracket satisfies the Leibniz rule.

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

The bracket $\{\cdot,\cdot\} : \mathcal{P} \times \mathcal{P} \to \mathcal{P}$ is called a **Poisson bracket**. //

We will now show that $C^\infty(\mathcal{M})$ admits a natural Poisson algebra structure on a symplectic manifold.

Proposition 5.15. *For a symplectic manifold $(\mathcal{M},\omega)$, the pair $(C^\infty(\mathcal{M}), \{\cdot,\cdot\})$ forms a Poisson algebra, with*

$$\{\cdot,\cdot\} : C^\infty(\mathcal{M}) \times C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M}) : (f,g) \mapsto \omega(X_f, X_g),$$

where X_f and X_g are the Hamiltonian vector fields of f and g respectively.

Proof. Let $(\mathcal{M}, \omega)$ be a symplectic manifold and define the bracket $\{\cdot,\cdot\}$ as in the proposition. It should be clear that $C^\infty(\mathcal{M})$ is a commutative associative algebra with pointwise multiplication.

The asymmetry is a consequence of the symplectic form being alternating, such that for $f, g \in C^\infty(\mathcal{M})$

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

To prove the bilinearity, we only need to do it in the first component as the second component then follows with the asymmetry. By the linearity of the mapping $f \mapsto X_f$, we find that

$$\begin{aligned}\{af + bg, h\} &= \omega(X_{af+bg}, X_h) = \omega(aX_f + bX_g, X_h) \\ &= a\omega(X_f, X_h) + b\omega(X_g, X_h) = a\{f, h\} + b\{g, h\}.\end{aligned}$$

To prove the Jacobi identity we will use the fact that ω is closed. Remark the following relation between the symplectic form and Hamiltonian vector fields:

$$\omega(X_f, Y) = \left(\iota_{X_f}\omega\right)(Y) = df(Y) = Yf. \tag{5.4}$$

It follows from the definition of the bracket that $\{f, g\} = X_g f$. We can then calculate the action of $d\omega$ on three vector fields using Proposition 14.32 in [13] and Equation 5.4. For some arbitrary functions $f, g, h \in C^\infty(\mathcal{M})$ we find

$$\begin{aligned}0 = d\omega(X_f, X_g, X_h) &= X_f\omega(X_g, X_h) - X_g\omega(X_f, X_h) + X_h\omega(X_f, X_g) \\ &- \omega([X_f, X_g], X_h) + \omega([X_f, X_h], X_g) - \omega([X_g, X_h], X_f).\end{aligned}$$

If we then set $\mathrm{I}(X_f, X_g, X_h) = X_f\omega(X_g, X_h) - X_g\omega(X_f, X_h) + X_h\omega(X_f, X_g)$, we can deduce that

$$\begin{aligned}\mathrm{I}(X_f, X_g, X_h) &= X_f\omega(X_g, X_h) - X_g\omega(X_f, X_h) + X_h\omega(X_f, X_g) \\ &= X_f\{g, h\} - X_g\{f, h\} + X_h\{f, g\} \\ &= \{\{g, h\}, f\} - \{\{f, h\}, g\} + \{\{f, g\}, h\} \\ &= -\{f, \{g, h\}\} - \{g, \{h, f\}\} - \{h, \{f, g\}\}.\end{aligned}$$

If we also take $\mathrm{II}(X_f, X_g, X_h) = -\omega([X_f, X_g], X_h) + \omega([X_f, X_h], X_g) - \omega([X_g, X_h], X_f)$, it follows that

$$\begin{aligned}\mathrm{II}(X_f, X_g, X_h) &= -\omega([X_f, X_g], X_h) + \omega([X_f, X_h], X_g) - \omega([X_g, X_h], X_f) \\ &= \omega(X_h, [X_f, X_g]) - \omega(X_g, [X_f, X_h]) + \omega(X_f, [X_g, X_h]) \\ &= [X_f, X_g]h - [X_f, X_h]g + [X_g, X_h]f \\ &= X_f X_g h - X_g X_f h - X_f X_h g + X_h X_f g + X_g X_h f - X_h X_g f \\ &= X_f\{h, g\} - X_g\{h, f\} - X_f\{g, h\} + X_h\{g, f\} + X_g\{f, h\} - X_h\{f, g\} \\ &= \{\{h, g\}, f\} - \{\{h, f\}, g\} - \{\{g, h\}, f\} + \{\{g, f\}, h\} + \{\{f, h\}, g\} - \{\{f, g\}, h\} \\ &= \{f, \{g, h\}\} + \{g, \{h, f\}\} + \{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} \\ &= 2\{f, \{g, h\}\} + 2\{g, \{h, f\}\} + 2\{h, \{f, g\}\}.\end{aligned}$$

Hence, it follows from the fact that $0 = d\omega(X_f, X_g, X_h) = \mathrm{I}(X_f, X_g, X_h) + \mathrm{II}(X_f, X_g, X_h)$ that

$$\begin{aligned}0 &= -\{f, \{g, h\}\} - \{g, \{h, f\}\} - \{h, \{f, g\}\} + 2\{f, \{g, h\}\} + 2\{g, \{h, f\}\} + 2\{h, \{f, g\}\} \\ &= \{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\}.\end{aligned}$$

This proves that the Jacobi identity holds. Moreover, Equation 5.4 implies the bracket satisfies the product rule. Take arbitrary $f, g, h \in C^\infty(\mathcal{M})$, then by Proposition A.8

$$\{f, gh\} = -X_f(gh) = -(X_f h)g - (X_f g)h = \{f, h\}g + \{f, g\}h.$$

We can conclude that $(C^\infty(\mathcal{M}), \{\cdot,\cdot\})$is a Poisson algebra, where the bracket is defined by the symplectic form. □

Proposition 5.16. *On a symplectic manifold $(\mathcal{M},\omega)$, the induced Poisson bracket is given locally in Darboux coordinates $\left(U,\left(x^1,\ldots,x^n,y^1,\ldots,y^n\right)\right)$ by*

$$\{f,g\}|_U=\sum_{i=1}^{n}\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial y^i}-\frac{\partial f}{\partial y^i}\frac{\partial g}{\partial x^i}. \tag{5.5}$$

Proof. Suppose that $(\mathcal{M},\omega)$ is a symplectic manifold and $\left(U,\left(x^1,\ldots,x^n,y^1,\ldots,y^n\right)\right)$ are some Darboux coordinates. The result follows from a calculation in coordinates where we use the coordinate expression of Hamiltonian vector fields of Equation 5.1 and the fact that $\{f,g\}=X_gf$ as we mentioned in the proof of Proposition 5.15

$$\{f,g\}|_U=X_g|_Uf=\left(\sum_{i=1}^{n}\frac{\partial g}{\partial y^i}\frac{\partial}{\partial x^i}-\frac{\partial g}{\partial x^i}\frac{\partial}{\partial y^i}\right)f=\sum_{i=1}^{n}\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial y^i}-\frac{\partial f}{\partial y^i}\frac{\partial g}{\partial x^i}.$$

This was the exact expression we wanted. □

Combining Bracket Algebras

We have seen that a symplectic manifold induces a Lie algebra, $\left(\mathfrak{X}_{\text{Ham}}\left(\mathcal{M}\right),[\cdot,\cdot]\right)$, and a Poisson algebra, $\left(C^\infty\left(\mathcal{M}\right),\{\cdot,\cdot\}\right)$. The connection between both is by the algebra anti-homomorphism created by taking the Hamiltonian vector field of a function.

Proposition 5.17. *Given a symplectic manifold $(\mathcal{M},\omega)$, we have an induced Lie algebra anti-homomorphism, $C^\infty\left(\mathcal{M}\right)\to\mathfrak{X}_{Ham}\left(\mathcal{M}\right):h\mapsto X_h$, with $\{\cdot,\cdot\}\rightsquigarrow-[\cdot,\cdot]$.*

Proof. Suppose that $(\mathcal{M},\omega)$ is a symplectic manifold. We then have an induced map $C^\infty\left(\mathcal{M}\right)\to\mathfrak{X}_{\text{Ham}}\left(\mathcal{M}\right):f\mapsto X_f$. To prove that this map is a Lie algebra anti-homomorphism, we argue it is enough to check whether

$$\iota_{X_{\{f,g\}}}\omega=-\iota_{[X_f,X_g]}\omega.$$

This is a consequence of the unicity of Hamiltonian vector fields and linearity of the interior multiplication in X.

We can then rewrite $\iota_{[X_f,X_g]}\omega$ if by using Equation 5.3 and get our result.

$$\iota_{[X_f,X_g]}\omega=-d\left(\omega\left(X_f,X_g\right)\right)=-d\{f,g\}=-\iota_{X_{\{f,g\}}}\omega.$$

This proves $X_{\{f,g\}}=-[X_f,X_g]$ and thus that the map is an anti-homomorphism of Lie algebras. □

5.2 Hamiltonian Systems

Now that we have seen the interplay between symplectic forms, smooth functions and vector fields, we can build their relation to physics. We will use symplectic manifolds in combination with a smooth function to create a model of classical physics, more specifically Hamiltonian mechanics. Such a triplet of a manifold, symplectic form and smooth function is what we will call a Hamiltonian system and it is what we will use to model classical mechanics.

Definition 5.18. A **Hamiltonian system** is a triplet $(\mathcal{M},\omega,\mathscr{H})$. Such that $(\mathcal{M},\omega)$ is a symplectic manifold and $\mathscr{H}\in C^\infty\left(\mathcal{M}\right)$ which is called the **Hamiltonian**. The associated Hamiltonian vector field $X_{\mathscr{H}}$ is called the **Hamiltonian phase flow** and its integral curves are called **trajectories**. //

Using these systems, we can quite easily recover the Hamiltonian equations. Suppose we have some Hamiltonian system $(\mathcal{M},\omega,\mathscr{H})$ and Darboux coordinates $\left(U,\left(x^1,\ldots,x^n,y^1,\ldots,y^n\right)\right)$. We see that a trajectory of the system $\gamma\left(t\right)=\left(x^i\left(t\right),y^i\left(t\right)\right)$ satisfies

$$\dot{x}^i\left(t\right)=\frac{\partial\mathscr{H}}{\partial y^i}\left(x,y\right)\quad\text{and}\quad\dot{y}^i\left(t\right)=-\frac{\partial\mathscr{H}}{\partial x^i}\left(x,y\right).$$

These are equivalent to Hamilton's equations we saw in Section 4.3. Let us showcase how we model physical systems using this method. In this method, we will most often consider a configuration space, $\mathcal{M}$, and then take the cotangent bundle with the canonical symplectic form as the symplectic form. This gives a clear distinction between position, the coordinates on $\mathcal{M}$, and the moment, the coordinates in the cotangent spaces. When setting up a Hamiltonian system as a model, we can still derive the Hamiltonian from the Lagrangian using a Legendre transform. In this case, the Lagrangians are defined as functions on the tangent bundle of the configuration space. The Legendre transform can then locally, at a point, transform the Lagrangian to a Hamiltonian, see [4, Chapter 20] for a rigorous construction of the Legendre transform. Here, we will work in coordinates, in which case the construction of Section 4.3 is sufficient. Let us consider some simple examples using Hamiltonian systems as a model.

Example 5.19. Let us consider a free particle of mass m moving through three-dimensional space. The configuration space is $\mathbb{R}^3$ and the phase space is the cotangent bundle with the canonical symplectic form $(\mathrm{T}^*\mathbb{R}^n, \omega_{\mathrm{can}})$. In the global coordinates generated by the global chart on $\mathbb{R}^3$ we can write the Lagrangian as

$$\mathscr{L}(x, \dot{x}) = \frac{1}{2} m \sum_i \dot{x}_i^2.$$

To transform this into a Hamiltonian, we first determine the generalised momenta of the Lagrangian.

$$p_i = \frac{\partial \mathscr{L}}{\partial \dot{x}_i}(x, \dot{x}) = m\dot{x}_i.$$

Hence, we can derive the Hamiltonian for this system using the Legendre transform

$$\mathscr{H}(x, p) = \sum_i \dot{x}_i p_i - \mathscr{L}(x, \dot{x}) = \frac{1}{2m} \sum_i p_i^2.$$

The Hamiltonian vector field is then given by

$$X_{\mathscr{H}} = \sum_i \frac{p_i}{m} \frac{\partial}{\partial q_i}.$$

The trajectories of the system are then the flow of the Hamiltonian phase flow. As the Hamiltonian phase flow is linear we find that the trajectories are as expected

$$\phi_{X_{\mathscr{H}}}((q, p), t) = (q + pt/m\ , p).$$

//

Example 5.20. Let us consider the example of a particle with charge q moving through a transversal constant magnetic field and some electric potential. Due to the transversality of the magnetic field, we can consider the system in just two dimensions: the ones perpendicular to the direction of the magnetic field. Let us adopt the coordinates x and y to describe the positions of the system. Hence, our configuration space is given by $\mathbb{R}^2$ and the phase space by $\left(\mathrm{T}^*\mathbb{R}^2, \omega_{\mathrm{can}}\right)$. We can derive from Example 4.6 that the potential energy can be expressed as

$$U(x, y, \dot{x}, \dot{y}) = V(x, y) - A(x, y)\dot{x} - B(x, y)\dot{y}.$$

In our case, we will choose the potential functions to be defined as

$$V(x, y) = qE\left(x^2 + y^2\right), \quad A(x, y) = \frac{My}{2}, \quad B(x, y) = -\frac{Mx}{2}. \tag{5.6}$$

This leads to the following Lagrangian for the system:

$$\mathscr{L}(x, y, \dot{x}, \dot{y}) = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2\right) - qE\left(x^2 + y^2\right) + \frac{M}{2} y\dot{x} - \frac{M}{2} x\dot{y}.$$

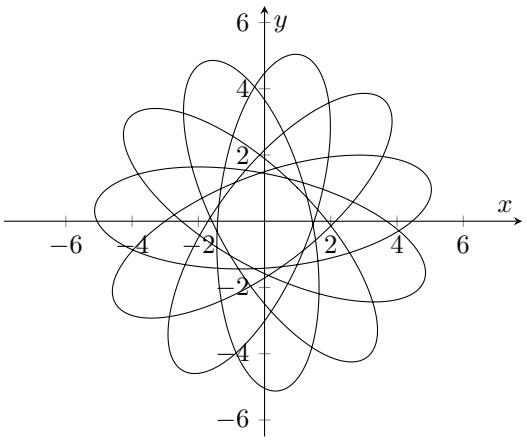


Figure 5.1: A plot of the (x, y)-trajectory of a charged particle in an electric and magnetic field, generated by the potentials in Equation 5.6. The equations in Equation 5.7 were solved numerically using the SciPy package in Python, with the initial values of $(x_0, y_0, p_{x0}, p_{y0}) = (1, 1, -2, 2)$ and $M = 0.1$, $q = 0.5$ and $E = 0.3$.

From the Lagrangian, we obtain the generalised momenta associated with x and y.

$$p_x = \frac{\partial \mathscr{L}}{\partial \dot{x}} = \dot{x} + \frac{My}{2}, \qquad p_y = \frac{\partial \mathscr{L}}{\partial \dot{y}} = \dot{y} - \frac{Mx}{2}.$$

Using the Legendre transform, we can deduce that the Hamiltonian is given by

$$\begin{aligned}
\mathscr{H}(x, y, p_x, p_y) &= \dot{x}p_x + \dot{y}p_y - \mathscr{L}(x, y, \dot{x}, \dot{y}), \\
&= p_x\left(p_x - \frac{My}{2}\right) + p_y\left(p_y + \frac{Mx}{2}\right) - \frac{1}{2}\left(p_x - \frac{My}{2}\right)^2 - \frac{1}{2}\left(p_y + \frac{Mx}{2}\right)^2 \\
&\quad + qE\left(x^2 + y^2\right) - \frac{M}{2}y\left(p_x - \frac{My}{2}\right) + \frac{M}{2}x\left(p_y + \frac{Mx}{2}\right), \\
&= \frac{1}{2}\left(\left(p_x - \frac{My}{2}\right)^2 + \left(p_y + \frac{Mx}{2}\right)^2\right) + qE\left(x^2 + y^2\right).
\end{aligned}$$

Remark that this is not the total energy, yet we recover the equations of motion from the Hamiltonian vector field

$$X_{\mathscr{H}}(d\alpha) = \begin{cases} p_x - \dfrac{My}{2}, & \alpha = x, \\ p_y + \dfrac{Mx}{2}, & \alpha = y, \\ -\left(p_y + \dfrac{Mx}{2}\right)\dfrac{M}{2} - 2qEx, & \alpha = p_x, \\ \left(p_x - \dfrac{My}{2}\right)\dfrac{M}{2} - 2qEy, & \alpha = p_y. \end{cases} \tag{5.7}$$

It is quite tricky to solve for the flow of this system analytically. Luckily, we can easily solve it numerically, see Figure 5.1 //

5.3 Conserved Quantities

The last section introduced our basic model for classical physics: Hamiltonian systems. We will now investigate a special class of functions which are conserved over the trajectories of the system.

Definition 5.21. An **first integral**, also called a **conserved quantity** or **integral of motion**, of a Hamiltonian system $(\mathcal{M}, \omega, \mathscr{H})$ is a function $f \in C^\infty(\mathcal{M})$ that is constant along all trajectories of the system. //

It can be quite cumbersome to check this condition, as one would need to know the exact flow of the Hamiltonian phase flow. Luckily, we can use the connection between $C^\infty(\mathcal{M})$ and the Hamiltonian vector fields to get an easier check for first integrals.

Proposition 5.22. *If $(\mathcal{M}, \omega, \mathscr{H})$ is a Hamiltonian system, then a function $f \in C^\infty(\mathcal{M})$ is a first integral of the system if and only if $\{f, \mathscr{H}\} = 0$.*

Proof. Let $(\mathcal{M}, \omega, \mathscr{H})$ be an Hamiltonian system and $f \in C^\infty(\mathcal{M})$. Using Proposition 3.34 and Equation 5.4, we deduce that the change of f along the flow is given by

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} \left(f \circ \phi^t_{X_\mathscr{H}}\right) = \left(\phi^{t_0}_{X_\mathscr{H}}\right)^* \mathcal{L}_{X_\mathscr{H}} f = \left(\phi^{t_0}_{X_\mathscr{H}}\right)^* \iota_{X_\mathscr{H}} df = \left(\phi^{t_0}_{X_\mathscr{H}}\right)^* \{f, \mathscr{H}\}.$$

This proves the equivalence of the two statements. □

Using Proposition 5.22, we can easily verify whether a function is a first integral of a system.

5.3.1 Symmetries

We will show that these first integrals give quite a lot of information about the physical system in terms of symmetries. Here a symmetry of a Hamiltonian system is defined as an infinitesimal symmetry, i.e. a vector field under whose flow the Hamiltonian and symplectic form are invariant.

Definition 5.23. A vector field is an **infinitesimal symmetry** of a Hamiltonian system $(\mathcal{M}, \omega, \mathscr{H})$ if both ω and $\mathscr{H}$ are invariant under its flow. //

Lemma 5.24. *A vector field X is an infinitesimal symmetry of $(\mathcal{M}, \omega, \mathscr{H})$ if and only if $X\mathscr{H} = 0$ and $\iota_X\omega$ is closed.*

Proof. Let X be a vector field on a Hamiltonian system, $(\mathcal{M}, \omega, \mathscr{H})$. Suppose that X is an infinitesimal symmetry of the Hamiltonian system, then by Theorem 12.37 in [13] we know that the invariance under the flow is equivalent to

$$\mathcal{L}_X \omega = 0 = \mathcal{L}_X \mathscr{H}. \tag{5.8}$$

We can easily calculate the action of X on a Hamiltonian system by the definition of the Lie derivative on smooth functions and Equation 5.8.

$$X\mathscr{H} = \mathcal{L}_X \mathscr{H} = 0.$$

Furthermore, as the symplectic form is closed we can rewrite $d \circ \iota_X\omega$ to a Lie derivative using Cartan's magic formula, such that

$$d \circ \iota_X \omega = (d \circ \iota_X + \iota_X \circ d)\,\omega = \mathcal{L}_X \omega = 0.$$

This shows the implication one way. For the other way, suppose that $X\mathscr{H} = 0$ and $\iota_X\omega$ is closed. Calculate the Lie derivative of $\mathscr{H}$ along X

$$\mathcal{L}_X \mathscr{H} = (d \circ \iota_X + \iota_X \circ d)\,\mathscr{H} = d\mathscr{H}(X) = Xh = 0.$$

Thus $\mathscr{H}$ is invariant under the flow of X. As for ω, we can do a similar calculation and use the fact that ω and $\iota_X\omega$ are both closed

$$\mathcal{L}_X \omega = (d \circ \iota_X + \iota_X \circ d)\,\omega = d(\iota_X \omega) = 0.$$

This shows that the implication holds the other way as well. □

We will now show that there is a correspondence between the first integrals of a Hamiltonian system and the possible infinitesimal symmetries. This is a variation of Noether's theorem, which is most often stated in terms of moment maps and Lie group actions, see Chapter 24 in [4]. The following theorem in terms of infinitesimal symmetries is Theorem 22.22 in [13]

Theorem 5.25. *Let $(\mathcal{M}, \omega, \mathscr{H})$ be a Hamiltonian system. If f is a first integral of the system, then its Hamiltonian vector field is an infinitesimal symmetry of the system. Furthermore, if all closed 1-forms are exact, then all infinitesimal symmetries are generated as the Hamiltonian vector field of some function, which is unique up to some function that is constant on each component.*

Proof. Let $(\mathcal{M}, \omega, \mathscr{H})$ be a Hamiltonian system and $f \in C^\infty(\mathcal{M})$ a first integral of this system. The Hamiltonian is invariant under the flow of X_f by definition and ω is invariant under the flow of any Hamiltonian vector field, see Proposition 5.9.

Now suppose that all closed 1-forms on $\mathcal{M}$ are exact and let X be an infinitesimal symmetry of $(\mathcal{M}, \omega, \mathscr{H})$. If follows from Lemma 5.24 that $X\mathscr{H} = 0$ and that $\iota_X \omega$ is closed. By assumption, this form is also exact and hence there is a function f such that $X = X_f$. We can then easily deduce that it commutes with the Hamiltonian

$$\{\mathscr{H}, f\} = X_f \mathscr{H} = X\mathscr{H} = 0.$$

Thus f is a first integral of the Hamiltonian system. Remark that f is defined uniquely up to some element of $Z^0(\mathcal{M})$ as in Proposition 5.6. □

Example 5.26. Let us describe a spherical pendulum. This consists of a rigid rod of unit length and negligible mass which has one point fixed in space around which it is free to rotate. An object of unit mass is attached to the other end of the rod and is under the influence of a constant gravitational field, whose acceleration constant we set to 1. No other external forces are acting on the system. The position of the object can then be described as a point on $\mathbb{S}^2$, which is therefore the configuration space. The phase space is then given by $\left(\mathrm{T}^*\mathbb{S}^2, \omega_{\mathrm{can}}\right)$.

In Cartesian coordinates, we can describe the Lagrangian as

$$\mathscr{L}(x, y, z, \dot{x}, \dot{y}, \dot{z}) = \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{2} - z.$$

We can transform these to the more natural coordinates θ and ϕ by the usual spherical coordinate transformation,

$$\begin{aligned} x &= \sin(\theta)\cos(\phi), & \dot{x} &= \cos(\theta)\cos(\phi)\dot{\theta} - \sin(\theta)\sin(\phi)\dot{\phi} \\ y &= \sin(\theta)\sin(\phi), & \dot{y} &= \cos(\theta)\sin(\phi)\dot{\theta} + \sin(\theta)\cos(\phi)\dot{\phi} \\ z &= \cos(\theta), & \dot{z} &= -\sin(\theta)\dot{\theta}. \end{aligned}$$

Using this transformation, we can write the Lagrangian and the associated general momenta as

$$\mathscr{L}\left(\theta, \phi, \dot{\theta}, \dot{\phi}\right) = \frac{\dot{\theta}^2 + \sin^2(\theta)\dot{\phi}^2}{2} - \cos(\theta), \quad p_\theta = \frac{\partial \mathscr{L}}{\partial \dot{\theta}} = \dot{\theta}, \quad p_\phi = \frac{\partial \mathscr{L}}{\partial \dot{\phi}} = \sin^2(\theta)\dot{\phi}.$$

This lets us determine the Hamiltonian as

$$\mathscr{H}(\theta, \phi, p_\theta, p_\phi) = p_\theta \dot{\theta} + p_\phi \dot{\phi} - \mathscr{L} = \frac{p_\theta^2 + \csc^2(\theta) p_\phi^2}{2} + \cos(\theta).$$

Now take the function $J = p_\phi$, we can deduce that this is a first integral of the system using Proposition 5.22 and the fact that the Poisson bracket is given by Equation 5.5

$$\{\mathscr{H}, J\} = \frac{\partial \mathscr{H}}{\partial \theta}\frac{\partial J}{\partial p_\theta} - \frac{\partial \mathscr{H}}{\partial p_\theta}\frac{\partial J}{\partial \theta} + \frac{\partial \mathscr{H}}{\partial \phi}\frac{\partial J}{\partial p_\phi} - \frac{\partial \mathscr{H}}{\partial p_\phi}\frac{\partial J}{\partial \phi}.$$

As J is only dependent on p_ϕ and $\mathscr{H}$ is independent of ϕ, it follows that $\{\mathscr{H}, J\} = 0$ and thus J is a first integral of the Hamiltonian system. //

5.3.2 Integrable Systems

Next up, we will research the solvability of a system depending on the first integrals. We already saw that we were able to find the Hamiltonian phase flow of some systems, see Examples 5.19 and 5.20, but we could not always easily solve these equations. We will see that a system is solvable if sufficient and well-behaved first integrals exist. We will see that we can find suitable coordinates on such systems in which the motions are trivial and that the coordinates can be constructed by quadratures. In this section, we will follow Chapter 10 of [1], however, more general theorems are also found in Chapter 18 of [4].

Definition 5.27. A Hamiltonian system $(\mathcal{M}, \omega, \mathscr{H})$ is called an **integrable system** if there exist $n = \frac{1}{2}\dim\mathcal{M}$ first integrals, $f_1 = \mathscr{H}, f_2, \ldots, f_n$, that commute, $\{f_i, f_j\} = 0$ for all i and j, and have linearly independent differentials. Such a system is denoted as $(\mathcal{M}, \omega, f)$, where $f = (\mathscr{H}, f_2, \ldots, f_n)$. We abbreviate the flow of the Hamiltonian vector field of f_i as $\phi_{X_{f_i}} = \phi_i$ and also abbreviate $\phi^{(p)}_{X_{f_i}} = \phi^{(p)}_i$ and $\phi^t_{X_{f_i}} = \phi^t_i$. //

Remark. In some other texts, [4] for example, the differentials of an integrable system are required to be linearly independent on a dense subset instead of the whole manifold. Here, we follow [1] which defines them to be linearly independent everywhere. //

Given an integrable system, we would like to show that we can determine the trajectories of the Hamiltonian system. To do this, we will consider the following level sets of f,

$$\mathcal{M}_c = f^{-1}(c) = \{p \in \mathcal{M} : \ \forall i \ f_i(p) = c_i\}.$$

Remark that this is a regular level set as f has a non-singular differential. This set has some nice properties.

Proposition 5.28. *Let $(\mathcal{M}, \omega, f)$ be an integrable system, then $\mathcal{M}_c$ is a smooth manifold that is invariant under the flow of any of the first integrals.*

Proof. It is clear that $\mathcal{M}_c$ is a manifold as it is a regular level set of f. Furthermore, we know that any first integral f_i is invariant under the flow of another f_j, and thus for any $p \in \mathcal{M}_c$, we can deduce that

$$f_i\left(\phi^t_j(p)\right) = \left(\phi^t_j\right)^* f_i(p) = f_i(p) = c_i.$$

This implies that $\phi^t_j : \mathcal{M}_c \to \mathcal{M}_c$, and thus that $\mathcal{M}_c$ is invariant under the flow of the first integrals. □

As the flows of all the first integrals commute, we can sometimes define a general global flow of an integrable system.

Definition 5.29. If $(\mathcal{M}, \omega, f)$ is an integrable system for which the Hamiltonian vector fields of each first integral are complete on $\mathcal{M}_c$, we define the **simultaneous global flow** on $\mathcal{M}_c$ as

$$g_f : \mathbb{R}^n \times \mathcal{M}_c \to \mathcal{M}_c : (t, p) \mapsto \left(\phi^{t_1}_1 \circ \cdots \circ \phi^{t_n}_n\right)(p).$$

We then also define $g^{(p)}_f : \mathbb{R}^n \to \mathcal{M}_c : t \mapsto g_f(t, p)$ and $g^t_f : \mathcal{M}_c \to \mathcal{M}_c : p \mapsto g_f(t, p)$, similar to $\phi^{(t)}_X$ and ϕ^t_X in Definition A.21. //

By examining the behaviour of this flow we can uncover more information about the level set. We will focus our attention on level sets that are compact and connected as these are of physical interest. However, many results can be generalised to any system as long as all the flows are complete, as this ensures the existence of the simultaneous flow, see [4] for more details on this.

Lemma 5.30. *Let $(\mathcal{M}, \omega, f)$ be an integrable system and $\mathcal{M}_c$ is compact and connected, then the simultaneous flow $g^{(p)}_f$ is surjective, but not injective.*

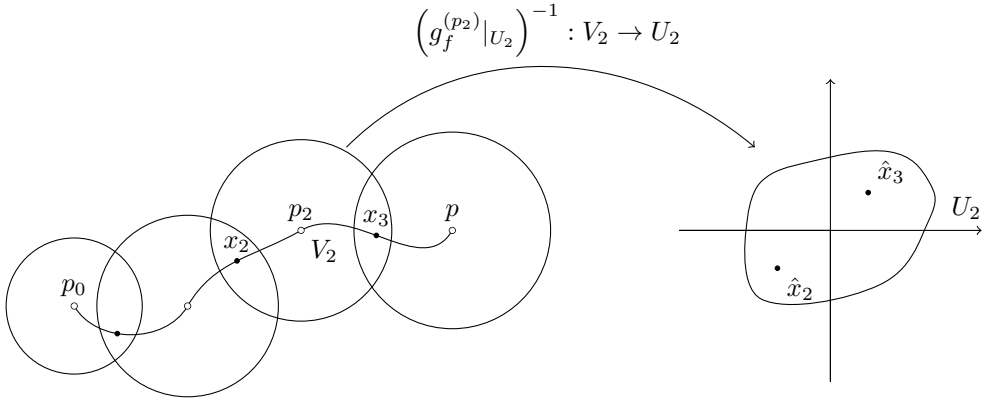


Figure 5.2: Here an example of the path connecting two points p_0 and p and how we cover this. Here, we also see how $\left(g_f^{(p_2)}|_{U_2}\right)^{-1}$ maps elements from V_2 to U_2. Here, an element mapped by $\left(g_f^{(p_2)}|_{U_2}\right)^{-1}$ is denoted with a hat. We could also draw t_2 as the vector between $\hat{x}_2$ and $\hat{x}_3$.

Proof. Suppose we have some integrable system $(\mathcal{M}, \omega, f)$ and some level set $\mathcal{M}_c$ which is compact and connected. Remark that the flows of the first integrals then act on a compact manifold and are therefore complete. We can thus define the simultaneous flow g_f and let us fix some $p_0 \in \mathcal{M}_c$. Remark that $g_f^{(p_0)}$ can not be a bijection as it is continuous, $\mathbb{R}^n$ is Hausdorff but not compact, and $\mathcal{M}_c$ is compact. If it was a bijection, this would imply it is a homeomorphism such that $\mathbb{R}^n$ would be compact. Hence, it is enough to show that $g_f^{(p_0)}$ is surjective.

Firstly, we will show that we can take some small steps using $g_f^{(p)}$ for any $p \in \mathcal{M}$ to somewhere in the neighbourhood of p. Remark that $dg_f^{(p)}$ is non-singular as the differentials of the first integrals are linearly independent. Hence, by the inverse function theorem, there exist some neighbourhoods $V \subset \mathcal{M}_c$ and $U \subset \mathbb{R}^n$ such that $(g_f^{(p)}|_U)^{-1} : V \to U$ exists and defines a chart. For any $q \in V$ we can then define $t = (g_f^{(p)}|_U)^{-1}(q)$ such that

$$g_f^{(p)}(t) = g_f^{(p)} \circ \left(g_f^{(p)}|_U\right)^{-1}(q) = q.$$

Suppose that q is an arbitrary point on $\mathcal{M}_c$. We know by the connectedness of $\mathcal{M}_c$ that there exists a continuous path $\gamma : [0,1] \to \mathcal{M}_c$ such that $\gamma(0) = p_0$ and $\gamma(1) = q$. Because $\gamma([0,1])$ is compact, we can find $\{p_i\}_{i=0}^n$, with $p_n = q$, such that for each p_i there exists a neighbourhood $V_i \subset \mathcal{M}_c$ of p and neighbourhood $U_i \subset \mathbb{R}^n$ of 0, such that $g_f^{(p_i)}|_{U_i} : U_i \to V_i$ is a diffeomorphism. Furthermore, we can assume that $V_{i-1} \cap V_i \neq \emptyset$. See Figure 5.2 for a visualisation of this process. Let us now choose some points $\{x_i\}_{i=0}^{m+1}$, such that $x_0 = p_0$, $x_i \in V_{i-1} \cap V_i$ and $x_{m+1} = q \in V_m$.

We can then step from x_i to x_{i+1} by using the flow of p_i. Let us define a time step in U_i as $t_i = (g_f^{(p_i)})^{-1}(x_{i+1}) - (g_f^{(p_i)})^{-1}(x_i)$, such that

$$g_f^{(x_i)}(t_i) = g_f^{(x_i)}\left((g_f^{(p_i)})^{-1}(x_{i+1}) - (g_f^{(p_i)})^{-1}(x_i)\right) = x_{i+1}.$$

Hence, we can walk from x_i to x_{i+1} with a time step t_i. If we then add all of these steps together, $t = \sum_{i=0}^m t_i$, we see that this lets us walk from $x_0 = p_0$ to $x_{m+1} = q$ using $g_f^{(p_0)}$. This implies that there exists some t such that $g_f^{(p_0)}(t) = q$, and thus that $g_f^{(p_0)}$ is surjective. □

Theorem 5.31. *Let $(\mathcal{M}, \omega, f)$ be an integrable system. If $\mathcal{M}_c$ is compact and connected, then $\mathcal{M}_c$ is diffeomorphic to $\mathbb{T}^n$.*

Proof. Let $(\mathcal{M}, \omega, f)$ be an integrable system such that $\mathcal{M}_c$ is compact and connected and take some $p \in \mathcal{M}_c$. By Lemma 5.30 we know that $g_f^{(p)}$ is not injective, such that we can consider its

stable points defined as

$$\Gamma(p) = \left\{ t \in \mathbb{R}^n : \; g_f^{(p)}(t) = p \right\}.$$

We can easily check that this is a subgroup of $\mathbb{R}^n$ as the simultaneous flow is also a one-parameter group action. Furthermore, we can prove that it is independent of p. Given any $q \in \mathcal{M}_c$, we know there is some $s \in \mathbb{R}^n$ such that $g_f^{(p)}(s) = q$. Thus, for any $t \in \Gamma(p)$ we deduce that

$$g_f^t(q) = g_f^t \circ g_f^s(p) = g_f^s \circ g_f^t(p) = g_f^s(p) = q.$$

This implies that $\Gamma(p) = \Gamma(q)$, such that we can define some unique Γ associated with g_f. This Γ has the structure of a discrete subgroup of $\mathbb{R}^n$. In other words, any point in Γ has a neighbourhood in the subspace topology which only contains itself.

We will first show the existence of such a neighbourhood around $0 \in \Gamma$ and we will use translations of this neighbourhood to show that it works for any $t \in \Gamma$. First choose an arbitrary $p \in \mathcal{M}_c$, as $g_f^{(p)}(0) = \mathrm{Id}$ we know that $0 \in \Gamma$. Now suppose U is a neighbourhood on which $g_f^{(p)}|_U$ is a diffeomorphism. For an arbitrary $s \in U \cap \Gamma$ we know that $g_f^{(p)}|_U(s) = p_0 = g_f^{(p)}|_U(0)$ as $s \in \Gamma$. By the injectivity of $g_f^{(p)}|_U$ it follows that $s = 0$ and thus $U \cap \Gamma = \{0\}$. We can generalise this to any element of Γ by translating it using the flow. Take an arbitrary $t \in \Gamma$ and define $q = g_f^{(p)}(t)$. We can then use the same neighbourhood U of 0 such that $g_f^{(q)}|_U$ is a diffeomorphism and $U \cap \Gamma = \{0\}$. This neighbourhood can be transformed back to a neighbourhood around t using $(g_f^{(p)})^{-1} \circ g_f^{(q)}$. This shows that Γ is a discrete subgroup of $\mathbb{R}^n$ and remark that we could choose a single neighbourhood U such that each $t \in \Gamma$ satisfies $(t + U) \cap \Gamma = \{t\}$.

We can now show that this group is generated by some basis $\{e_1, \ldots, e_k\}$ such that any element of Γ is a unique integral linear combination of this basis. If $\Gamma = \{0\}$, this would be trivial, else we can take a $x_1 \in \Gamma$ such that $x_1 \neq 0$. Consider the set $\Delta_1 = \{t \in \mathbb{R}^n : \|t\| \leq \|x_1\|\}$. Remark that we showed that there exists some neighbourhood U of 0 such that $(t + U) \cap \Gamma = \{t\}$. Hence, any element $t \in \Delta_1 \cap \Gamma$ is covered by $t + U$ and as the volume of Δ_1 is finite, we can conclude that $\Delta_1 \cap \Gamma$ only contains finite elements. We can then choose $e_1 \in \mathbb{R}x_1 \setminus \{0\}$ such that it closest to 0. Let us show that $\mathbb{Z}e_1 = \Gamma \cap \mathbb{R}x_1$. Suppose this is not the case, hence, there exists some $u_1 \in (\Gamma \cap \mathbb{R}x_1) \setminus \mathbb{Z}e_1$. Then there must also exists some $m \in \mathbb{Z}$ such that $u_1 \in \{(m + t)\, e_1 : \; t \in (0, 1)\}$, i.e. there exists some $t \in (0, 1)$ and $m \in \mathbb{Z}$ such that $u_1 = (m + t)\, e_1$. It then follows for these t and m that

$$\|u_1 - me_1\| = \|(m + t)\, e_1 - me_1\| = \|te_1\| = t\|e_1\|.$$

As $t \in (0, 1)$ it follows that $u_1 - me_1$ is closer to 0 than e_1 but is also not the zero vector. This is a contradiction with the construction of e_1 and thus we can conclude that $\mathbb{Z}e_1 = \Gamma \cap \mathbb{R}x_1$.

If $\mathbb{Z}e_1 = \Gamma$ we are done, else we can choose an $x_2 \in \Gamma \setminus \mathbb{Z}e_1$. Let m be the integer such that the projection of x_2 onto $\mathbb{R}e_1$ lies within $A_2 = \{te_1 : 0 \leq \mathrm{sign}(m)\, t \leq |m|\}$. If $\mathbb{Z}e_1 = \Gamma$, we are done, else we can find an $x_2 \in \Gamma \setminus \mathbb{Z}e_1$. Then there must exist some $m \in \mathbb{Z}$ such that the projection of x_2 onto $\mathbb{R}e_1$ lies within $A_1 = \{te_1 : \; t \in \mathbb{R},\; 0 \leq t \leq m\}$. Now let Δ_2 denote the set of points whose projections onto $\mathbb{R}e_1$ also lie in A_2 and whose distance to A_2 is smaller than that of x_2, see Figure 5.3. Remark that this set has some finite volume and thus contains finite points of Γ as we discussed above. We can therefore choose come e_2 in Δ_2 which is closest to $\mathbb{R}e_1$ but not on it, remark that this choice is not unique. We now want to show that $\mathbb{Z}e_1 + \mathbb{Z}e_2 = (\mathbb{R}e_1 + \mathbb{R}e_2) \cap \Gamma$. Suppose that this is not true, we can then find an $u_2 \in (\Gamma \cap (\mathbb{R}e_1 + \mathbb{R}e_2)) \setminus (\mathbb{Z}e_1 + \mathbb{Z}e_2)$. Remark that there exist $m_1, m_2 \in \mathbb{Z}$ such that $u_2 \in \{(m_1 + t)\, e_1 + (m_2 + s)\, e_2 : t, s \in (0, 1)\}$, i.e. there exist some $t, s \in (0, 1)$ and $m_1, m_2 \in \mathbb{Z}$ such that $u_1 = (m_1 + t)\, e_1 + (m_2 + s)\, e_2$. It follows that the distance from $u_1 - m_1 e_1 - m_2 e_2$ would be closer to $\mathbb{R}e_1$ than e_2 and we can move this vector to one in Δ_2 by adding some multiple of e_1 while not changing the distance to $\mathbb{R}e_1$. This would lead to a contradiction with the construction of e_2 and we can therefore conclude that $\mathbb{Z}e_1 + \mathbb{Z}e_2 = \Gamma \cap (\mathbb{R}e_1 + \mathbb{R}e_2)$.

Again if this covers the whole of Γ, we would be done, else we can repeat the process but look for a linearly independent vector closest to a higher dimensional plane, i.e. e_3 would be the closest point to $\mathbb{R}e_1 + \mathbb{R}e_2$ and e_4 would be closest to $\mathbb{R}e_1 + \mathbb{R}e_2 + \mathbb{R}e_3$, etc. We iterate this process

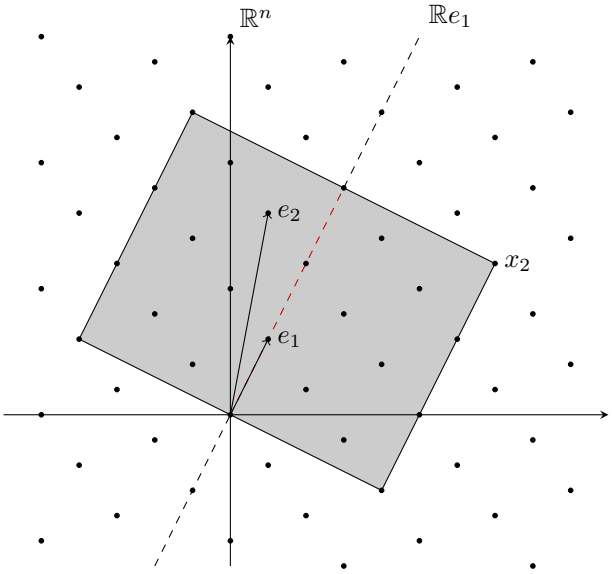


Figure 5.3: A plot of a Lattice with the dots representing the lattice points. Shown are the already chosen vector e_1 and the line $\mathbb{R}e_1$ this generates, which is drawn with a dashed line. An arbitrary point $x_2 \in \Gamma \backslash \mathbb{R}e_1$ is chosen, for which A_2 is drawn in red and Δ_2 as the shaded area. We can see that this contains finite points, and one point in $\Gamma \backslash \mathbb{R}e_1$ closest to $\mathbb{R}e_1$ is marked as e_2. Remark that this choice is not unique.

to get a set $\{e_1, \ldots, e_k\}$. As this is a set of independent vectors, this process must terminate as the dimension of $\mathbb{R}^n$ is finite. Thus we conclude that there must exist some $0 \leq k \leq n$ and a basis $\{e_1, \ldots, e_k\}$ for Γ.

Using this basis, we will generate a diffeomorphism between $\mathbb{T}^k \times \mathbb{R}^{n-k}$ and $\mathcal{M}_c$. This diffeomorphism should make the diagram in Figure 5.4 commute. To show the existence of this diffeomorphism, we must of course first define the functions ρ and F.

Let us first define $\rho : \mathbb{R}^n \to \mathbb{T}^k \times \mathbb{R}^{n-k}$. Remark that we can write a vector $v \in \mathbb{R}^n$ as a vector $v = (\phi, y) \in \mathbb{R}^k \times \mathbb{R}^{n-k}$. Notice that we have a surjective mapping $\tilde{\rho} : \mathbb{R}^k \to \mathbb{T}^k$ given by

$$\tilde{\rho}\left(\phi_1, \ldots, \phi_k\right) = \left(\phi_1 \mod 1, \ldots, \phi_k \mod 1\right).$$

Remark that this function is well-defined as S^1 is diffeomorphic to $\mathbb{R}/\mathbb{Z}$. This mapping lets us define a mapping $\rho : \mathbb{R}^n \to \mathbb{T}^k \times \mathbb{R}^{n-k} : (\phi, y) \mapsto (\tilde{\rho}(\phi), y)$. Remark that ρ is a local diffeomorphism.

Next up, we can go on and define F. First, we choose the standard basis for $\mathbb{R}^n$ and denote it as $\beta = \{f_1, \ldots, f_n\}$. Let $\{e_1, \ldots, e_n\}$ denote an extension of the basis $\{e_1, \ldots, e_k\}$ of Γ with vectors from β, see [8, Theorem 1.10]. Remark that the first k vectors in this basis are still the basis for Γ. We can then define a linear mapping F as the basis transformation $F(f_i) = e_i$. By definition, this mapping is bijective and as it is linear also a diffeomorphism.

Using Theorem 4.30 in [13], we can show that there must exist a diffeomorphism $\tilde{F}$ which makes the diagram in Figure 5.4 commute. First of all, remark that ρ and $g_f^{(p_0)}$ are smooth submersions as they are both local diffeomorphisms, see [13, Proposition 4.8]. The existence of smooth maps $\tilde{F}$ and $\tilde{F}^{-1}$ with the properties that $\tilde{F} \circ \rho = g_f^{(p_0)} \circ F$ and $\tilde{F}^{-1} \circ g_f^{(p_0)} = \rho \circ F^{-1}$ is then a consequence of the fact that the definition of F implies that $g_f^{(p_0)} \circ F$ and $\rho \circ F^{-1}$ are constant on the fibres of respectively ρ and $g_f^{(p_0)}$. We now have to show that $\tilde{F}^{-1}$ is indeed the inverse of $\tilde{F}$, here we will use that ρ and $g_f^{(p_0)}$ are surjective such that they both have a right inverse, which we denote with ρ^{-1} and $\left(g_f^{(p_0)}\right)^{-1}$. We can see that $\tilde{F}^{-1}$ is the right inverse of

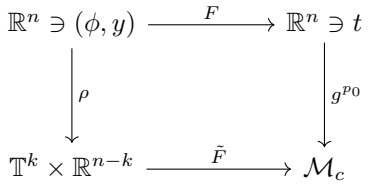


Figure 5.4: Commutative diagram which describes the relation between $F : \mathbb{R}^n \to \mathbb{R}^n$ and $\tilde{F} : \mathbb{T}^k \times \mathbb{R}^{n-k}$ by the natural mapping ρ and the simultaneous flow $g_f^{(p_0)}$.

$\tilde{F}$ using the property that they make the diagram of Figure 5.4 commute.

$$\begin{aligned}\tilde{F} \circ \tilde{F}^{-1} &= \tilde{F} \circ \tilde{F}^{-1} \circ g_f^{(p_0)} \circ \left(g_f^{(p_0)}\right)^{-1} = \tilde{F} \circ \rho \circ F \circ \left(g_f^{(p_0)}\right)^{-1} \\ &= g_f^{(p_0)} \circ F^{-1} \circ F \circ \left(g_f^{(p_0)}\right)^{-1} = \mathrm{Id}_{\mathcal{M}_c}\,.\end{aligned}$$

Similarly, we can deduce that $\tilde{F}^{-1}$ is the left inverse of $\tilde{F}$.

$$\tilde{F}^{-1} \circ \tilde{F} = \tilde{F}^{-1} \circ \tilde{F} \circ \rho \circ \rho^{-1} = \tilde{F}^{-1} \circ g_f^{(p_0)} \circ F^{-1} \circ \rho^{-1} = \rho \circ F \circ F^{-1} \circ \rho^{-1} = \mathrm{Id}_{\mathbb{T}^k \times \mathbb{R}^{n-k}}\,.$$

Hence, it follows that there exists a diffeomorphism $\tilde{F}$ between $\mathbb{T}^k \times \mathbb{R}^{n-k}$ and $\mathcal{M}_c$. Now remark that $\mathcal{M}_c$ is compact, implying that $n = k$ which was the result we wanted. □

This shows that the trajectories of the integrable system all lie on $\mathcal{M}_c$ and that this space is homeomorphic to some torus. Furthermore, the maps in this proof let us define coordinates on $\mathcal{M}_c$ such that our simultaneous flow is of a simple form.

Corollary 5.32. *There exists a chart on a compact and connected $\mathcal{M}_c$ such that the flows of the first integrals are linear.*

Proof. Let $(\mathcal{M}, \omega, f)$ be an integrable system and let $\mathcal{M}_c$ be compact and connected. Define F, ρ and $\tilde{F}$ as in the proof of Theorem 5.31. Let us now consider the function $(g_f^{(p_0)} \circ F)^{-1}$, where $p_0 \in \mathcal{M}_c$. We can define a chart around p_0 as $g_f^{(p_0)}$ is a local diffeomorphism. The representation of g_f in these coordinates can be determined using the bases f_i and e_i for $\mathbb{R}^n$ as defined in the proof of Theorem 5.31

$$\left(F^{-1} \circ \left(g_f^{(p_0)}\right)^{-1} \circ g_f^t \circ g_f^{(p_0)} \circ F\right)\left(\sum_i \phi_i f_i\right) F^{-1}\left(\sum_i \phi_i e_i + t\right) = \sum_i \phi_i f_i + F^{-1}(t)\,.$$

Hence, the flows of the first integrals are all linear. From this, we can easily deduce that the flow associated with a first integral is linear. □

These coordinates are what we call **angle coordinates** as they correspond to the angles on the torus. These angle coordinates give a representation on $\mathcal{M}_c$, but not on the whole of $\mathcal{M}$. We want to extend these angle coordinates to coordinates on $\mathcal{M}$ in a natural manner.

Theorem 5.33. *Let $(\mathcal{M}, \omega, f)$ be an integrable system and suppose that $\mathcal{M}_c$ is compact and connected. Then there exist Darboux coordinates $(\phi_1, \ldots, \phi_n, \psi_1, \ldots, \psi_n)$, called the **action-angle coordinates**, such that (ϕ_i) are the angle coordinates described in Corollary 5.32 and (ψ_i) are some first integrals.*

We will now go over the proof of this theorem here, but it can be found in [1, 7, 10]. The most important consequence of the proof given in [1] is the fact that we can solve an integrable system using quadratures. This goes to show that any integrable system hence has a somewhat 'nice' solution. Let us consider a relevant physical problem: two bodies attracted by gravity, also called the Kepler problem. We will show that this system is actually an integrable system.

Example 5.34. Let us consider the Kepler problem which consists of two bodies with masses m_1 and m_2 with an attractive gravitational interaction and a stationary centre of mass. We have seen in Example 4.8 that the potential for such a gravitational interaction is given by

$$U(r_1, r_2) = \frac{G m_1 m_2}{\|r_2 - r_1\|}.$$

Here, we will assume that $G = 1$ for ease of notation. The Lagrangian in Cartesian coordinates is then simply the kinetic energy minus this potential energy

$$\mathscr{L}(r_1, r_2, \dot{r}_1, \dot{r}_2) = \frac{1}{2}\left(m_1\|\dot{r}_1\|^2 + m_2\|\dot{r}_2\|^2\right) - \frac{m_1 m_2}{\|r_2 - r_1\|}.$$

Let us consider the coordinate transformation to the centre of mass system, defined by

$$R = \frac{m_1 r_1 + m_2 r_2}{m_1 + m_2}, \quad r = r_2 - r_1.$$

Inverting this transformation leads to the following expressions for r_1 and r_2

$$\begin{aligned} r_1 &= R - \frac{m_2}{m_1 + m_2} r, \quad \dot{r}_1 = \dot{R} - \frac{m_2}{m_1 + m_2}\dot{r} \\ r_2 &= R + \frac{m_1}{m_1 + m_2} r, \quad \dot{r}_2 = \dot{R} + \frac{m_1}{m_1 + m_2}\dot{r}. \end{aligned}$$

Using this coordinate transformation we find that the Lagrangian in these coordinates can be expressed as

$$\begin{aligned} \mathscr{L}\left(R, r, \dot{R}, \dot{r}\right) &= \frac{1}{2}\left(m_1\left\|\dot{R} - \frac{m_2}{m_1 + m_2}\dot{r}\right\|^2 + m_2\left\|\dot{R} + \frac{m_1}{m_1 + m_2}\dot{r}\right\|^2\right) - \frac{m_1 m_2}{\|r\|} \\ &= \frac{1}{2}\left(m_1\left\|\dot{R}\right\|^2 + m_1\left\|\frac{m_2}{m_1 + m_2}\dot{r}\right\|^2 - \frac{2 m_1 m_2}{m_1 + m_2}\dot{R} \cdot \dot{r}\right. \\ &\qquad \left. + m_2\left\|\dot{R}\right\|^2 + m_2\left\|\frac{m_1}{m_1 + m_2}\dot{r}\right\|^2 + \frac{2 m_1 m_2}{m_1 + m_2}\dot{R} \cdot \dot{r}\right) - \frac{m_1 m_2}{\|r\|} \\ &= \frac{1}{2}\left((m_1 + m_2)\left\|\dot{R}\right\|^2 + \frac{m_1 m_2}{m_1 + m_2}\|\dot{r}\|^2\right) - \frac{m_1 m_2}{\|r\|}. \end{aligned}$$

Now define $M = m_1 + m_2$ and $\mu = m_1 m_2/(m_1 + m_2)$ and remark that $\dot{R} = 0$ as the centre of mass is stationary. It then follows that the Lagrangian reduces to

$$\mathscr{L}\left(R, r, \dot{R}, \dot{r}\right) = \mathscr{L}(r, \dot{r}) = \frac{1}{2}\mu\|\dot{r}\|^2 - \frac{\mu M}{\|r\|}.$$

We can see that we can model our configuration space with $\mathbb{R}^3 \setminus \{0\}$, as we divide by the norm of r, and hence the phase space as $\left(\mathrm{T}^*\left(\mathbb{R}^3 \setminus \{0\}\right), \omega_{\mathrm{can}}\right)$. The Hamiltonian can be obtained by calculating the generalised momentum.

$$p = \frac{\partial \mathscr{L}}{\partial \dot{r}} = \mu \dot{r}.$$

Hence, the Hamiltonian is given by

$$\mathscr{H}(r, p) = p \cdot \dot{r} - \mathscr{L} = \frac{\|p\|^2}{\mu} - \frac{1}{2}\frac{\|p\|^2}{\mu} + \frac{\mu M}{\|r\|} = \frac{\|p\|^2}{2\mu} + \frac{\mu M}{\|r\|}.$$

Now let us define some function L on $\mathrm{T}^*\mathbb{R}^n$ which represents the angular momentum

$$L = r \times p = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \times \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} y p_z - z p_y \\ z p_x - x p_z \\ x p_y - y p_x \end{pmatrix} = \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix}.$$

We will now show that $f = \left(\mathscr{H}, \|L\|^2, L_z\right)$ gives an integrable system $(\mathcal{M}, \omega, f)$. Let us first check the commutation relations between the functions, which also shows that the L_z and $\|L\|^2$ are first integrals. We will start by reducing $\left\{\mathscr{H}, \|L\|^2\right\}$ to simpler cases,

$$\left\{\mathscr{H}, \|L\|^2\right\} = \left\{\mathscr{H}, L_x^2 + L_y^2 + L_z^2\right\} = \left\{\mathscr{H}, L_x^2\right\} + \left\{\mathscr{H}, L_y^2\right\} + \left\{\mathscr{H}, L_z^2\right\}.$$

Remark that $\left\{\mathscr{H}, L_i^2\right\} = 2\{\mathscr{H}, L_i\}L_i$ for $i \in \{x, y, z\}$, hence, it is enough to determine $\{\mathscr{H}, L_i\}$. Let us do this for $i = z$ as we have to determine this commutation relation directly as well,

$$\begin{aligned}\{\mathscr{H}, L_z\} &= \frac{\partial \mathscr{H}}{\partial x}\frac{\partial L_z}{\partial p_x} - \frac{\partial \mathscr{H}}{\partial p_x}\frac{\partial L_z}{\partial x} + \frac{\partial \mathscr{H}}{\partial y}\frac{\partial L_z}{\partial p_y} - \frac{\partial \mathscr{H}}{\partial p_y}\frac{\partial L_z}{\partial y} \\ &= -\frac{\mu M x y}{\|r\|^3} - \frac{p_x p_y}{\mu} + \frac{\mu M x y}{\|r\|^3} + \frac{p_x p_y}{\mu} = 0.\end{aligned}$$

It follows that $\{\mathscr{H}, L_z\} = 0$, and thus $\left\{\mathscr{H}, L_z^2\right\} = 0$ as well. This follows similarly for $i \in \{x, y\}$. As the above calculation proves that $\left\{\mathscr{H}, \|L\|^2\right\} = 0$ and $\{\mathscr{H}, L_z\} = 0$, we only have to check $\left\{\|L\|^2, L_z\right\}$.

$$\left\{\|L\|^2, L_z\right\} = \left\{L_x^2 + L_y^2 + L_z^2, L_z\right\} = 2\left[\{L_x, L_z\}L_x + \{L_y, L_z\}L_y + \{L_z, L_z\}L_z\right]. \tag{5.9}$$

Remark that $\{L_z, L_z\} = 0$ by the skew-symmetry of the Poisson bracket. Hence, we only need to calculate $\{L_z, L_x\}$ and $\{L_z, L_x\}$

$$\begin{aligned}\{L_x, L_z\} &= \frac{\partial L_x}{\partial y}\frac{\partial L_z}{\partial p_y} - \frac{\partial L_x}{\partial p_y}\frac{\partial L_z}{\partial y} = x p_z - p_x z = -L_y \\ \{L_y, L_z\} &= \frac{\partial L_y}{\partial x}\frac{\partial L_z}{\partial p_x} - \frac{\partial L_y}{\partial p_x}\frac{\partial L_z}{\partial x} = y p_z - p_y z = L_x.\end{aligned}$$

Combining these with Equation 5.9 gives us the result

$$\left\{\|L\|^2, L_z\right\} = 2\left[-L_y L_x + L_x L_y\right] = 0.$$

So indeed $\left\{\mathscr{H}, \|L\|^2, L_z\right\}$ is a set of commuting function. Let us consider the Jacobian of $f = \left(\mathscr{H}, \|L\|^2, L_z\right)$, which at a point $(r, p) = (x, y, z, p_x, p_y, p_z)$ is given by

$$dF_{(r,p)} = \begin{pmatrix} -\dfrac{\mu M x}{\|r\|^3} & -\dfrac{\mu M y}{\|r\|^3} & -\dfrac{\mu M z}{\|r\|^3} & \dfrac{p_x}{\mu} & \dfrac{p_y}{\mu} & \dfrac{p_z}{\mu} \\ 2\left[p \times L\right]_1 & 2\left[p \times L\right]_2 & 2\left[p \times L\right]_3 & 2\left[L \times r\right]_1 & 2\left[L \times r\right]_2 & 2\left[L \times r\right]_3 \\ 0 & p_z & -p_y & 0 & -z & y \end{pmatrix},$$

where the subscript of $[p \times L]$ and $[L \times r]$ denote the i-th component. We can recognise that this Jacobian has full rank on at least a dense subset of $\mathrm{T}^*\left(\mathbb{R}^3 \backslash \{0\}\right)$, as we mentioned one can show that this is enough for the system to be integrable by quadratures, see [4, Section 18.4]. Please refer to [10, Section 4.4] for more details on this problem. //

Appendix A

Vector Fields

Throughout this thesis, we assume the reader is familiar with vector fields on a manifold, and even with time-dependent vector fields. In this appendix, we will discuss some of the main results and showcase the notation we use. Furthermore, we will also introduce some theory regarding time-dependent vector fields. We follow Chapters 8 and 9 in [13], but for flows, we follow the approach of [18]. As a natural complement to these texts on manifolds, we also use some basic existence and uniqueness theorems of solutions to ordinary differential equations, see for example [11, 17] or Appendix D of [13].

In this appendix, we will first discuss vector fields and their algebraic properties. Then expand this to integral curves and flows and use these to take a special kind of derivative: the Lie derivative. After which, we go into time-dependent vector fields and again discuss their flows and derivatives.

A.1 Vector fields

Let us start by defining time-independent vector fields, or simply vector fields. Naturally, we identify these as tangent vectors appended to each point of a manifold. This is formally defined using sections of the projectile map $\pi : \mathrm{T}\mathcal{M} \to \mathcal{M}$.

Definition A.1. A **rough vector field** on a manifold is a section of the projection map $\pi : \mathrm{T}\mathcal{M} \to \mathcal{M}$. More concretely, $X : \mathcal{M} \to \mathrm{T}\mathcal{M}$ is a rough vector field if it is a continuous map such that $\pi \circ X = \mathrm{Id}_{\mathcal{M}}$, in other words, $X(p) \in \mathrm{T}_p\mathcal{M}$ for each $p \in \mathcal{M}$. We often denote $X(p)$ as X_p and identify this as the tangent vector in $\mathrm{T}_p\mathcal{M}$.

If a rough vector field is a smooth function with respect to the natural smooth structure on $\mathrm{T}\mathcal{M}$ imposed by $\mathcal{M}$, it is called a **smooth vector field** or **simply vector field**. The set of all vector fields on a manifold $\mathcal{M}$ is denoted by $\mathfrak{X}(\mathcal{M})$. //

Remark that we will use the name vector field solely for smooth vector fields and mention roughness only for not necessarily smooth vector fields. If $\left(U, \left(x^i\right)\right)$ is some chart around the point $p \in \mathcal{M}$, we can write

$$X_p = X^i(p) \left.\frac{\partial}{\partial x^i}\right|_p .$$

The functions $X^i : U \to \mathbb{R}$ are called the **component functions** of a vector field. These give us a more simple condition for the smoothness of vector fields.

Proposition A.2. *Let X be a rough vector field on $\mathcal{M}$ and $\left(U, \left(x^i\right)\right)$ be a coordinate chart. The restriction X to U is smooth if and only if its component functions in this chart are smooth.*

Proof. Let X and $\left(U, \phi = \left(x^i\right)\right)$ be as in the proposition. The coordinate representation of X is then given by

$$\hat{X}(x) = \left(x^1, \ldots, x^n, X^1\left(\phi^{-1}(x)\right), \ldots, X^n\left(\phi^{-1}(x)\right)\right).$$

Here X^i are the component functions of X in the given chart. As ϕ^{-1} and all x^i are smooth functions, the fact that $\hat{X}$ is smooth is equivalent to each X^i being smooth. □

Before we move on to more of the geometric structure of vector fields and their interactions with the manifold, we will discuss some of the algebraic structures.

A.1.1 Algebraic Structures

We can induce many different algebraic structures onto $\mathfrak{X}(\mathcal{M})$. Here, we will discuss its module structure over $C^\infty(\mathcal{M})$ and its Lie algebra structure. In this process, we also discuss some intermediate steps, namely, its vector space structure over $\mathbb{R}$ and the identification with derivations.

Vector Spaces

At each point, a vector field is a vector from a real vector space, namely the tangent space. This lets us define the addition and scalar multiplication of vector field pointwise in each of these vector spaces. For any $X, Y \in \mathfrak{X}(\mathcal{M})$, $a, b \in \mathbb{R}$ and $p \in \mathcal{M}$ define $aX + bY$ as

$$(aX + bY)_p = aX_p + bY_p.$$

Remark that this is well-defined as $X_p, Y_p \in \mathrm{T}_p\mathcal{M}$. With these operations, we notice that $\mathfrak{X}(\mathcal{M})$ inherits the same structure as the tangent spaces.

Proposition A.3. *The set of vector fields with pointwise addition and scalar multiplication form a real vector space.*

Proof. It is clear that for any $X, Y \in \mathfrak{X}(\mathcal{M})$ and $a, b \in \mathbb{R}$ the function $aX + bY$ is a rough vector field. We can check the smoothness using Proposition A.6. The coordinate functions of $aX + bY$ are given by

$$(aX + bY)^i = aX^i + bY^i.$$

As these are the addition of smooth functions, these are also smooth. □

Modules

However, what happens if we vary the scalar over the manifold using some smooth function? We can then again define the multiplication pointwise, such that for any $X, Y \in \mathfrak{X}(\mathcal{M})$, $f, g \in C^\infty(\mathcal{M})$ and $p \in \mathcal{M}$ we define

$$(fX + gX)_p = f(p) X_p + g(p) Y_p.$$

Again we might ask ourselves what kind of structure this has.

Proposition A.4. *The set of vector fields forms a module over $C^\infty(\mathcal{M})$ with pointwise multiplication.*

Proof. It is again evident that addition and multiplication give back a rough vector field. The smoothness is a result of Proposition A.6. Given some $X, Y \in \mathfrak{X}(\mathcal{M})$ and $f, g \in C^\infty(\mathcal{M})$, the coordinate functions of $fX + gY$ are given by

$$(fX + gY)^i = fX^i + gY^i.$$

Which is smooth as $C^\infty(\mathcal{M})$ is a ring. Hence, it follows that $\mathfrak{X}(\mathcal{M})$ is a module. □

Derivation

As a vector field is an element of the tangent space at each point, it inherits the derivation property at each point. To extend this globally, we define the action of vector fields on smooth functions.

Definition A.5. Let X be a vector field on $\mathcal{M}$ and $f \in C^\infty(\mathcal{M})$. Then define the function Xf pointwise such that $(Xf)(p) = X_p f$. //

Remark the difference between fX and Xf, where the first one is a vector field, and the second is a smooth function. Using the action of vector fields on smooth functions we get the following result.

Proposition A.6. *Let X be a rough vector field. If the function Xf is smooth for each $f \in C^\infty(\mathcal{M})$, then X is smooth as well.*

Proof. Suppose that X is a rough vector field. Firstly, we will prove that the assumption in the proposition implies we only need to look at open subsets of $\mathcal{M}$. We will then combine this with Proposition A.2.

Suppose that the conditions in the proposition hold, i.e. if $f \in C^\infty(\mathcal{M})$, we can assume that $Xf \in C^\infty(\mathcal{M})$. Now suppose that U is an open subset of $\mathcal{M}$ and that $g \in C^\infty(U)$. For any $p \in U$, we can find a smooth bump function $\psi \in C^\infty(\mathcal{M})$ and neighbourhood V of p such that $V \subset \operatorname{supp}\psi \subset U$. We can then extend g to the whole manifold in a smooth manner.

$$\tilde{g}(p) = \begin{cases} \psi g(p) & \text{, if } p \in U. \\ 0 & \text{, else.} \end{cases}$$

We can then conclude that $X\tilde{g}$ is smooth by our assumption. It should be clear that $X\tilde{g}|_V = Xg|_V$. Hence, Xg is smooth in a neighbourhood of every $p \in U$ and thus on the whole of U. As U was arbitrary, we conclude that for any arbitrary open subset $U \subset \mathcal{M}$ and $f \in C^\infty(U)$ the function $Xf \in C^\infty(U)$.

Take an arbitrary $p \in \mathcal{M}$ and some coordinate chart $\left(U, \left(x^i\right)\right)$ around it. Notice that x^i is smooth on U and thus is $X|_U x^i$ smooth as well. However, in these same coordinates, we can write

$$X|_U x^i = X^j \frac{\partial x^i}{\partial x^j} = X^i.$$

Hence, every coordinate function of X is smooth in U. Therefore X is smooth on U by Proposition A.2. As U was arbitrary, we can conclude that X is smooth on the whole of $\mathcal{M}$. □

We can also recognise that the vector fields inherit the local structure of the tangent space in terms of derivations as well.

Proposition A.7. *Any $X \in \mathfrak{X}(\mathcal{M})$ induces a map $X : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ using Definition A.5 which satisfies the following product rule*

$$X(fg) = fXg + gXf, \tag{A.1}$$

where $f, g \in C^\infty(\mathcal{M})$.

Proof. First, we should check whether Xf is indeed a smooth function. This can easily be done in some chart $\left(U, \left(x^i\right)\right)$.

$$Xf(p) = \left(X^i(p) \left.\frac{\partial}{\partial x^i}\right|_p \right) f = X^i(p) \frac{\partial f}{\partial x^i}(p).$$

As the coordinate functions are smooth and $f \in C^\infty(\mathcal{M})$, we conclude that Xf is smooth. The fact that it is linear and satisfies Equation A.1, is a direct consequence of the definition and the fact that X_p is a linear derivation for each p. □

We can even go further to say that any linear map $D : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ that satisfies the product rule, we call these **derivations**, can be identified with a vector field.

Proposition A.8. *A map $D : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ is a derivation if and only if it is of the form $Df = Xf$ for some $X \in \mathfrak{X}(\mathcal{M})$.*

Proof. Proposition A.7 tells us that any vector field induces a linear map that satisfies the product rule. Hence, if $D : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ is a map for which there exists an $X \in \mathfrak{X}(\mathcal{M})$ such that $Df = Xf$, it is clear that it is a derivation.

On the other hand, suppose that $D : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ is a derivation. Then define a rough vector field X at each $p \in \mathcal{M}$ by its action of functions,

$$X_p f = (Df)(p).$$

As D is a derivation, we can deduce that $X_p \in \mathrm{T}_p\mathcal{M}$. It then follows from Proposition A.6 that X is smooth. □

Lie algebra

Now, we will combine the vector space structure and the derivation property of vector fields to create an algebraic structure. Let us define a multiplication using the composition of vector fields as functions on $C^\infty(\mathcal{M})$, such that for some $X, Y \in \mathfrak{X}(\mathcal{M})$ and $f \in C^\infty(\mathcal{M})$

$$XYf = X(Yf). \tag{A.2}$$

We can verify that XY is not necessarily a vector field again, see Example A.9. Here, we make use of the derivation property of vector fields and show that the composition of two is not necessarily a derivation any longer.

Example A.9. Take $\mathcal{M} = \mathbb{R}^2$ with the global coordinates $\left(\mathbb{R}^2, (x, y)\right)$. Define the vector fields $X = \partial/\partial x$ and $Y = x\,\partial/\partial y$ and the functions $f(x, y) = x$ and $g(x, y) = y$. We then compute

$$XY(fg) = X(Y(xy)) = \frac{\partial}{\partial x}\left(x\frac{\partial}{\partial y}(xy)\right) = \frac{\partial}{\partial x}\left(x^2\right) = 2x.$$

Meanwhile, if we write out the product rule, we get the following computation.

$$fXYg + gXYf = x\frac{\partial}{\partial x}\left(x\frac{\partial y}{\partial y}\right) + y\frac{\partial}{\partial x}\left(x\frac{\partial x}{\partial y}\right) = x\frac{\partial x}{\partial x} + y\frac{\partial 0}{\partial x} = x.$$

We can then remark that $XY(fg) \neq fXYg + gXYf$, and therefore XY is not a derivation. //

Surprisingly, even though this multiplication does not work, it is still salvageable.

Definition A.10. The **Lie bracket** of two vector fields $X, Y \in \mathfrak{X}(\mathcal{M})$ is defined by its action on a function $f \in C^\infty(\mathcal{M})$

$$[X, Y]f = XYf - YXf.$$

The multiplication on the right-hand side is as in Equation A.2. //

Corollary A.11. *The Lie bracket of two vector fields is again a vector field.*

Proof. By Proposition A.8, it is enough to check whether $[X, Y]$ is a derivation. Take two vector fields $X, Y \in \mathfrak{X}(\mathcal{M})$ and smooth functions $f, g \in C^\infty(\mathcal{M})$.

$$\begin{aligned}
[X, Y](fg) &= XY(fg) - YX(fg) = X(fYg + gYf) - Y(fXg + gXf) \\
&= XfYg + fXYg + XgYf + gXYf - YfXg - fYXg - YgXf - gYXf. \\
&= fXYg + gXYf - fYXg - gYXf = f[X, Y]g + g[X, Y]f.
\end{aligned}$$

This proves that $[X, Y]$ is a vector field. □

Corollary A.12. *For $X, Y \in \mathfrak{X}(\mathcal{M})$ we can write $[X,Y]$ in a coordinate chart $\left(U, \left(x^i\right)\right)$ at a point $p \in U$ as*

$$[X,Y]\Big|_U = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i}\right) \frac{\partial}{\partial x^j},$$

where $X\big|_U = X^i \, \partial/\partial x^i$ and $Y\big|_U = Y^i \, \partial/\partial x^i$.

Proof. Let X and Y be vector fields and $f \in C^\infty(\mathcal{M})$. Then take an arbitrary coordinate chart $\left(U, \left(x^i\right)\right)$ of $\mathcal{M}$. We then express $X|_U = X^i \, \partial/\partial x^i$ and $Y|_U = Y^i \frac{\partial}{\partial x^i}$ and consider the action of $[X,Y]$ on f.

$$\begin{aligned}
[X,Y]|_U f &= (XY - YX)\,|_U f = X|_U \left(Y^j \frac{\partial f}{\partial x^j}\right) - Y|_U \left(X^j \frac{\partial f}{\partial x^j}\right) \\
&= X|_U \left(Y^j\right) \frac{\partial f}{\partial x^j} + Y^j X|_U \left(\frac{\partial f}{\partial x^j}\right) - Y|_U \left(X^j\right) \frac{\partial f}{\partial x^j} - X^j Y|_U \left(\frac{\partial f}{\partial x^j}\right) \\
&= X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial f}{\partial x^j} + Y^j X^i \frac{\partial^2 f}{\partial x^j \partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \frac{\partial f}{\partial x^j} - X^j Y^i \frac{\partial^2 f}{\partial x^j \partial x^i} \\
&= X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial f}{\partial x^j} - Y^i \frac{\partial X^j}{\partial x^i} \frac{\partial f}{\partial x^j} \\
&= \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i}\right) \frac{\partial f}{\partial x^j}.
\end{aligned}$$

This proves our result. □

We now use $[\cdot,\cdot] : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$ as the multiplication on $\mathfrak{X}(\mathcal{M})$. The pair $(\mathfrak{X}(\mathcal{M}), [\cdot,\cdot])$ is called the Lie algebra of vector fields.

Corollary A.13. *For any $X, Y, Z \in \mathfrak{X}(\mathcal{M})$ the Lie bracket is:*

(1) Bilinear, i.e. or any $a, b \in \mathbb{R}$

$$\begin{aligned}
[aX + bY, Z] &= a[X,Z] + b[Y,Z] \\
[Z, aX + bY] &= a[Z,X] + b[Z,Y].
\end{aligned}$$

(2) Antisymmetric

$$[X,Y] = -[Y,X].$$

(3) A derivation on itself

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

Proof. The first two properties follow from the basic definitions of the Lie bracket and the multiplication and addition of vector fields. The last property can then be proven by proving the equivalent **Jacobi identity**:

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

Which follows by writing it out and seeing that every term cancels out. □

A.1.2 Integral Curves and Flow

Let us turn back to the geometric picture of vector fields in the sense that they give a direction on a manifold. We make use of this sense of direction by defining integral curves, which are paths following the vector field.

Definition A.14. An **integral curve of X**, where $X \in \mathfrak{X}(\mathcal{M})$, is a function $\gamma : I \to \mathcal{M}$, with I being an interval in $\mathbb{R}$, such that for each $t \in I$

$$\dot{\gamma}(t) = X_{\gamma(t)}.$$

Furthermore, if 0 is contained in I, we call $\gamma(0)$ the starting point of γ. //

As our vector fields are independent of time, it follows that the integral curves are invariant under translations in time.

Corollary A.15. *Let $\gamma : I \to \mathcal{M}$ be an integral curve of $X \in \mathfrak{X}(\mathcal{M})$. Then for any $b \in \mathbb{R}$, the curve $\tilde{\gamma} : \tilde{J} \to \mathcal{M} : t \mapsto \gamma(t+b)$, where $\tilde{J} = \{t \in \mathbb{R} : t + b \in J\}$, is also an integral curve of X.*

Proof. Let $\gamma : J \to \mathcal{M}$ be an integral curve of an arbitrary vector field $X \in \mathfrak{X}(\mathcal{M})$ and suppose that $b \in \mathbb{R}$. By defining $\tilde{\gamma}$ and $\tilde{J}$ as in the lemma we can check that it is still an integral curve by letting it act on $f \in C^\infty(\mathcal{M})$.

$$\dot{\tilde{\gamma}}(s) f = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=s} (f \circ \tilde{\gamma})(t) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=s} (f \circ \gamma)(t+b) = \dot{\gamma}(s+b) f = X_{\tilde{\gamma}(s)} f.$$

This shows that, $\tilde{\gamma}$ is an integral curve of X. □

The existence of these integral curves is a result of the existence of solutions for ordinary differential equations. To showcase this connection, we will calculate the integral curves of a vector field in an example.

Example A.16. Let $\mathcal{M} = \mathbb{R}^n$ and let $\left(x^i\right)$ denote the global coordinates on $\mathbb{R}^n$. Take the point $x \in \mathbb{R}^n$ with $x^i(x) = 1$ for each $1 \leq i \leq n$ and define the vector field $X = x^i \, \partial/\partial x^i$.

We then look for the integral curve of X that has x as its starting point. Such an integral curve $\gamma : J \to \mathbb{R}^n$ should then satisfy, with $\hat{\gamma}$ as the coordinate representation of γ and $\hat{\gamma}^i$ its i-th component function

$$\frac{\mathrm{d}\hat{\gamma}^i}{\mathrm{d}t} \frac{\partial}{\partial x^i}\bigg|_{\gamma(t)} = \dot{\gamma}(t) = X_{\gamma(t)} = \hat{\gamma}^i(t) \frac{\partial}{\partial x^i}\bigg|_{\gamma(t)}.$$

and $\hat{\gamma}^i(0) = 1$ for each $1 \leq i \leq n$.

This results in a simple autonomous system of first-order linear differential equations with some initial values.

$$\dot{u}(t) = \left(\dot{u}^1(t), \cdots, \dot{u}^n(t)\right)^T = \left(u^1(t), \cdots, u^n(t)\right)^T \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = \mathrm{Id}\, u(t).$$

With the initial condition $u^i(0) = 1$ for all $1 \leq i \leq n$. The solution curves to these initial value problems are given by

$$u(t) = e^{t\,\mathrm{Id}} u(0) = e^t \,\mathrm{Id}\, u(0) = \left(e^t, \cdots, e^t\right)^T.$$

Thus the integral curve starting at x are given by $\hat{\gamma}^i(t) = e^t$. //

Example A.16 shows that finding the integral comes down to solving the differential equation generated by the vector field in coordinates. With this, we can ensure the local existence of integral curves as the differential equation theory only ensures local solutions and this is done in the local coordinates.

Proposition A.17. *For a vector field $X \in \mathfrak{X}(\mathcal{M})$ and a point $p \in \mathcal{M}$ there exists an $\epsilon > 0$ and curve $\gamma : (-\epsilon, \epsilon) \to \mathcal{M}$ that is an integral curve of X with starting point p. Furthermore, it is uniquely defined on this interval.*

Proof. Take an arbitrary vector field $X \in \mathfrak{X}(\mathcal{M})$, point $p \in \mathcal{M}$ and chart $\left(U, \phi = \left(x^i\right)\right)$ centred around p. The vector field can then be expressed in these coordinates as $X = X^i \, \partial/\partial x^i \,\big|_p$. Using this coordinate expression, we can translate our problem to an initial value problem on $\mathbb{R}^n$.

$$\dot{u}^i = X^i\left(\phi^{-1}(u)\right), \quad u(0) = \phi(p) = 0.$$

We can ensure the local existence of a unique solution to this problem, i.e. there exists an $\epsilon > 0$ and a function $u : (-\epsilon, \epsilon) \to \mathbb{R}^n$ that satisfies the initial value problem. We then define $\gamma = \phi^{-1} \circ u : (-\epsilon, \epsilon) \to \mathcal{M}$ such that it satisfies $\dot{\gamma}(t) = X_{\gamma(t)}$ and is therefore an integral curve of X. □

We would like to show that the solution for an integral curve starting at a point is unique, however, we can only show that the integral curves are unique on their domain. Therefore, we will focus on a special type of domain and integral curve, whose existence is unique.

Definition A.18. An integral curve $\gamma : I \to \mathcal{M}$ of $X \in \mathfrak{X}(\mathcal{M})$ starting at $p \in \mathcal{M}$ is called a **maximal integral curve** if its domain cannot be extended to a larger open interval while remaining an integral curve of X starting at p. //

Corollary A.19. *Let X be a vector field on a manifold $\mathcal{M}$ and suppose that p is an arbitrary point on $\mathcal{M}$. Then there exists a unique **maximal integral curve** $\gamma_p : I_p \to \mathcal{M}$ of X starting at p.*

Proof. Suppose that $X \in \mathfrak{X}(\mathcal{M})$ and $p \in \mathcal{M}$. By Theorem A.17 we know that there exists an integral curve of X starting at p.

Suppose that $\gamma_1 : I_1 \to \mathcal{M}$ and $\gamma_2 : I_2 \to \mathcal{M}$ are both integral curves of X starting at p. Then define the set of points where the integral curves coincide,

$$J = \{t \in I_1 \cap I_2 : \ \gamma_1(t) = \gamma_2(t)\}.$$

Remark that this is a non-empty set as $0 \in I_1 \cap I_2$ and $\gamma_1(0) = p = \gamma_2(0)$. We can easily see that it is closed by the continuity of the integral curves. Furthermore, we can prove that it is open using Corollary A.15. Translate the integral curves by some $t \in J$, for which we define $\sigma^i(s) = \gamma^1(s+t)$ for any $s \in \{t \in \mathbb{R} : t + s \in J\}$. Remark that $\sigma^1(0) = \gamma^1(t) = \gamma^2(t) = \sigma^2(0) := q$. Hence, by the uniqueness of Proposition A.17 it follows that $\sigma^1 = \sigma^2$ in a neighbourhood of 0. It follows that J contains an neighbourhood of t in $I_1 \cap I_2$, hence J is open in $I_1 \cap I_2$. As J is open, closed and non-empty in $I_1 \cap I_2$, it follows that J is the whole of $I_1 \cap I_2$. This implies that any two integral curves starting at p agree on the intersection of their domains.

Now let I_p be the union of all the domains of integral curves starting at p. Then define $\gamma_p(t)$ for some $t \in I_p$ be the common value of all integral curves starting at p whose domain contain t. Then $\gamma_p : I_p \to \mathcal{M}$ is maximal. The uniqueness of this integral curve follows from the fact that any two integral curves must agree on their common domain and we can not extend I_p any further. □

Next up, we would like to not only consider the integral curve of a vector field at a point but on the whole manifold simultaneously. This is motivated by the fact that we can generate a vector field simply from a so-called global flow.

Proposition A.20. *Given a smooth function $\phi : \mathbb{R} \times \mathcal{M} \to \mathcal{M}$, from which we also define $\phi^t : \mathcal{M} \to \mathcal{M} : p \mapsto \phi(t,p)$ and $\phi^{(p)} : \mathbb{R} \to \mathcal{M} : t \mapsto \phi(t,p)$, which is a global flow. This means that it satisfies $\phi^t \circ \phi^s = \phi^{t+s}$ and $\phi^0 = \mathrm{Id}_{\mathcal{M}}$. There exists a vector field X such that $X_p = \dot{\phi}^{(p)}(0)$ and $\phi^{(p)}$ is the integral curve of this vector field with starting point p.*

Proof. Let ϕ be as in the proposition and define X pointwise as $X_p = \dot{\phi}^{(p)}(0)$. We will prove that X is smooth using Proposition A.6, as it is clear that X is a section of $\pi : \mathrm{T}\mathcal{M} \to \mathcal{M}$ by definition. Hence, let f be an arbitrary smooth function on some $U \subset \mathcal{M}$ and suppose that $p \in U$. We then calculate $Xf(p)$,

$$Xf(p) = X_p(f) = \dot{\phi}^{(p)}(0) f = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f\left(\phi^{(p)}(t)\right) = \left.\frac{\partial}{\partial t}\right|_{t=0} f(\phi(t,p)).$$

As both f and ϕ are smooth this shows that Xf is smooth as well. As f and U are arbitrary this shows that X is smooth.

We should still check that $\phi^{(p)}$ is an integral curve of X. This comes down to an easy calculation on an arbitrary smooth function f.

$$\begin{aligned} X_{\phi^{(p)}(s)} f = \dot{\phi}^{(\phi(s,p))} f &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(\phi(t, \phi(s,p))) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(\phi(t+s,p)) = \left.\frac{\mathrm{d}}{\mathrm{d}u}\right|_{u=s} f(\phi(u,p)) = \dot{\phi}^{(p)}(s) f. \end{aligned}$$

This proves that $X_{\phi^{(p)}(s)} = \dot{\phi}^{(p)}(s)$ implying that $\phi^{(p)}$ are the integral curves of X. Furthermore, the assumption that $\phi(0,p) = p$ implies that $\phi^{(p)}$ is the integral curve starting at p. □

We would now like to inverse these operations. Unfortunately, this is not always possible as the flow of a vector field may not be global, i.e. not defined on the whole of $\mathbb{R} \times \mathcal{M}$. Hence, we weaken our condition and look for some local flow or flow as a function $\phi_X : \mathcal{D}(X) \to \mathcal{M}$ where $\mathcal{D}(X) \subset \mathbb{R} \times \mathcal{M}$.

Definition A.21. Suppose that $X \in \mathfrak{X}(\mathcal{M})$. Define $\mathcal{D}(X)$, which we call the **flow domain of** X, as

$$\mathcal{D}(X) = \{(t,p) \in \mathbb{R} \times \mathcal{M} : \ t \in I_p\},$$

where I_p is the domain of the maximal integral curve starting at p. The associated map is called the **flow of** X, $\phi_X : \mathcal{D}(X) \to \mathcal{M}$, and is defined as

$$\phi_X(t,p) = \gamma_p(t),$$

where γ_p is the maximal integral curve starting at p. Naturally, we can define two maps $\phi_X^{(p)} : I_p \to \mathcal{M} : t \mapsto \phi(t,p)$ and $\phi_X^t : \mathscr{D}(X) \to \mathcal{M} : p \mapsto \phi(t,p)$, where $\mathscr{D}(X)$ is a subset of $\mathcal{M}$ given by $\mathscr{D}(X) = \{p \in \mathcal{M} : \ (t,p) \in \mathcal{D}(X)\}$ //

Even though our definition is not algebraic, the flow does have some group structure and it is therefore sometimes called the local one-parameter group action.

Proposition A.22. *For an $X \in \mathfrak{X}(\mathcal{M})$ and arbitrary $p \in \mathcal{M}$. Then for any $s \in I_p$, we have that $t \in I_{\phi(s,p)}$ if and only if $t + s \in I_p$. Furthermore, the flow of X obeys the following*

$$\phi_X(t, \phi_X(s,p)) = \phi_X(t+s, p),$$

where $s \in I_p$ and $t \in I_{\phi_X(s,p)}$.

Proof. Let X, p, s and t be as in the lemma. Then call $q = \phi_X(s,p)$ and $\tau_s(t) = t + s$. Then it follows that $\gamma_p \circ \tau_s : \tau_{-s}(I_p) \to \mathcal{M}$ is an integral curve of X starting at q and therefore $\tau_{-s}(I_p) \subset I_q$ and $\gamma_p \circ \tau_s = \gamma_q\big|_{\tau_{-s}(I_p)}$. This implies that $\phi_X^{t+s} = \phi_X^t \circ \phi_X^s$.

In a similar manner, we can conclude that $\gamma_q \circ \tau_{-s} : \tau_s(I_q) \to \mathcal{M}$ is an integral curve starting at p, and hence $\tau_s(I_q) \subset I_p$. But we had already seen that $\tau_{-s}(I_p) \subset I_q$, hence $\tau_s(I_q) = I_p$. □

Lastly, we will go into some of the topological properties of the flow and flow domain.

Theorem A.23. *For a vector field X on $\mathcal{M}$, the flow domain $\mathcal{D}(X)$ is open and ϕ_X is smooth.*

Proof. Suppose that X is a vector field on $\mathcal{M}$. For an arbitrary $p \in \mathcal{M}$ with a surrounding chart $\left(U, \phi = \left(x^i\right)\right)$, we can solve the flow in the coordinate representation

$$\dot{\phi}_X^i(t,p) = X^i(\phi_X(t,p)).$$

We know that there is some smooth solution to this differential equation exists, see [5, Appendix C]. Hence, there exists a neighbourhood U of p such that ϕ_X is smooth on $(-\epsilon, \epsilon) \times U$.

Suppose that we define the $W \subset \mathcal{D}(X)$ as the set of all point (t,p) such that ϕ_X is defined and smooth on some neighbourhood of (t,p) of the form $J \times U$. Then this is an open subset of $\mathbb{R} \times \mathcal{M}$ and ϕ_X restricted to W is smooth as well. We will show by contradiction that $W = \mathcal{D}(X)$. The idea is to find a point up until which the flow is smooth and then remark that we can extend the flow smoothly using Proposition A.22.

Suppose that $\overline{W} = \mathcal{D}(X) - W$ is non-empty, and assume that there is some $(t, p_0) \in \overline{W}$ with $t > 0$. Define some τ as follows

$$t_0 = \inf\left(t \in \mathbb{R}_{\geq 0} : \ (t, p_0) \in \overline{W}\right).$$

Clearly, the flow is smooth in some neighbourhood $(-\epsilon_0, \epsilon_0) \times W_0$ of p_0, hence $t_0 > 0$. Now define the point $q = \phi_X(t_0, p_0)$. We can then assure the smoothness of the flow in some neighbourhood

$(-\epsilon_q, \epsilon_q) \times W_q$ of q. Now take some $t_1 < t_0$ such that $t_1 + \epsilon_q > t_0$. Then $(t_1, p_0) \in W$ such that there is some neighbourhood $(t_1 - \epsilon_1, t_1 + \epsilon_1) \times W_1 \subset W$. Thus the flow is smooth in this neighbourhood, which lets us define the following mapping

$$\tilde{\phi} : [0, t_1 + \epsilon_q) \times W_1 \to \mathcal{M} : (t, p) \mapsto \begin{cases} \phi_X(t, p) & \text{if } 0 \leq t < t_1 \\ \phi_X(t - t_1, \phi_X(t_1, p)) & \text{if } t_1 \leq t < t_1 + \epsilon_q. \end{cases}$$

This then forms a natural smooth extension of the flow in a neighbourhood of (t_0, p_0), which was a contradiction with the definition of t_0. Hence, it follows that $\overline{W}$ is empty, implying that $\mathcal{D}(X) = W$ and that it is therefore open. □

Derivations along vector fields

With the interpretation of the vector fields as a sense of direction and the flow as the paths we can walk along, we can see how different functions change in the direction of the vector field. We call such derivatives the Lie derivative along a vector field. We can define these for both functions and vector fields, and we will see that they coincide with some simpler expressions. In Section 3.3.1, we go deeper into the action of the Lie derivative on tensor fields.

Definition A.24. For an $X \in \mathfrak{X}(\mathcal{M})$ and $f \in C^\infty(\mathcal{M})$ we define the **Lie derivative of f along X** as

$$\mathcal{L}_X f = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\phi_X^t\right)^* f.$$

The existence of this operator is ensured by the fact that the flow locally exists //

Corollary A.25. *For a vector field X and smooth function f we have $\mathcal{L}_X f = Xf$.*

Proof. Take some arbitrary $X \in \mathfrak{X}(\mathcal{M})$ and $f \in C^\infty(\mathcal{M})$. Remark that the pullback of a function is defined such that

$$\left(\phi_X^t\right)^* f(p) = f \circ \phi_X^{(p)}(t).$$

We can calculate the Lie derivative at an arbitrary point p. This reduces to the derivative of the composition $f \circ \phi_X^{(p)}$,

$$(\mathcal{L}_X f)(p) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f \circ \phi_X^{(p)} = df_p\left(\dot{\phi}_X^{(p)}(0)\right) = df_p(X_p) = Xf(p).$$

This proves that $\mathcal{L}_X f = Xf$ as p is arbitrary. □

We can extend this definition to vector fields.

Definition A.26. Lie derivative of vector field Y along X is defined as

$$\mathcal{L}_X Y = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\phi_X^t\right)^* Y.$$

By the local existence of the flow, this is well-defined. //

We can also express this operator in simpler terms, namely in terms of the Lie bracket defined in Definition A.10. This proof is based on Theorem 20.4 in [22].

Corollary A.27. *For arbitrary vector fields X and Y, the Lie derivative of Y along X is given by $\mathcal{L}_X Y = [X, Y]$.*

Proof. Let $\mathcal{M}$ be an arbitrary manifold and suppose that $X, Y \in \mathfrak{X}(\mathcal{M})$ and $f \in C^\infty(\mathcal{M})$. We know we can write the Lie bracket of X and Y in some coordinate chart $\left(U, \left(x^i\right)\right)$ using Corollary A.12.

$$[X, Y]_p = \left(X^i \left.\frac{\partial Y^j}{\partial x^i}\right|_p - Y^i \left.\frac{\partial X^j}{\partial x^i}\right|_p\right) \left.\frac{\partial}{\partial x^j}\right|_p.$$

So we just need to check whether the Lie derivative also satisfies this equation.

$$\begin{aligned}(\mathcal{L}_X Y)_p &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(\left(\phi_X^t\right)^* Y\right)_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(d\phi_X^{-t}\right)_{\phi_X(t,p)} \left(Y_{\phi_X(t,p)}\right) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} Y^i\left(\phi_X(t,p)\right) \left(d\phi_X^{-t}\right)_{\phi_X(t,p)} \left(\left.\frac{\partial}{\partial x^i}\right|_{\phi_X(t,p)}\right) \\ &= \left[\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} Y^i\left(\phi_X(t,p)\right) \left.\frac{\partial x^j \circ \phi_X^{-t}}{\partial x^i}\right|_{\phi_X(t,p)}\right] \left.\frac{\partial}{\partial x^j}\right|_p\end{aligned}$$

Using the product and chain rule, we can simplify this equation.

$$\begin{aligned}&= \left[\left.\frac{\partial Y^i}{\partial x^k}\right|_p \left.\frac{\partial x^k \circ \phi_X^{(p)}}{\partial t}\right|_{t=0} \left.\frac{\partial x^j \circ \phi_X^0}{\partial x^i}\right|_p - Y^i(p) \left.\frac{\partial^2 x^j \circ \phi_X^{-t}}{\partial t \partial x^i}\right|_{t=0,\phi_X(t,p)}\right] \left.\frac{\partial}{\partial x^j}\right|_p \\ &= \left[\left.\frac{\partial Y^j}{\partial x^k}\right|_p X^k(p) - Y^i(p) \left.\frac{\partial X^j}{\partial x^i}\right|_p\right] \left.\frac{\partial}{\partial x^j}\right|_p = [X,Y]_p\end{aligned}$$

□

Thus, we see that the Lie derivative has not yet given us any new operations in the sense that $\mathcal{L}_X f = Xf$ and $\mathcal{L}_X Y = [X,Y]$ were already defined. It does give us more intuition behind the definition of these operators.

A.2 Time-dependent Vector Fields

Now we will generalise our theory of vector fields, to ones that change over time. These turn up naturally in the proof of Darboux's theorem, see Theorem 3.10, but their theory is slightly more involved than time-independent vector fields. However, we can come to a nearly equivalent result.

Definition A.28. A **time-dependent vector field** on $\mathcal{M}$ is a smooth function $X : J \times \mathcal{M} \to \mathrm{T}\mathcal{M}$, where J is an interval in $\mathbb{R}$, such that for each $(t,p) \in J \times \mathcal{M}$ we have $X(t,p) \in \mathrm{T}_p\mathcal{M}$. In other words, the map $X_t : \mathcal{M} \to \mathrm{T}\mathcal{M}$ defined by $X_t(p) = X(t,p)$ is a vector field of $\mathcal{M}$. //

We will not go into the algebraic structure of the vector fields here, but we solely focus on the geometric interpretation.

A.2.1 Integral Curves and Flow

Much like time-independent vector fields, we can use a time-dependent vector field to generate motion over a manifold in the form of an integral curve.

Definition A.29. Let X be a time-dependent vector field on $\mathcal{M}$ defined on the time interval J. An **integral curve** of X is a curve $\gamma : I \to \mathcal{M}$, with $I \subset J$, such that

$$\dot{\gamma}(t) = X(t, \gamma(t)),$$

with $t \in I$. We define maximality similarly to Definition A.18. //

We should note that it is much harder to define a flow in this case, as different integral curves may go over the same point without being equal to each other in a neighbourhood of this point, see Example A.30.

Example A.30. Let $\mathcal{M} = \mathbb{R}^2$ and define $X(t,p) = -\sin(t)\,\partial/\partial x + \cos(t)\,\partial/\partial y$.

Let $\gamma_1 : [0,\infty] \to \mathcal{M}$ be the integral curve such that $\gamma_1(0) = (1,0)^T$. It then needs to be a solution of the following differential equation

$$\dot{\gamma}(t) = X(t, \gamma(t)) = \begin{pmatrix} -\sin(t) \\ \cot(t) \end{pmatrix}.$$

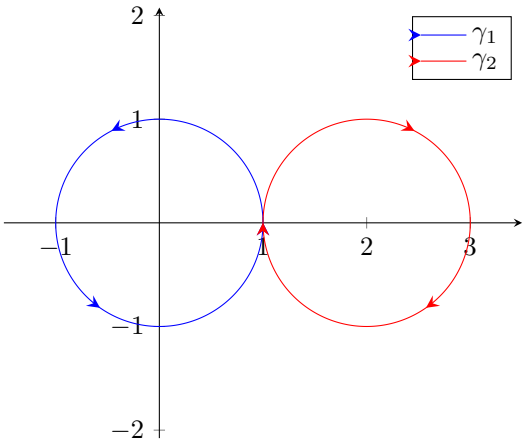


Figure A.1

The obvious solution then is the following

$$\gamma_1(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

Now suppose that $\gamma_2 : [\pi, \infty] \to \mathcal{M}$ is the integral curve such that $\gamma(\pi) = (1,0)^T$. For time-independent vector fields, this would lead to the same integral curve up to some time translation. However, given our time-dependent vector field, we find a different solution to the differential equation, namely

$$\gamma_2(t) = \begin{pmatrix} 2 + \cos(t) \\ -\sin(t) \end{pmatrix}.$$

We can see in Figure A.1, that these are very different, even though they pass through the same point, $(1,0)^T$. This is very different from time-independent vector fields. //

The previous example clearly shows that it is not very informative to look at a single integral curve passing through a point. We will instead focus on the more global behaviour over time in terms of a time-dependent flow.

Theorem A.31. *Let X be a time-dependent vector field on $\mathcal{M}$ defined on some open interval J. There exists an open subset $\mathscr{E}(X) \subset J \times J \times \mathcal{M}$ and a smooth function $\phi_X : \mathscr{E}(X) \to \mathcal{M}$ called the **time-dependent flow** with the following properties:*

(1) For each $t_0 \in J$ and $p \in \mathcal{M}$, the set $\mathscr{E}^{(t_0,p)}(X) = \{t \in J : (t, t_0, p) \in \mathscr{E}(X)\}$ is an open interval containing t_0. Furthermore, $\phi_X^{(t_0,p)}(t) = \phi_X(t, t_0, p)$ is the unique maximal integral curve of V, with the condition $\phi_X^{(t_0,p)}(t_0) = p$.

(2) For any $t_1 \in \mathscr{E}^{(t_0,p)}(X)$ and $q = \phi_X^{(t_0,p)}(t_1)$, we have $\mathscr{E}^{(t_1,q)}(X) = \mathscr{E}^{(t_0,p)}(X)$ and $\phi_X^{(t_1,q)} = \phi_X^{(t_0,p)}$.

(3) For each $(t_1, t_0) \in J \times J$, the set $\mathcal{M}_{t_1,t_0} = \{p \in \mathcal{M} : (t_1, t_0, p) \in \mathscr{E}(X)\}$ is open in $\mathcal{M}$, and the map $\phi_X^{t_1,t_0} : \mathcal{M}_{t_1,t_0} \to \mathcal{M}$ defined by $\phi_X^{t_1,t_0}(p) = \phi_X(t_1, t_0, p)$ is a diffeomorphism from $\mathcal{M}_{t_1,t_0}$ onto $\mathcal{M}_{t_0,t_1}$.

(4) If $p \in \mathcal{M}_{t_1,t_0}$ and $\phi_X^{t_1,t_0}(p) \in \mathcal{M}_{t_2,t_1}$, then $p \in \mathcal{M}_{t_2,t_1}$ and

$$\phi_X^{t_2,t_1} \circ \phi_X^{t_1,t_0}(p) = \phi_X^{t_2,t_0}(p).$$

Proof. Let X be a time-dependent vector field on $\mathcal{M}$ defined on a time interval J. We can translate this to a vector field $\widetilde{X}$ on $J \times \mathcal{M}$ defined as follows

$$\widetilde{X}_{(s,p)} = \left(\left.\frac{\partial}{\partial s}\right|_s, X(s,p) \right).$$

We use s as the standard standard coordinate of J and remark that an element of $\mathrm{T}_{(s,p)}(J \times \mathcal{M})$ can be identified in $\mathrm{T}_s J \oplus \mathrm{T}_p \mathcal{M}$. This vector field has a flow $\phi_{\widetilde{X}}$ that is of the following form

$$\phi_{\widetilde{X}}(t,(s,p)) = (\alpha(t,(s,p)), \beta(t,(s,p))).$$

As it is the flow of $\widetilde{X}$ we can see that α satisfies the following initial value problem

$$\frac{\partial \alpha}{\partial t}(t,(s,p)) = 1, \quad \alpha(0,(s,p)) = s.$$

Hence, we get $\alpha(t,(s,p)) = s + t$. For β we then see it should satisfy the following

$$\frac{\partial \beta}{\partial t}(t,(s,p)) = X(t+s, \beta(t,(s,p))). \tag{A.3}$$

Hence, we see that this can function as an integral curve of the vector field, enticing us to define the flow of X, ϕ_X, as

$$\phi_X(t,t_0,p) = \beta(t-t_0,(t_0,p)). \tag{A.4}$$

The smoothness of ϕ_X is implied by the smoothness of $\phi_{\widetilde{X}}$ and β. The domain of this function, $\mathscr{E}(X)$, can be made as large as possible by defining it as

$$\mathscr{E}(X) = \left\{(t,t_0,p) \in \mathbb{R} \times J \times \mathcal{M} : \ (t-t_0,(t_0,p)) \in \mathcal{D}(\widetilde{X})\right\}.$$

Remark that the function α maps $\mathcal{D}(\widetilde{X})$ to J, such that for any point $(t,t_0,p) \in \mathscr{E}(X)$ we have $t \in J$, implying that $\mathscr{E}(X) \subset J \times J \times \mathcal{M}$. Similarly, we can deduce that $\mathscr{E}(X)$ is open from the fact that $\mathcal{D}(\widetilde{X})$ is open.

Now take some arbitrary $t_0 \in J$ and $p \in \mathcal{M}$ and define $\mathscr{E}^{(t_0,p)}(X)$ as in the theorem, which is open by definition. If we define $\phi_{\widetilde{X}}^{((t_0,p))} = \phi_{\widetilde{X}}(t,t_0,p)$ we can see that it is an integral curve by combining Equation A.3 and A.4. The uniqueness and maximality are direct consequences of the definition of the flow of a vector field. This proves 1.

If we have some $t_0 \in J$ and $p \in \mathcal{M}$, we can take some arbitrary $t_1 \in \mathscr{E}^{(t_0,p)}$ and define $q = \phi_X(t_1,t_0,p)$. Then $\phi_X^{((t_0,p))}$ and $\phi_X^{((t_1,q))}$ are both integral curves that go through q at $t = t_1$. Hence, by the uniqueness of the flow of $\widetilde{X}$ these are the same curve and hence $\mathscr{E}^{(t_0,p)} = \mathscr{E}^{(t_1,q)}$. This proves 2.

We will now skip to proving 4. Assuming $p \in \mathcal{M}_{t_1,t_0}$ and $\phi_X^{t_1,t_0}(p) \in \mathcal{M}_{t_2,t_1}$, then 2 implies that

$$\phi_X(t_2,t_1,\phi_X(t_1,t_0,p)) = \phi_X(t_2,t_0,p) \implies \phi_X^{t_2,t_1} \circ \phi_X^{t_1,t_0}(p) = \phi_X^{t_2,t_0}(p).$$

This proves 4.

Lastly, take a look at 3. Suppose that $(t_1,t_0) \in J$, the openness of $\mathcal{M}_{t_1,t_0}$ is implied by the openness of $\mathscr{E}(X)$. We can easily define the inverse function of $\phi_X^{t_1,t_0}$ as this is simply $\phi_X^{t_0,t_1}$. We should remark that for an arbitrary $p \in M_{t_1,t_0}$ we know that $\mathscr{E}^{(t_0,p)} = \mathscr{E}^{(t_1,q)}$, with $q = \phi_X^{t_1,t_0}(p)$. Implying that $q \in M_{t_0,t_1}$, or in other words, $\phi_X^{t_1,t_0}(M_{t_1,t_2}) = M_{t_0,t_1}$. This concludes the proof. □

Derivations along time-dependent vector fields

With the existence of the flow, we can once again see how functions and vector fields change along the vector field.

Definition A.32. The Lie derivative of $f \in C^\infty(\mathcal{M})$ along a time-dependent vector field X on $\mathcal{M}$ at some $t \in J$, where J is an open interval in $\mathbb{R}$ on which X is defined, is given by

$$\mathcal{L}_{X_t} f = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} \left(\phi_X^{s,t}\right)^* f.$$

The existence is ensured by Theorem A.31. //

Proposition A.33. *If X is a time-dependent vector field on $\mathcal{M}$ and $s \in J$, then*

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} \left(\phi_X^{s,0}\right)^* f = \left(\phi_X^{t,0}\right)^* \mathcal{L}_{X_t} f$$

for all $f \in C^\infty(\mathcal{M})$.

Proof. Suppose that X is a time-dependent vector field on $\mathcal{M}$ and $t \in J$, where J is the time-interval on which X is defined. It follows for an arbitrary $f \in C^\infty(\mathcal{M})$ from the definition that

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} \left(\phi_X^{s,0}\right)^* f &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} \left(\phi_X^{s,t} \circ \phi_X^{t,0}\right)^* f = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} \left(\phi_X^{t,0}\right)^* \circ \left(\phi_X^{s,t}\right)^* f \\ &= \left(\phi_X^{t,0}\right)^* \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} \left(\phi_X^{s,t}\right)^* f = \left(\phi_X^{t,0}\right)^* \mathcal{L}_{X_t} f\end{aligned}$$

This proves our statement. □

Furthermore, remark that the action of the Lie derivative of a time-dependent vector field can be calculated rather simply.

Proposition A.34. *Let X be a time-dependent vector field on $\mathcal{M}$ defined on a time interval J. The Lie derivative of an $f \in C^\infty(\mathcal{M})$ at $t \in J$ is given by*

$$\mathcal{L}_{X_t} f = X_t f.$$

Proof. Let X be a time-dependent vector field on $\mathcal{M}$, defined on the time interval J. Take an arbitrary $f \in C^\infty(\mathcal{M})$, $p \in \mathcal{M}$ and $t \in J$, we can then conclude that

$$\begin{aligned}\left(\mathcal{L}_{X_t} f\right)(p) &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} \left(\phi_X^{s,t}\right)^* f(p) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} \left(f \circ \phi_X^{(t,p)}\right)(s) \\ &= df_p\left(\dot{\phi}_X^{(t,p)}(t)\right) = df_p\left(X(t,p)\right) = \left(X_t f\right)(p).\end{aligned}$$

As p is arbitrary it follows that $\mathcal{L}_{X_t} f = X_t f$ on the whole of $\mathcal{M}$. □

We can see that the Lie derivative along a time-dependent vector field acts in the same manner as the Lie derivative of a time-independent vector field does.

Bibliography

[1] V. I. Arnol'd, *Mathematical method of classical mechanics*, Second, Graduate texts in mathematics, Springer-Verlag, New York, 1989.

[2] V. I. Arnol'd, V. V. Kozlov, and A. I. Neĭshtadt, *Mathematical Aspects of Classical Mechanics*, Third, Encyclopeedia of Mathematical Sciences, Springer, Berlin, 2006.

[3] O. Bolza, *Vorlesungen über Variationsrechnung*, BG Teubner Verlag, Leibniz, 1909.

[4] A. Cannas da Silva, *Lectures on Symplectic Geometry*, Lecture Notes in Mathematics, Springer, 2008.

[5] L. Conlon, *Differentiable Manifolds A First Course*, Advanced Texts, Birkhäuser, Boston, 1993.

[6] Gaston Darboux, *Sur le probleme de Pfaff*, Bulletin des sciences mathématiques et astronomiques **6** (1882), no. 1, 14–36.

[7] J. J. Duistermaat, *On global action-angle coordinates*, Communications on Pure and Applied Mathematics **33** (1980nov), no. 6, 687–706.

[8] S. H. Friedberg, A. J. Insel, and L. E. Spence, *Linear algebra*, Fourth, Pearson Education, Upper Saddle River, N.J., 2003.

[9] A. Gray, *Tubes*, Second, Progress in mathematics, Springer, Basel, 2004.

[10] G. Heckman, *Symplectic Geometry. Lecture Notes.*, 2014.

[11] E. L. Incle, *Ordinary differential equations*, Dover books in advanced mathematics, Dover publications, New York, 1956.

[12] Hansjörg Kielhöfer, *Calculus of Variations*, Springer, Cham, 2018.

[13] J. M. Lee, *Introduction to Smooth Manifolds*, Second, Springer, New York, 2013.

[14] ———, *Introduction to Riemannian Manifolds*, Second, Springer International Publishing, 2019.

[15] J. Moser, *On the volume elements on a manifold*, Transactions of the American Mathematical Society **120** (February 1965), no. 2, 286–286.

[16] J. Munkres, *Topology*, Second, Pearson Education Limited, 2013.

[17] T. Myint-U, *Ordinary Differential Equations*, Elsevier North-Holland, New York, 1978.

[18] I. T. Mărcut, *Manifolds. Lecture Notes - Fall 2017*, 2017.

[19] B. Riemann, *Über die hypothesen, welche der geometrie zu grunde liegen.*, Second (H. Weyl, ed.), Berlin, 1921.

[20] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Second, Vol. 2, Publish or Perish, Berkeley, 1979.

[21] J. R. Taylor, *Classical Mechanics*, University Science Books, Mill Valley, Califonia, 2005.

[22] L. W. Tu, *An Introduction to Manifolds*, Second, Universitytext, Springer, New York, 2011.

[23] A. Weinstein, *Symplectic manifolds and their lagrangian submanifolds*, Advances in Mathematics **6** (1971), no. 3, 329 –346.