BACHELOR'S THESIS

Degree in Mathematics

Title: Geometric Mechanics and Hamilton-Jacobi Theory

Author: Iñigo Urtiaga Erneta

Advisor: Xavier Gràcia Sabaté

Department: Mathematics

Academic year: 2016-2017

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya
Facultat de Matemàtiques i Estadística

Degree in Mathematics

Bachelor's Degree Thesis

Geometric Mechanics and Hamilton–Jacobi Theory

Iñigo Urtiaga Erneta

Supervised by Xavier Gràcia Sabaté

Department of Mathematics

January 2017

To my parents, for their trust and support.

Abstract

A review of analytical mechanics in the language of differential geometry is given. The classical formulations of Newton, Lagrange and Hamilton are discussed in detail, with special interest in the Hamilton-Jacobi equation. The latter is studied in a new framework, which allows to identify the interesting geometric structures underlying the classical Hamilton-Jacobi theory. Some relevant examples are analysed in this context.

Keywords

Lagrangian mechanics, Hamiltonian mechanics, Hamilton-Jacobi equation, slicing equation.

MSC2010

70H20, 70H06, 70G45, 70H03, 70H05.

Contents

Chapter 1

Introduction

Analytical Mechanics

Mathematics is essential for the foundation and progress of the natural sciences. Perhaps the best example is found in Physics, where the interplay between both disciplines is so profound that it has led to the development of many active research topics in Mathematics. These conform the broad subject of Mathematical Physics, which provides methods that can be applied to real problems.

The physicist starts from a set of assumptions about reality that do not contradict experience to some extent. These principles are taken as axioms that can be included in some existing mathematical theory, and constitute a theoretical model for the physical situation under study. When the model is sufficiently exact, the problem of Physics is reduced to translating the features of the original system into the objects and relations of the corresponding theory and applying its characteristic techniques to extract the desired information.

Historically, one of the first and most successful theoretical model was achieved in the context of mechanics. Mechanics is the physical science concerned with the motion of bodies under the action of forces. This was an old problem of which many qualitative investigations had been made, most notably by Copernicus, Galileo and Kepler. In the seventeenth century, Newton developed a quantitative mathematical theory that described the motion of a wide variety of systems and reproduced the previously known experimental facts. According to this formalism, the spatial coordinates of a mechanical system as a function of time satisfied a system of second order differential equations, *Newton's equations*.

The theory worked remarkably well and began the mathematical study of mechanics, which is nowadays known as *analytical mechanics*. The problem of mechanics was thus reduced to integrating Newton's equations, a problem which turned out to be quite involved as the equations often appeared in the form of complex systems of ODEs. This difficulty led other scientists to look for new formulations of mechanics that could be regarded as a generalization of Newton's work: producing equations that were easier to integrate while recovering the implications of Newton's theory. Among these, there are two prominent formulations in analytical mechanics which include *Newtonian mechanics* as a particular case.

The first is *Lagrangian mechanics*, which was elaborated by Lagrange in the eighteenth century. Lagrange succeeded in deriving Newton's equations from a more general type of equations, independent of the coordinates employed. This fact allowed the use of special coordinates that simplified several problems in mechanics, particularly those involving constraints in the motion of the system. In the modern Lagrangian formalism, the equations governing the motion are obtained from a scalar function, the *Lagrangian* of the system, by taking its partial derivatives.

These equations are a coupled system of second order ODEs in the spatial coordinates.

Half a century later, Hamilton was able to derive Lagrange's equations from a variational principle, *Hamilton's principle of least action*, and gave his own formulation of mechanics, *Hamiltonian mechanics.* Hamilton defined an additional set of variables from the Lagrangian function, the *generalized momenta*, and considered a new function containing all the dynamical information of the system, the *Hamiltonian.* The equations of motion are once again derived from the scalar function, in the form of first order ODEs in the positions and momenta.

Hamilton–Jacobi Theory

In this context, Hamilton recognized various features of the original theory which had been ignored until then. One of these dealt with the variational principle whence Lagrange's equations arose, which came from the problem of determining the extrema of a particular integral functional. Under certain conditions, the functional could be made into a function of the coordinates, the *action function.*

Hamilton studied the action function and noted that it determined the dynamical variables he had considered. Furthermore, he observed that the action was a solution to a certain partial differential equation of first order. It was not clear, however, how to recover the motion of a mechanical system from a solution of the PDE.

It was Jacobi who later found the relationship between the systems of ODEs appearing in mechanics and the PDE discovered by Hamilton, producing valid solutions of the former from a special class of solutions of the latter. He looked for solutions of the PDE depending on arbitrary parameters, obtaining a family of functions that accounted for different dynamical trajectories. From the family, he derived smaller systems of ODEs which gave the original motion and were easier to integrate. Jacobi applied his method to various problems in mechanics successfully, becoming the most powerful technique for integrating the equations of motion. The celebrated PDE is consequently known as the *Hamilton-Jacobi equation.*

Hamilton found the equation by relating his theory of mechanics with his previous work on optics. In the latter, he was able to establish an unprecedented relation between the particle properties and the wave nature of light, through a PDE accounting for both behaviours. This analogy, together with the work of Jacobi, unveiled a deep connection between the trajectories of a mechanical system and partial differential equations. These ideas would pave the way for a more profound generalization of the theory of mechanics in the next century, where Schrödinger was able to produce the fundamental PDE accounting for the evolution of a quantum-mechanical system from the Hamilton-Jacobi equation.

Although it would seem that due to its antiquity the Hamilton-Jacobi equation is already well understood, there are several aspects of the theory surrounding it that had not been considered until recently. From a geometrical point of view, the situation suggests that the solutions of the Hamilton-Jacobi equation give a way of describing the dynamics of a mechanical system by the dynamics of a smaller system. This way, it becomes interesting to investigate this phenomenon in general, identifying the notable geometrical objects that allow for the dynamics of a system to be simplified through the solutions of a PDE.

Goals of the thesis

- To formulate analytical mechanics in the language of differential geometry, which is the natural framework for the classical theory ([JS98]) .

- To study the solutions and properties of the Hamilton-Jacobi equation as a particular aspect of the general theory.
- To present a new general framework, as seen in the works of Cariñena et al. ([CGM+06], [CGM+16]), where the geometrical structures underlying the Hamilton-Jacobi theory are best understood.
- To describe examples in the new framework.

Contents

We give a brief description of each chapter separately:

- **Chapter 2. Complements on Differential Geometry.**

  Here we define some of the basic geometrical concepts that are needed in the development of the theory. The chapter is divided into two parts: one dedicated to symplectic geometry, and a second one dealing with structures arising in vector bundles. In symplectic geometry, we consider a manifold with a 2-form satisfying certain conditions and study the objects that can be defined with it. These play an important role in Hamiltonian mechanics, where manifolds are *symplectic*. In the second part, we study vector bundles in general, with special attention to the natural bundles of a manifold: the tangent and cotangent bundles. There are several canonical objects appearing in these bundles, which are required to define the interesting objects of mechanics. The content of this chapter is taken from parts 1 and 2 in [AM78].

- **Chapter 3. Analytical Mechanics**

  In this chapter, we develop the classical theory of analytical mechanics from the point of view of differential geometry. We start with Newtonian dynamical systems, stating the axioms of Newtonian mechanics and writing the equations of motion for manifolds. We then consider Lagrangian systems, showing their relation with Newton's equations in the context of constrained motion, and writing the dynamical equations in an intrinsic manner. Finally, we present Hamiltonian systems as a generalization of Lagrangian systems via the Legendre transform. Throughout the chapter, several classical problems from mechanics are solved in each formalism, to give a picture of each situation. A great part of the theory is taken from [Arn74], [AM78] and [Gra15], while the examples are taken from [LL77].

- **Chapter 4. Classical Hamilton-Jacobi Theory**

  Here the classical Hamilton-Jacobi theory is developed. The first two sections are devoted to the action function and its relation with geometrical optics, explaining the motivation that led Hamilton to his formulation of mechanics and the discovery of the Hamilton-Jacobi equation. The rest of the chapter deals with the relationship of the equations of motion of a mechanical system and the PDE, as Jacobi understood it. We study Jacobi's method to produce *complete solutions* of the Hamilton-Jacobi equation, with application to the integration of motion in completely integrable systems. Additionally, we solve some important problems in mechanics with Jacobi's method. The content of the chapter is a combination of ideas from classical books in Physics [LL77] and Mathematics [Arn74], [CH37] discussing the Hamilton-Jacobi PDE. The classical examples are taken from [CH37].

- **Chapter 5. A Generalized Geometric Setting for Hamilton–Jacobi Theory.**

  In this last chapter, the new framework for the Hamilton-Jacobi theory is presented. We show how the Hamilton-Jacobi problem is equivalent to the description of the dynamics in a manifold by means of a family of dynamics in a lower dimensional manifold (a *slicing* of the dynamics). We find the conditions under which the motion of a system is equivalent to a solution of a PDE for different situations, depending on additional structures on the carrier manifold. More precisely, we consider Hamiltonian systems and fibered manifolds separately, and construct their corresponding analogue to the Hamilton-Jacobi equation. Furthermore, we recover the classical results as a particular case and we tackle several problems in the new framework. The first part of the chapter is a summary of the article [CGM+16], relating its content with the previous development. The solved examples are new, and give a novel insight into the role of the Hamilton-Jacobi equation in practice, particularly when considering consecutive slicings of the dynamics.

Chapter 2

Complements on Differential Geometry

An intrinsically geometrical formulation of analytical mechanics requires several algebraic and geometric objects, and the development of some foundational theory. It is assumed that the reader knows the basics of differential geometry, which can be found in many books (e.g. [Lan99], [Lee03]). In order to facilitate the reading of the text, we present here some topics from geometry that are not so common but still necessary to formulate mechanics, namely: the subject of symplectic geometry and certain notable structures arising in vector bundles. We take most concepts from [AM78, chapters 1,2,3 and 5].

Throughout the work, the maps are always assumed to be $\mathcal{C}^\infty$ and the words "differentiable" and "smooth" are used indistinctly. It is clear how to translate the definitions and theorems for more general classes of differentiability. Given a smooth manifold M, we use $\mathcal{F}(M)$ to denote the functions on the manifold, $\mathfrak{X}(M)$ for the smooth vector fields and $\Omega^k(M)$ for the differentiable k-forms.

2.1 Symplectic Geometry

2.1.1 Symplectic Manifolds

Definition 2.1.1. Let E be a finite-dimensional real vector space and ω a bilinear form on E, $\omega : E \times E \to \mathbb{R}$. We say that the form is *nondegenerate* when $\omega(u, v) = 0$ for all $v \in E$ implies $u = 0$.

Every bilinear form ω defines a linear map from the space to its dual E^*

$$\omega^\flat : E \to E^*, \quad u \mapsto \omega(u, \cdot). \tag{2.1}$$

We cannot say anything *a priori* about the properties of the map, but when ω is nondegenerate, $\omega^\flat$ is injective by definition. Because the dimensions of E and E^* coincide in finite dimension, we conclude that $\omega^\flat$ defines a vector space isomorphism if and only if ω is nondegenerate. Under these conditions, we define the inverse of the map $\omega^\sharp = (\omega^\flat)^{-1}$.

Definition 2.1.2. A bilinear form $\omega : E \times E \to \mathbb{R}$ is *symplectic* when it is both nondegenerate and skew-symmetric.

For a skew-symmetric bilinear form to be nondegenerate, a necessary condition is that E be of even dimension. Indeed, let $n = \dim E$ and let $\langle \cdot, \cdot \rangle_E$ denote the canonical pairing between E and its dual (when the context allows it, we will omit the subindex E). For $u, v \in E$, we have

$$\langle {}^t\omega^\flat(v), u \rangle_E = \langle v, \omega^\flat(u) \rangle_{E^*} = \langle \omega^\flat(u), v \rangle_E$$

where ${}^t\omega^\flat : E^{**} \to E$ is the transpose of $\omega^\flat$ and we have identified $E^{**} \cong E$. By skew-symmetry, $\omega^\flat = -{}^t\omega^\flat$ whence

$$\det \omega^\flat = \det {}^t\omega^\flat = \det -\omega^\flat = (-1)^n \det \omega^\flat.$$

But $\omega^\flat$ is an isomorphism, so $\det \omega^\flat \neq 0$ and we get a contradiction for n odd. From now on, we let $n = 2m$.

It can be shown that we can always find a basis $\{e_1, \cdots, e_n\}$ of E such that ω takes the form

$$\omega = \sum_{i=1}^{m} e_*^i \wedge e_*^{i+m} \tag{2.2}$$

where $\{e_*^1, \cdots, e_*^n\}$ is the dual basis and $\wedge$ denotes the wedge product of forms. To construct the basis, we take two vectors $e_1, e_{m+1} \in E$ such that $\omega(e_1, e_{m+1}) \neq 0$ which is possible by nondegeneracy. Dividing by a constant, we may further assume that $\omega(e_1, e_{m+1}) = 1$. The proof proceeds by induction on the orthogonal complement of the span of e_1, e_{m+1} [AM78, proposition 3.1.2].

For any other basis $\{u_1, \cdots, u_n\}$ of E, ω is written as $\omega = \omega_{ij} u_*^i \otimes u_*^j$, where the coefficients ω_{ij} conform an antisymmetric matrix ($\omega_{ij} = -\omega_{ji}$). Note that we have used the index summation convention in the expression of ω, that is, a subindex equal to a superindex indicates a summation. We shall be using this convention throughout the report, unless explicitly stated.

It is clear that the nondegeneracy condition is equivalent to 0 not being an eigenvalue of $\omega^\flat$. Using coordinates, we can write an arbitrary vector $v \in E$ as $v = v^i u_i$ and consequently $\omega^\flat(v^i u_i) = v^i \omega_{ij} u_*^j$, whence we see that nondegeneracy is also equivalent to the condition $|\omega_{ij}| \neq 0$.

In order to extend the notion of symplectic algebra to manifolds, we present an important theorem of Darboux that relates the particularly simple expression of simplectic forms to closedness of a global 2-form.

Theorem 2.1.3 (Darboux)**.** *Let ω be a nondegenerate differential 2-form on a $2m$ dimensional smooth manifold M. Then ω is closed (i.e. $d\omega = 0$) if and only if there exists a chart (U, φ) at each $p \in M$ such that $\varphi(p) = 0$, with $\varphi(u) = (x^1(u), \cdots, x^m(u), y_1(u), \cdots, y_m(u))$ for $u \in U$ and*

$$\omega|_U = dx^i \wedge dy_i. \tag{2.3}$$

To prove the theorem, one must looks for a chart where ω is constant. The result follows from considering the coordinates mentioned earlier on the chart (see [AM78, theorem 3.2.2]). We refer to the charts in the theorem as *symplectic charts* and to their coordinates as *symplectic coordinates* or *Darboux coordinates.*

We can finally give the definition of a symplectic form:

Definition 2.1.4. Let M be a smooth manifold. A 2-form $\omega \in \Omega^2(M)$ is *symplectic* when it closed and nondegenerate at each point of M.

It is clear that a symplectic form ω defines a vector bundle isomorphism $\omega^\flat : \mathrm{T}M \to \mathrm{T}^*M$ by globalizing 2.1, and we write again $\omega^\sharp = (\omega^\flat)^{-1}$. If we let $\mathscr{F}(M)$ denote the functions in M, then we have a map $\mathfrak{X}(M) \to \Omega^1(M)$, $X \mapsto i_X\omega = \omega^\flat \circ X$ which gives a $\mathscr{F}(M)$-module isomorphism.

Definition 2.1.5. A *symplectic manifold* is a pair (M, ω) where M is a manifold and $\omega \in \Omega^2(M)$ is symplectic.

The symplectic structure imposes some restrictions on the topology of the manifold M.

The closedness of the 2-form guarantees the existence of Darboux coordinates, which shows that symplectic manifolds are locally the same. In particular, they are all locally isomorphic to $\mathbb{R}^{2m}$ with the canonical 2-form $dx^1 \wedge dx_{1+m} \wedge \cdots \wedge dx^m \wedge dx_{2m}$, and they do not posses any local invariant other than the dimension.

Also, symplectic manifolds must be orientable, since $\Omega = (\omega)^m = \omega \wedge \overset{m)}{\cdots} \wedge \omega$ is a volume form on M. In symplectic coordinates,

$$\Omega = (-1)^{m(m-1)/2} dx^1 \wedge \cdots \wedge dx^m \wedge dy_1 \wedge \cdots \wedge dy_m. \tag{2.4}$$

We state Poincaré's lemma, which will be used many times throughout the text. For a proof, see [Spi65, theorem. 4.11].

Theorem 2.1.6 (Poincaré Lemma)**.** *Let $\omega \in \Omega^k(M)$ be closed. Then for each $x \in M$ there exists a neighbourhood U of x on which $\omega|_U \in \Omega^k(U)$ is an exact form (i.e. $\omega|_U = d\theta$ for some $\theta \in \Omega^{k-1}(U)$).*

We now study a special class of fields arising in the context of symplectic manifolds. Let $H : M \to \mathbb{R}$ be a function defined on a symplectic manifold (M, ω).

Definition 2.1.7. The *symplectic gradient* of H is the vector field

$$Z_H = \omega^\sharp \circ dH. \tag{2.5}$$

Equivalently, it is the only vector field such that

$$i_{Z_H}\omega = dH. \tag{2.6}$$

H is the *Hamiltonian* of Z_H. A vector field X is said to be *Hamiltonian* when $X = Z_H$ for some function H. We call it *locally Hamiltonian* if every point in M has an open neighbourhood where the restriction of X is Hamiltonian. It is useful to keep in mind the following characterization of locally Hamiltonian fields:

Proposition 2.1.8. *The following are equivalent:*

(i). X is locally Hamiltonian.

(ii). $i_X\omega$ is closed.

(iii). $\mathscr{L}_X\omega = 0$

Proof. $d^2 = 0$ proves that (i) implies (ii). The converse is also true by the Poincaré Lemma. Using Cartan's formula and the closedeness of ω, we find $\mathscr{L}_X\omega = i_X d\omega + d(i_X\omega) = d(i_X\omega)$ which shows the equivalence between (ii) and (iii). □

We know that the integral curves of a vector field define a local flow, which is an automorphism of the manifold. By the definition of the Lie derivative, the third condition in 2.1.8 tells us that locally Hamiltonian fields are exactly the fields whose flow leaves ω invariant. In particular,

$$\mathscr{L}_X\Omega = \sum_{i=1}^{m} (-1)^{(i+1)} \omega \wedge \cdots \wedge \mathscr{L}_X \overset{(i)}{\omega} \wedge \cdots \wedge \omega = 0 \tag{2.7}$$

and we obtain a well-known theorem of Liouville as a corollary:

Theorem 2.1.9 (Liouville)**.** *The symplectic volume form Ω is invariant by locally Hamiltonian fields.*

It is useful to remember the expressions in Darboux coordinates for the objects we have considered so far. These take the form:

$$\omega^\flat\left(X^i\frac{\partial}{\partial x^i}+Y_i\frac{\partial}{\partial y_i}\right)=X^idy_i-Y_idx^i \tag{2.8}$$

$$\omega^\sharp\left(A_idx^i+B^idy_i\right)=-A^i\frac{\partial}{\partial y_i}+B_i\frac{\partial}{\partial x^i} \tag{2.9}$$

$$Z_H=\frac{\partial H}{\partial y_i}\frac{\partial}{\partial x^i}-\frac{\partial H}{\partial x^i}\frac{\partial}{\partial y_i} \tag{2.10}$$

The Lie derivative defined by a Hamiltonian field has some remarkable properties when acting on functions. Let $F,G,H\in\mathscr{F}(M)$ be functions on M. We consider their variation with respect to the flow of symplectic gradients.

Definition 2.1.10. The *Poisson bracket* between F and G is the function

$$\{F,G\}=\mathscr{L}_{Z_G}F \tag{2.11}$$

or equivalently, it is the variation of F along the flow of the Hamiltonian field defined by G.

From the definition, it is clear that

$$\{F,G\}=\mathscr{L}_{Z_G}F=i_{Z_G}dF=i_{Z_G}i_{Z_F}\omega=\omega(Z_F,Z_G) \tag{2.12}$$

and in particular, we see that the Poisson bracket is skew-symmetric:

$$\{F,G\}=-\{G,H\}. \tag{2.13}$$

We also have the identities

$$\{FG,H\}=F\{G,H\}+G\{F,H\} \tag{2.14}$$
$$\{F,GH\}=H\{F,G\}+G\{F,H\} \tag{2.15}$$

which follow from applying Leibniz's rule to the expression $\{FG,H\}=-\mathscr{L}_{Z_H}(FG)$.

Let us consider momentarily a nondegenerate 2-form $\omega\in\Omega^2(M)$ which is not necessarily closed. Applying the well-known formulas for the Lie derivative of forms we find that

$$\begin{aligned}i_{[Z_F,Z_G]}\omega&=\mathscr{L}_{Z_F}i_{Z_G}\omega-i_{Z_G}\mathscr{L}_{Z_F}\omega\\&=\mathscr{L}_{Z_F}dG-i_{Z_G}(i_{Z_F}d\omega+d(i_{Z_F}\omega))\\&=i_{Z_F}d^2G+d(i_{Z_F}dG)-i_{Z_G}i_{Z_F}d\omega-i_{Z_G}(d^2F)\\&=-d\{F,G\}-i_{Z_G}i_{Z_F}d\omega\end{aligned}$$

whence

$$i_{[Z_F,Z_G]+Z_{\{F,G\}}}\omega=-i_{Z_G}i_{Z_F}d\omega \tag{2.16}$$

and by nondegeneracy we conclude that ω is closed if and only if

$$[Z_F,Z_G]+Z_{\{F,G\}}=0. \tag{2.17}$$

We observe that 2.17 is also equivalent to *Jacobi's identity.* Indeed, by definition of the exterior derivative:

$$\begin{aligned}-d\omega(X_F, X_G, X_H) &= -\mathscr{L}_{Z_F}\omega(Z_G, Z_H) - \mathscr{L}_{Z_G}\omega(Z_H, Z_F) - \mathscr{L}_{Z_H}\omega(Z_F, Z_G)\\ &= \{F,\{G,H\}\} + \{G,\{H,F\}\} + \{H,\{F,G\}\}\end{aligned}$$

and ω is closed if and only if

$$\{F,\{G,H\}\} + \{G,\{H,F\}\} + \{H,\{F,G\}\} = 0 \tag{2.18}$$

which is Jacobi's identity.

It is worth translating the previous remarks to algebraic terms. Clearly, $\mathscr{F}(M)$ is a $\mathbb{R}$-algebra with the Poisson bracket, and ω is closed if and only if $(\mathscr{F}(M), \{\cdot,\cdot\})$ is a Lie-algebra. Moreover, we observe that the set of Hamiltonian fields in M, which we denote by $\mathfrak{X}_h(M)$, must be closed under the Lie bracket by 2.17. This also implies that the map $\mathscr{F}(M) \to \mathfrak{X}_h(M)$, $f \mapsto Z_f$ is a Lie-algebra antihomomorphism between $(\mathscr{F}(M), \{\cdot,\cdot\})$ and $(\mathfrak{X}_h(M), [\cdot,\cdot])$.

In symplectic coordinates, the bracket is written as

$$\{F, G\} = \frac{\partial F}{\partial x^i}\frac{\partial G}{\partial y_i} - \frac{\partial F}{\partial y_i}\frac{\partial G}{\partial x^i} \tag{2.19}$$

and in particular, for the coordinate functions we obtain the *canonical relations*

$$\{x^i, x^j\} = 0, \quad \{y_i, y_j\} = 0, \quad \{x^i, y_j\} = \delta^i_j. \tag{2.20}$$

In the next chapter, we will see that Hamilton's formulation of mechanics requires a symplectic structure to write the equations of motion. In this connection, one must often consider diffeomorphisms which preserve the symplectic structure, as they can be used to transform the dynamics of the mechanical system under study into a more tractable form. The following definitions and results are concerned with the properties of these mappings.

Definition 2.1.11. Let $\varphi : M \to N$ be a differentiable map between symplectic manifolds (M, ω), (N, ρ). φ is *symplectic* when $\varphi^*(\rho) = \omega$. If φ is a diffeomorphism, we call it a *symplectomorphism* or *symplectic isomorphism.*

We prove a couple of lemmas that will make our lives easier:

Lemma 2.1.12. *Let $\varphi : M \to N$ be a diffeomorphism, $X \in \mathfrak{X}(N)$ a field and $\omega \in \Omega^k(N)$ a differential k-form. Then,*

$$\varphi^*(i_X\omega) = i_{\varphi^*(X)}\varphi^*(\omega) \tag{2.21}$$

Proof. If $k = 1$, letting $x \in M$ we have

$$\begin{aligned}\varphi^*(i_X\omega)(x) &= i_X\omega(\varphi(x))\\ &= \langle\omega(\varphi(x)), X(\varphi(x))\rangle\\ &= \langle\omega(\varphi(x)), \mathrm{T}\varphi \circ \mathrm{T}(\varphi^{-1}) \circ X(\varphi(x))\rangle\\ &= \langle\varphi^*(\omega)(x), \varphi^*(X)(x)\rangle\\ &= i_{\varphi^*(X)}\varphi^*(\omega)(x)\end{aligned}$$

If $k > 1$, letting $v_1, \dots, v_{k-1}$ denote $k-1$ vectors in $\mathrm{T}_x M$, we have

$$\begin{aligned}
\varphi^*(i_X\omega)(v_1, \dots, v_{k-1}) &= (i_X\omega) \circ \varphi(x)(\mathrm{T}_x\varphi \cdot v_1, \dots, \mathrm{T}_x\varphi \cdot v_{k-1}) \\
&= \omega_{\varphi(x)}(X(\varphi(x)), \mathrm{T}_x\varphi \cdot v_1, \dots, \mathrm{T}_x\varphi \cdot v_{k-1}) \\
&= \omega_{\varphi(x)}(\mathrm{T}_x\varphi \circ \mathrm{T}_{\varphi(x)}\varphi^{-1} \cdot X(\varphi(x)), \mathrm{T}_x\varphi \cdot v_1, \dots, \mathrm{T}_x\varphi \cdot v_{k-1}) \\
&= \omega_{\varphi(x)}(\mathrm{T}_x\varphi \cdot \varphi^*(X)(x), \mathrm{T}_x\varphi \cdot v_1, \dots, \mathrm{T}_x\varphi \cdot v_{k-1}) \\
&= i_{\varphi^*(X)}\varphi^*(\omega)(v_1, \dots, v_{k-1})
\end{aligned}$$

□

Lemma 2.1.13. *Let $\varphi : M \to N$ be a diffeomorphism, $X \in \mathfrak{X}(N)$ a field and $\omega \in \Omega^k(N)$ a differential k-form. Then,*

$$\varphi^*(\mathscr{L}_X\omega) = \mathscr{L}_{\varphi^*(X)}\varphi^*(\omega). \tag{2.22}$$

Proof. Using Cartan's formula for forms and 2.1.12, we have

$$\begin{aligned}
\varphi^*(\mathscr{L}_X\omega) &= \varphi^*(i_X d\omega + d\, i_X\omega) \\
&= i_{\varphi^*(X)}\varphi^*(d\omega) + d\,\varphi^*(i_X\omega) \\
&= i_{\varphi^*(X)} d\,\varphi^*(\omega) + d\, i_{\varphi^*(X)}\varphi^*(\omega) \\
&= \mathscr{L}_{\varphi^*(X)}\varphi^*(\omega)
\end{aligned}$$

□

Proposition 2.1.14. *Let $\varphi : M \to N$ be a diffeomorphism, where (M, ω), (N, ρ) are symplectic manifolds. The following are equivalent:*

(i). φ is a symplectomorphism.

(ii). For every $K \in \mathscr{F}(N)$, $\varphi^*(Z_K) = Z_{\varphi^*(K)}$.

(iii). For every $F, G \in \mathscr{F}(N)$, $\varphi^*\{F, G\} = \{\varphi^*(F), \varphi^*(G)\}$.

Proof. We first prove the equivalence $(i) \iff (ii)$:

If (i) holds, we have $i_{\varphi^* Z_K}\omega = \varphi^*(i_{Z_K}\rho) = \varphi^*(dK) = d(\varphi^* K)$ and (ii) also holds.

Assume (ii) holds, then $\varphi^*(dK) = d(\varphi^*K) = i_{\varphi^*(Z_K)}\omega = \varphi^*(i_{Z_K}\varphi_*\omega)$ whence it follows that $i_{Z_K}\rho = dK = i_{Z_K}\varphi_*\omega$. It can be shown that given $v_x \in \mathrm{T}_x N$ there always exists a function $K \in \mathscr{F}(N)$ such that $Z_K(x) = v_x$. Thus, $\varphi_*\omega = \rho$ and we conclude $\omega = \varphi^*\rho$.

For $(ii) \iff (iii)$, we observe

$$\begin{aligned}
\varphi^*\{F, G\} &= \varphi^*\mathscr{L}_{Z_G}F = \mathscr{L}_{\varphi^*(Z_G)}\varphi^*(F) \\
\{\varphi^*(F), \varphi^*(G)\} &= \mathscr{L}_{Z_{\varphi^*(G)}}\varphi^*(F)
\end{aligned}$$

so $\varphi^*\{F, G\} = \{\varphi^*(F), \varphi^*(G)\}$ if and only if $\varphi^*(Z_G) = Z_{\varphi^*(G)}$, and we are done. □

In particular, we may consider symplectic automorphisms from a symplectic manifold onto itself. If (M, ω) is a symplectic manifold and $X \in \mathfrak{X}(M)$ a field on M, we say that X is an *infinitesimal symplectomorphism* if the local flow it defines is a symplectic automorphism. Equivalently, it is a field such that $\mathscr{L}_X\omega = 0$ and we see from 2.1.8 that the infinitesimal symplectomorphisms are actually the locally Hamiltonian fields.

Later on, we will see that the symplectic structure accounts for the geometry of the space where we describe the motion of a mechanical system, while the Hamiltonian fields contain all the dynamical information. In 2.1.14, we have seen that the maps preserving the symplectic structure also preserve the Hamiltonian character of a field. We investigate a possible converse relation.

Definition 2.1.15. Let $\varphi : M \to N$ be a diffeomorphism between symplectic manifolds (M, ω), (N, ρ). φ is a *canonical transformation* if it transforms locally Hamiltonian fields into locally Hamiltonian fields.

In this context, there is a strong theorem by Lee Hwa Chung concerning the shape of invariant forms. For a proof, see [Lee47].

Theorem 2.1.16 (Lee Hwa Chung). *Let $\rho \in \Omega^k(M)$ such that $\mathscr{L}_X \rho = 0$ for every locally Hamiltonian field $X \in \mathfrak{X}(M)$. Then:*

1. *$\rho = 0$ if k is odd.*
2. *$\rho = c\,\omega \wedge \overset{k}{\cdots} \wedge \omega$ for some $c \in \mathbb{R} \setminus \{0\}$, if k is even.*

We use 2.1.16 to prove a useful characterization of canonical transformations.

Proposition 2.1.17. *Let $\varphi : M \to N$ a diffeomorphism between connected symplectic manifolds (M, ω), (N, ρ). Then φ is a canonical transformation if and only if there exists $c \in \mathbb{R} \setminus \{0\}$ such that $\varphi^*(\rho) = c\omega$. Moreover, given a locally Hamiltonian field $X \in \mathfrak{X}(M)$ with Hamiltonian H, and considering $Y = \varphi_*(X) \in \mathfrak{X}(N)$ with Hamiltonian K, there exists $k \in \mathbb{R}$ such that $cH = \varphi^*(K) + k$.*

Proof. If φ is canonical, we have $\mathscr{L}_{\varphi_*(X)} \rho = 0$ for every locally Hamiltonian field $X \in \mathfrak{X}(M)$. Thus, $\mathscr{L}_X \varphi^*(\rho) = \varphi^*(\mathscr{L}_{\varphi_*(X)} \rho) = 0$ and we get the first implication by Lee Hwa Chung. Conversely, assume $\varphi^*(\rho) = c\omega$, then $\mathscr{L}_{\varphi_*(X)} \rho = \mathscr{L}_{\varphi_*(X)} \varphi_*(c\omega) = c\varphi_*(\mathscr{L}_X \omega)$ which is zero for locally Hamiltonian fields.

For the second part, observe that $d(cH - \varphi^*(K)) = c\,dH - \varphi^*(dK) = c\,i_{Z_H}\omega - \varphi^*(i_{Z_K}\rho) = c\,i_{Z_H}\omega - i_{\varphi^*(Z_K)}\varphi^*(\rho) = c\,i_{Z_H}\omega - i_{Z_H}c\omega = 0$ and the function $cH - \varphi^*(K)$ is locally equal to a certain constant k. By connectedness we get that it is globally equal to k. $\square$

Constant c in the above proposition is called the *valence* of the canonical transformation. Thus, symplectomorphisms are precisely the canonical transformations of valence $c = 1$.

From 2.1.17 we deduce a practical way of describing a canonical transformation, which will motivate the Hamilton-Jacobi theory in chapter 4.

Proposition 2.1.18. *Let $\varphi : M \to N$ be a canonical transformation of valence c between symplectic manifolds (M, ω), (N, ρ), with $\omega = -d\theta$, $\rho = -d\vartheta$. Then there exists locally a function $S \in \mathscr{F}(M)$ such that*

$$\varphi^*(\vartheta) - c\theta = dS.$$

S is the generating function *of the canonical transformation.*

Proof. We have $d(\varphi^*(\vartheta) - c\theta) = \varphi^* d\vartheta - c d\theta = -\varphi^*(\rho) + c\omega = 0$ and the existence of S follows from the Poincaré Lemma. $\square$

2.1.2 Lagrangian Submanifolds

We now investigate the notion of orthogonality in the context of symplectic manifolds. In analogy with Riemannian geometry, the orthogonality will be defined in terms of the symplectic form, which is the symplectic analogue of the metric. In this connection, we define the concepts of isotropic, coisotropic and Lagrangian submanifolds, with especial interest in the last, which will be needed to derive certain aspects of the classical Hamilton-Jacobi theory from the corresponding generalization in chapter 5.

Definition 2.1.19. Let E be a vector space, F a subspace and ω a symplectic form in E. The *ω-orthogonal complement* of F is the subspace defined by

$$F^{\perp} = \{u \in E | \omega(u, v) = 0, \text{ for all } v \in F\}.$$

We say that

(i). F is *isotropic* if $F \subset F^{\perp}$.

(ii). F is *coisotropic* if $F^{\perp} \subset F$.

(iii). F is *Lagrangian* if $E = F \oplus G$ where F and G are isotropic.

We have the following results, analogous to the ones for scalar products:

Proposition 2.1.20. *Let $F, G \subset E$ subspaces and ω symplectic in E. Then:*

(i). $F \subset G$ *implies* $G^{\perp} \subset F^{\perp}$.

(ii). $\dim E = \dim F + \dim F^{\perp}$.

(iii). $F = F^{\perp\perp}$.

(iv). $F^{\perp} \cap G^{\perp} = (F + G)^{\perp}$.

(v). $(F \cap G)^{\perp} = F^{\perp} + G^{\perp}$.

Proof. (i) and (iv) are trivial.

For (ii), consider $\omega^{\flat} : E \to E^*$. If $u \in F$, then $F^{\perp} \subset \ker \omega^{\flat}(u)$ and we get an induced map $\omega^{\flat}(u) : (E/F^{\perp}) \to \mathbb{R}$, which globally gives $\omega^{\flat}_F : F \to (E/F^{\perp})^*$. $\omega^{\flat}_F$ is injective by nondegeneracy and thus

$$\dim F \leq \dim(E/F^{\perp})^* = \dim(E/F^{\perp}) = \dim E - \dim F^{\perp}.$$

The inclusion $F \overset{j}{\hookrightarrow} E$ induces the natural map $E^* \overset{j^*}{\to} F^*$, and the kernel of the linear map $\omega^{\flat}_{F^*} = j^* \circ \omega^{\flat} : E \to F^*$ is exactly $F^{\perp}$, whence

$$\dim F = \dim F^* \geq \dim \operatorname{im} \omega^{\flat}_{F^*} = \dim E - \ker \omega^{\flat}_{F^*} = \dim E - \dim F^{\perp}$$

and the inequalities give the result.

For (iii), $F \subset F^{\perp\perp}$ is clear and by (ii) we have $\dim F = \dim F^{\perp\perp} = \dim E - \dim F^{\perp}$, whence $F = F^{\perp\perp}$.

To prove (v), we use (iii) and (iv) in

$$(F \cap G)^{\perp} = (F^{\perp\perp} \cap G^{\perp\perp})^{\perp} = (F^{\perp} + G^{\perp})^{\perp\perp} = F^{\perp} + G^{\perp}.$$

□

We give a characterization of Lagrangian spaces that is used in practice.

Proposition 2.1.21. *Let $F \subset E$ a subspace and ω symplectic in E. The following are equivalent:*

(i). F is Lagrangian.

(ii). $F = F^{\perp}$.

(iii). F is isotropic and $2 \dim F = \dim E$.

Proof. Assume (i), and let $u \in F^{\perp}$. We have $E = F \oplus G$ where F, G are isotropic, and we can write $u = v + w$ where $v \in F$ and $w \in G$. By isotropy, $w \in G^{\perp}$ and $w = u - v \in F^{\perp}$, so $w \in G^{\perp} \cap F^{\perp} = (G + F)^{\perp} = E^{\perp} = \{0\}$ by nondegeneracy of ω. This way $w = 0$, which gives $F^{\perp} \subset F$, and we have $F \subset F^{\perp}$ by hypothesis.

Assuming (ii), we obtain (iii) from 2.1.20(ii).

If we assume (iii), using 2.1.20(ii). and by isotropy, we have $F = F^{\perp}$. To construct G, we choose $v_1 \notin F$ and define $G_1 = \text{span }(v_1)$, which satisfies $F \cap G_1 = \{0\}$ and applying 2.1.20(v), $F + G_1^{\perp} = E$. Now let $v_2 \in G_1^{\perp}$, $v_2 \notin F + G_1$, and define $G_2 = \text{span }(v_1, v_2)$. Inductively, we obtain $F + G_k = E$ and $F \cap G_k = \{0\}$ by construction, which gives $F \oplus G_k = E$. We also have $G_2^{\perp} = (G_1 + \text{span }(v_2))^{\perp} = G_1^{\perp} \cap \text{span }(v_2)^{\perp} \supset \text{span }(v_1, v_2) = G_2$ and so on, whence G_k is isotropic, and the result follows by taking $G = G_k$. □

We globalize the definitions in 2.1.19 to submanifolds in the natural way. Let (M, ω) be symplectic, and consider an immersed submanifold $N \overset{j}{\hookrightarrow} M$. We say that N is isotropic, coisotropic or Lagrangian if the corresponding condition holds in T_pN for every $p \in N$, where we identify T_pN with a subspace of T_pM by means of T_pj.

We now relate the concept of Lagrangianity with symplectic maps. Let (M, ω) and (N, ρ) be symplectic manifolds, and consider the natural projections $\mathrm{pr}_1 : M \times N \to M$, $\mathrm{pr}_2 : M \times N \to N$. We define the 2-form $\omega \ominus \rho = \mathrm{pr}_1^*\omega - \mathrm{pr}_2^*\rho$. It is easy to verify that $\omega \ominus \rho$ is a symplectic form. Let $f : M \to N$ be a manifold diffeomorphism, and consider

$$j_f : \Gamma_f = \{(x, y) \in M \times N | y = f(x),\ x \in M\} \hookrightarrow M \times N \tag{2.23}$$

the graph of f. We have:

Theorem 2.1.22. *The following are equivalent:*

(i). f is symplectic.

(ii). $j_f^*(\omega \ominus \rho) = 0$.

(iii). $j_f : \Gamma_f \hookrightarrow M \times N$ is a Lagrangian submanifold of $(M \times N, \omega \ominus \rho)$.

Proof. Since $\Gamma_f = \{(x, f(x)) \in M \times N,\ \text{for } x \in M\}$ we can write the tangent space as

$$\mathrm{T}_{(x,f(x))}\Gamma_f = \{(v, \mathrm{T}_x f \cdot v) \in \mathrm{T}_xM \times \mathrm{T}_{f(x)}N,\ \text{for } v \in \mathrm{T}_xM\}.$$

Thus, $j_f^*(\omega \ominus \rho)((v, \mathrm{T}f \cdot v), (w, \mathrm{T}f \cdot w)) = \omega(v, w) - \rho(\mathrm{T}f \cdot v, \mathrm{T}f \cdot w) = (\omega - f^*\rho)(v, w)$ and the equivalence between (i) and (ii) follows.

Assuming (ii), we have $\mathrm{T}_{(x,f(x))}\Gamma_f \subset (\mathrm{T}_{(x,f(x))}\Gamma_f)^{\perp}$ by hypothesis, but $M \cong N \cong \Gamma_f$ whence $\dim \mathrm{T}_xM = \dim \mathrm{T}_{f(x)}N = \dim \mathrm{T}_{(x,f(x))}\Gamma_f$ and thus $2 \dim \mathrm{T}_{(x,f(x))}\Gamma_f = \dim(\mathrm{T}_xM \times \mathrm{T}_{f(x)}N)$, which implies (iii) by 2.1.21. It is clear that (iii) implies (ii), and the equivalence follows. □

2.2 Canonical Structures in Vector Bundles

A formal treatment of mechanics requires the use of manifolds, as the laws of nature are written in coordinates while their validity does not depend on them. The structure of vector bundle is also implicit in the classical formulations of mechanics, since the motion of a system is described in terms of the natural bundles of the manifold: the tangent bundle in the Lagrangian case, and the cotangent bundle in the Hamiltonian.

In this section, we define vector bundles and construct the canonical structures associated with them. These structures play a significant role in the geometrical formulation of mechanics (see chapter 3). Throughout the text, given a smooth manifold M, $\tau_M : \mathrm{T}M \to M$ will denote the tangent bundle of M, while the cotangent bundle will be written as $\pi_M : \mathrm{T}^*M \to M$.

2.2.1 Vector Bundles

Before studying the particular case of the tangent and cotangent bundles, we describe general vector bundles and some of the maps that one may always define on them.

Definition 2.2.1. A *vector bundle* is a triple (E, B, π), such that

(i). E, B are manifolds, and $\pi : E \to B$ is a surjective morphism.

(ii). For every $b \in B$ there exist a neighbourhood $U \ni b$ and a vector space F such that there is a diffeomorphism $\varphi : \pi^{-1}(U) \to U \times F$ making the following diagram commutative:

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & U \times F \arrow[dl, "\mathrm{pr}_1"] \\ & U & \end{tikzcd}$$

E is called the *space* of the fiber, B the *base* and π the *projection*. The diffeomorphisms φ are *local trivializations* of the bundle E.

We will assume that the vector spaces F in the definition are isomorphic to a common space. If $b \in B$, the subset $E_b = \pi^{-1}(b)$ is the *fiber* of b, and is isomorphic to F. The surjectivity of π implies that the tangent map $\mathrm{T}\pi$ is also surjective, whence it follows that the fibers E_b are regular submanifolds of E. We note that it is possible to consider more general spaces by relaxing the condition that F be a vector space, which gives the notion of a *fiber bundle*.

We can always construct a special type of coordinates in E, by considering coordinates in B and transporting them to E by some local trivialization. These are referred to as the *fiber coordinates* of the bundle.

Let $\pi : E \to B$ be a vector bundle. At each point $e \in E$, the tangent map $\mathrm{T}_e\pi : \mathrm{T}_eE \to \mathrm{T}_{\pi(e)}B$ is linear and gives a subspace $V_e(E) = \ker \mathrm{T}_e\pi$ in T_eE. We define the *vertical subbundle* as the subbundle of $\tau_E : \mathrm{T}E \to E$ obtained from the union of the $V_e(E)$ when $e \in E$. We write $VE = \ker \mathrm{T}\pi$ and refer to its elements as *vertical vectors*. We say that a field in E is *vertical* when it takes values in VE.

Vertical vectors can be seen as elements of the tangent space of the fibers of E. More precisely, a vector $v \in V_e(E)$ can be seen as an element of $\mathrm{T}_e(E_b)$, where $b = \pi(e)$ and $\mathrm{T}_e(E_b) \hookrightarrow \mathrm{T}_eE$. We can make this identification by using coordinates: let $(b; e)$ denote the fiber coordinates from a trivialization of E, and let $(b, e; u, v)$ be the corresponding natural coordinates in $\mathrm{T}E$. It is clear that vertical vectors are of the form $(b, e; 0, v)$. Adapting the coordinates above to $E_b \hookrightarrow E$ and $\mathrm{T}(E_b) \hookrightarrow \mathrm{T}E$, the map

$$\mathrm{T}_e(E_b) \to \mathrm{T}_eE, \quad (e; w) \mapsto (b, e; 0, w) \tag{2.24}$$

gives the desired identification.

It is known that the tangent space of a vector space V at a point $x \in V$ may be identified with V via the map

$$\lambda_x : V \to \mathrm{T}_x V, \quad v \mapsto [t \mapsto x + tv]. \tag{2.25}$$

Since the fibers of $VE \subset TE$ are vector spaces, by the observation above we have for each $v_b \in E_b$ an isomorphism

$$E_b \to \mathrm{T}_{v_b}(E_b) = V_{v_b}E, \quad w_b \mapsto \lambda_{v_b}(w_b). \tag{2.26}$$

Changing the vector $v_b \in E_b$, we get a map

$$\lambda_E(v_b, w_b) = \lambda_{v_b}(w_b) \tag{2.27}$$

known as the *vertical lift*, since $\lambda : E$ lifts vectors on the fibers to the vertical bundle. Note that λ_E is defined only for those pairs $(v, w) \in E \times E$ such that $\pi(v) = \pi(w)$. This subset of E is actually a vector bundle, the *product bundle* which is written $\mathrm{pr}_1 : E \times_B E \to E$.

The vertical lift defines a characteristic vector field $\Delta_E \in \mathfrak{X}(E)$, the *Liouville field* of E, which is given by

$$\Delta_E(v) = \lambda_E(v, v) \tag{2.28}$$

for $v \in E$. We will use the Liouville field to define the *energy function* of the Lagrangian formulation of mechanics (see chapter 3).

Finally, we define the *fiber derivative* of a function on the bundle $f : E \to \mathbb{R}$. The restriction $f_b \equiv f|_{E_b} : E_b \to \mathbb{R}$ is a map between vector spaces, and we have at each point $v_b \in E_b$ the usual derivative

$$Df_b(v_b) : E_b \to \mathbb{R} \tag{2.29}$$

so $Df_b(v_b) \in E_b^*$. Globally, we get a map

$$\mathcal{F}f : E \to E^*, \quad v_b \mapsto Df_b(v_b) \tag{2.30}$$

from a bundle to its dual, known as the *fiber derivative* of f.

Naturally, we can also consider the derivative of the fiber derivative, which gives a bilinear form $D^2 f_b(v_b) : E_b \times E_b \to \mathbb{R}$ and globally defines the *fiber hessian*

$$\mathcal{F}^2 f : E \to E^* \otimes E^*, \, v_b \mapsto D^2 f_b(v_b). \tag{2.31}$$

Under some regularity conditions over f which will be translated to a certain restriction on the Hessian, the map may give rise to a local or global diffeomorphism (see chapter 3).

It is useful to remember the expressions of the objects defined in coordinates:

$$\lambda_E(b; u, v) = (b, u; 0, v) \tag{2.32}$$

$$\Delta_E = v^i \frac{\partial}{\partial v^i} \tag{2.33}$$

$$\mathcal{F}f(b; v^1, \ldots, v^m) = \left(b; \frac{\partial f}{\partial v^1}, \ldots, \frac{\partial f}{\partial v^m}\right) \tag{2.34}$$

$$\mathcal{F}^2 f(b; v)(u^1, \ldots, u^m, \, w^1, \ldots, w^m) = \frac{\partial^2 f}{\partial v^i \partial v^j}(b, v) u^i w^j \tag{2.35}$$

2.2.2 The Tangent Bundle

We now work on the tangent bundle $\tau_M : \mathrm{T}M \to M$ of a manifold M. Since $\mathrm{T}M$ has a manifold structure, we can consider its tangent bundle $\tau_{\mathrm{T}M} : \mathrm{T}(\mathrm{T}M) \to \mathrm{T}M$. There exists a map $J : \mathrm{T}(\mathrm{T}M) \to \mathrm{T}(\mathrm{T}M)$ defined by

$$J(w_v) = \lambda_{\mathrm{T}(\mathrm{T}M)}(v, \mathrm{T}_v(\tau_M) \cdot w_v) \in V_v(\mathrm{T}M) \subset \mathrm{T}(\mathrm{T}M), \text{ for } w_v \in \mathrm{T}_v(\mathrm{T}M)$$

which is linear in the fibers. J is the *vertical endomorphism* of $\mathrm{T}(\mathrm{T}M)$.

Introducing coordinates (q) in M, and considering the corresponding natural coordinates $(q; v)$ in $\mathrm{T}M$ and $(q, v; a, b)$ in $\mathrm{T}(\mathrm{T}M)$, the vertical endomorphism is written as

$$J(q, v; a, b) = (q, v; 0, a). \tag{2.36}$$

We may think of J as a $(1, 1)$ tensor field on $\mathrm{T}M$ which in coordinates is given by

$$J = \frac{\partial}{\partial v^i} \otimes dq^i. \tag{2.37}$$

It is clear that $J^2 = 0$ and $\mathrm{Im}J = \ker J$, where the expressions make sense at each fiber.

Consider now the map $\kappa_M : \mathrm{T}(\mathrm{T}M) \to \mathrm{T}(\mathrm{T}M)$ defined in the natural coordinates as

$$\kappa_M(q, v; a, b) = (q, a; v, b). \tag{2.38}$$

Clearly, $\kappa_M^2 = \mathrm{Id}_{\mathrm{T}(\mathrm{T}M)}$. κ_M is the *canonical involution* of $\mathrm{T}(\mathrm{T}M)$. The involution satisfies the relations

$$\tau_{\mathrm{T}M} \circ \kappa_M = \mathrm{T}(\tau_M), \quad \mathrm{T}(\tau_M) \circ \kappa_M = \tau_{\mathrm{T}M} \tag{2.39}$$

and gives a vector bundle isomorphism between $\tau_{\mathrm{T}M} : \mathrm{T}(\mathrm{T}M) \to \mathrm{T}M$ and $\mathrm{T}\tau_M : \mathrm{T}(\mathrm{T}M) \to \mathrm{T}M$, the two bundle structures of $\mathrm{T}(\mathrm{T}M)$.

A field $X \in \mathfrak{X}(M)$ defines a flow F_t that exists for sufficiently small values of t. F_t is a diffeomorphism of M on itself, which is seen as a one parameter group of transformations on M. Taking the tangent map, the morphisms $T(F_t)$ inherit the group properties of F_t, and there exists a field X^{T} which the infinitesimal generator of the group. By definition, the flow of X^{T} at time t is $T(F_t)$. X^{T} is the *canonical lift* or *tangent lift* of X to $\mathrm{T}M$. In the natural coordinates $(q; v)$, if $X = X^i \dfrac{\partial}{\partial q^i}$ then

$$X^{\mathrm{T}} = X^i \frac{\partial}{\partial q^i} + \left(\frac{\partial X^i}{\partial q^j} v^j \right) \frac{\partial}{\partial v^i}. \tag{2.40}$$

We also have the relations

$$\mathrm{T}\tau_M \circ X^{\mathrm{T}} = X \circ \tau_M, \quad X^{\mathrm{T}} = \kappa_M \circ \mathrm{T}X, \tag{2.41}$$

which are easy to verify. X being a section of τ_M implies that $\mathrm{T}X$ is a section of $\mathrm{T}\tau_M$, as $\mathrm{T}\tau_M \circ \mathrm{T}X = \mathrm{T}(\tau_M \circ X) = \mathrm{T}\mathrm{Id}_M = \mathrm{Id}_{\mathrm{T}M}$. This way, the second condition gives as an interpretation of the tangent lift as the tangent map, via the bundle isomorphism κ. The firs one simply states that X and X^{T} are τ_M-related (see chapter 5).

It is clear that we can also lift X to $\mathrm{T}M$ by means of the vertical lift which was explained earlier. We define the *vertical lift* of X by

$$X^V(v_q) = \lambda_{\mathrm{T}(\mathrm{T}M)}(v_q, X(q)) \in V_{v_q}(\mathrm{T}M). \tag{2.42}$$

Thus, we obtain

$$X^V = X^i \frac{\partial}{\partial v^i} \tag{2.43}$$

and clearly

$$X^V = J \circ X^{\mathrm{T}}. \tag{2.44}$$

We finish the section on the tangent bundle by making a few comments on second order differential equations. Let $\sigma : I \to \mathrm{T}M$ be a path in $\mathrm{T}M$. σ can be written as $\sigma = \gamma'$ for some $\gamma : I \to M$ if and only if $\sigma = \mathrm{T}\tau_M \circ \sigma'$ as can be easily verified in coordinates. Given a field $X \in \mathfrak{X}(\mathrm{T}M)$, its integral curves σ will be lifts of curves γ in M if and only if

$$\mathrm{T}\tau_M \circ X = \mathrm{Id}_{\mathrm{T}M}. \tag{2.45}$$

This follows from the previous equivalence.

Observe that the last condition is equivalent to X being a section of the bundle $\mathrm{T}\tau_M$: $\mathrm{T}(\mathrm{T}M) \to \mathrm{T}M$. This requirement is known as a *second order condition*, and the fields that satisfy it will be called *second order fields* or *second order differential equations*. Indeed, a second order field X defines the second order differential equation

$$\gamma'' = X \circ \gamma' \tag{2.46}$$

for paths $\gamma : I \to M$ in M.

2.2.3 The Cotangent Bundle

The cotangent bundle T^*M has a natural symplectic structure. We first define a distinguished 1-form θ on T^*M and then take the exterior derivative to obtain a symplectic form ω.

Theorem 2.2.2. *Let θ be a map $\theta : \mathrm{T}^*M \to \mathrm{T}^*(\mathrm{T}^*M)$ defined at each $p \in \mathrm{T}^*(M)$ as*

$$\theta(p) = {}^t(\mathrm{T}_p\pi_M) \cdot p \tag{2.47}$$

for $p \in \mathrm{T}^*M$. Then θ is a smooth section of $\pi_{\mathrm{T}^*M} : \mathrm{T}^*(\mathrm{T}^*M) \to \mathrm{T}^*(M)$ and the 2-form $\omega = -d\theta$ is nondegenerate. In other words, $\theta \in \Omega^1(\mathrm{T}^*M)$ and (T^*M, ω) is a symplectic manifold.

Proof. It is clear that $\pi_{T^*M} \circ \theta = \mathrm{Id}_{\mathrm{T}^*M}$. Let $\xi_p \in \mathrm{T}_p(\mathrm{T}^*M)$, then

$$\langle \theta(p), \xi_p \rangle = \langle {}^t(T_p\pi_M) \cdot p, \xi_p \rangle = \langle p, \mathrm{T}_p\pi_M \cdot \xi_p \rangle.$$

Taking the natural coordinates $(q^i; p_i)$ on T^*M, the above expression reads $\theta = p_i dq^i$, which is clearly differentiable. In the same coordinates, we have

$$\omega = -d\theta = dq^i \wedge dp_i$$

and the associated matrix ω_{ij} of ω has entries

$$\omega_{ij} = \begin{cases} 1, & \text{when } i + m = j \\ -1, & \text{when } i = j + m \\ 0, & \text{otherwise} \end{cases}$$

Thus, $\det(\omega_{ij}) = \det \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix} = 1$ and ω is nondegenerate. $\square$

The forms above are known as the *canonical forms* on the cotangent bundle. We have that (T^*M, ω) is a symplectic manifold. Observe that the natural coordinates on T^*M directly give the symplectic coordinates for the canonical symplectic structure. The following characterisation will be useful later on:

Proposition 2.2.3. *The canonical 1-form θ on T^*M is the unique 1-form such that, for any 1-form α on M,*

$$\alpha^*\theta = \alpha \tag{2.48}$$

where we think of α as a map $\alpha : M \to \mathrm{T}^*M$ which is a section of T^*M.

Proof. Let $v_q \in \mathrm{T}_qM$, then

$$\begin{aligned}\langle \alpha^*\theta, v_q\rangle &= \langle \theta_{\alpha(q)}, \mathrm{T}_q\alpha \cdot v_q\rangle \\ &= \langle \tau_{\mathrm{T}^*M} \circ \mathrm{T}_q\alpha \cdot v_q, \mathrm{T}_{\alpha(q)}\pi_M \circ \mathrm{T}_q\alpha \cdot v_q\rangle \\ &= \langle \alpha(q), \mathrm{T}_q(\pi_M \circ \alpha) \cdot v_q\rangle \\ &= \langle \alpha(q), v_q\rangle\end{aligned}$$

whence it follows that $\alpha^*\theta = \alpha$.

θ is also unique, because $\alpha(q)$ and $\mathrm{T}_q\alpha \cdot v_q$ span all of T^*_qM and $\mathrm{T}_{\alpha(q)}(\mathrm{T}^*M)$ for variable α and v_q. □

It is always possible to generate symplectic mappings on the tangent bundle from diffeomorphisms in the base manifold.

Theorem 2.2.4. *Let $\varphi : M \to N$ be a diffeomorphism, and $(\mathrm{T}^*M, \omega = -d\theta)$, $(\mathrm{T}^*N, \rho = -d\vartheta)$ the canonical symplectic structures on T^*M and T^*N. Define the* canonical lift *of φ as $\mathrm{T}^*\varphi : \mathrm{T}^*M \to \mathrm{T}^*N$ such that $(\mathrm{T}^*\varphi)_x = ({}^t\mathrm{T}_x\varphi)^{-1}$ at each $x \in M$. Then $\mathrm{T}^*\varphi$ is a symplectomorphism and in fact*

$$(\mathrm{T}^*\varphi)^*(\vartheta) = \theta. \tag{2.49}$$

Proof. Let $p \in \mathrm{T}^*_xM$, $\xi \in \mathrm{T}_p(\mathrm{T}^*M)$. By definition:

$$\begin{aligned}\langle (\mathrm{T}^*\varphi)^*\vartheta, \xi\rangle &= \langle \vartheta(\mathrm{T}^*\varphi(p)), \mathrm{T}(\mathrm{T}^*\varphi) \cdot \xi\rangle \\ &= \langle \mathrm{T}^*\varphi(p), \mathrm{T}\pi_N \circ \mathrm{T}(\mathrm{T}^*\varphi) \cdot \xi\rangle \\ &= \langle \mathrm{T}^*\varphi(p), \mathrm{T}(\pi_N \circ \mathrm{T}^*\varphi) \cdot \xi\rangle \\ &= \langle \mathrm{T}^*\varphi(p), \mathrm{T}(\varphi \circ \pi_M) \cdot \xi\rangle \\ &= \langle \mathrm{T}^*\varphi(p), \mathrm{T}\varphi \circ \mathrm{T}\pi_M \cdot \xi\rangle \\ &= \langle p, \mathrm{T}\pi_M \cdot \xi\rangle \\ &= \langle \theta, \xi\rangle\end{aligned}$$

□

Writing the diffeomorphism $\varphi : M \to N$ in coordinates as

$$(Q^1, \cdots, Q^m) \mapsto (q^1, \cdots, q^m)$$

the canonical lift in the natural coordinates is written

$$(Q^1, \cdots, Q^m; P_1, \cdots, P^m) \mapsto (q^1, \cdots, q^m; p_1, \cdots, p^m)$$

where

$$p_i = \frac{\partial Q^j}{\partial q^i} P_j. \tag{2.50}$$

As we already noted, a field $X \in \mathfrak{X}(M)$ defines a flow that gives a diffeomorphism of M on itself. In particular, we may consider the canonical lift of the flow to obtain a symplectomorphism on the cotangent bundle. Because these symplectic maps inherit the group properties of the flow, they can be seen as the flow of a unique field $X^{\mathrm{T}*} \in \mathfrak{X}(\mathrm{T}^*M)$, the *canonical lift* of X.

If $X = X^i \dfrac{\partial}{\partial q^i}$, it is easy to see that in natural coordinates the canonical lift is written

$$X^{\mathrm{T}*} = X^i \frac{\partial}{\partial q^i} - \left(p_j \frac{\partial X^j}{\partial q^i} \right) \frac{\partial}{\partial p_i}. \tag{2.51}$$

Chapter 3

Analytical Mechanics

In this chapter, we present Newtonian mechanics and the formulations of Lagrange and Hamilton in the language of manifolds. This choice will simplify much of the work and will reveal the most prominent features of the classical theory in a very elegant way. The framework will prove itself essential when we carry out the generalizations of the principal topic in later chapters.

For the mathematical foundation of Newtonian mechanics, we follow mostly [Arn74, part I] and [Gra15]. For the Lagrangian and Hamiltonian formulations, we take the approach of [AM78, chapter 3]. Many of the examples are taken from [LL77, volume 1].

3.1 Newtonian Mechanics

In this section, we focus on the mathematical foundations of *Newtonian mechanics*. The theory is interested in describing the evolution in time of a system of *particles* in space. A particle may be thought of as a body whose dimensions may be neglected in describing its motion [LL77, volume 1, chapter 1]. In his celebrated *Principia*, Newton presented a quantitative theory of mechanics in agreement with the qualitative investigations of previous scientists [AM78]. His work allowed solving completely a series of important problems in mechanics, and its great exactness and wide range of applicability had a dramatic impact in the history of science.

3.1.1 Experimental Facts and the Galilean Structure

There are some basic experimental facts that lie at the foundation of mechanics. These facts are only approximately true and it can be shown by experiment that they do not hold in general. They do hold under certain conditions, in the so called *classical limit* when the velocities of the particles in question are "small" and their size is "big" in some sense [LL77, volumes 2, 3]. We take the following principles as axioms:

1 ***Space and Time****. There is* time*, which is one-dimensional and homogeneous. There is also* space*, which is three-dimensional, homogeneous, isotropic and Euclidean. Together, they form* spacetime*. A* coordinate system *introduces coordinates for the points in spacetime. The coordinates are written* $(t, \boldsymbol{r}) \in \mathbb{R} \times \mathbb{R}^3$ *and the* motion *of a particle is described in terms of maps* $t \mapsto \boldsymbol{r}(t)$.

2 ***Galileo's Principle of Relativity****. There exists a particular class of coordinate systems, the* inertial coordinate systems *with the following properties:*

(i). The laws of nature *are expressed in the same way in every inertial coordinate system.*

(ii). A coordinate system in uniform rectilinear motion with respect to an inertial system is also inertial.

3 ***Newton's Principle of Determinacy****. The motion of a mechanical system is determined by the positions and velocities of its particles at some moment of time.*

The first principle tells us about the structure of spacetime, which must necessarily be a four-dimensional affine space. In it, there must be a well-defined map which measures time differences, and a scalar product to measure distances in space. We arrive to the following definition:

Definition 3.1.1. A *Galilean spacetime* is a triple (A, τ, g) where:

- A is an affine space over $E \cong \mathbb{R}^4$.
- $\tau : E \to \mathbb{R}$ is an injective linear form, the *time interval.*
- g is a scalar product on $\ker \tau \subset E$.

We call the points in A *world events* or simply *events.* Because A is an affine space over E, we have a map $A \times A \to E$, $(x, y) \mapsto y - x$ with the usual properties, and E can be seen as the vector space of translations of A. Given two events $x, y \in A$ we define the *time interval* from x to y to be the number $\tau(y-x)$. When $y-x \in \ker \tau$ we say that the two events are *simultaneous.* $\ker \tau$ is a hyperplane in E, whose elements are the *spatial translations.* The points in the affine hyperplane $x + \ker \tau \subset A$ are simultaneous with x. $(\ker \tau, g)$ is a Euclidean space, and the scalar product defines a Euclidean norm $\|\cdot\|$. Thus, we may speak of the distance between two simultaneous events.

We can make Galilean spacetimes into a category where the objects are Galilean spacetimes and the morphisms are the maps between Galilean spacetimes that preserve the Galilean structure. More precisely, let (A, τ, g), (B, T, h) be Galilean spaces over E and F, respectively.

Definition 3.1.2. A *Galilean transformation* is a map $f : A \to B$ such that

- f is affine. We denote its linear part by $\hat{f}$, so $f(y) = f(x) + \hat{f}(y - x)$ for $x, y \in A$.
- f preserves time intervals, i.e. $(T \circ \hat{f})(v) = \tau(v)$ for every $v \in E$.
- f preserves the scalar product, i.e. $h(\hat{f}(v), \hat{f}(w)) = g(\hat{f}(v), \hat{f}(w))$ for every $v, w \in \ker \tau$.

It is clear that the composition of Galilean transformations is again a Galilean transformation. In particular, we define the *Galilean group* of (A, τ, g) as the set of its Galilean automorphisms.

We observe that Galilean transformations are always bijective. Taking the points $x_\circ, \in A$, $f(x_\circ) \in B$ as origins, it suffices to show that $\hat{f}$ is injective, which will imply surjectivity by linear algebra. Assume $v \in \ker \hat{f}$ and write v in the form $v = a + b$ where $a \notin \ker \tau$ and $b \in \ker \tau$. Then $0 = T\hat{f}(v) = \tau(v) = \tau(a)$ so $a = 0$, and $0 = h(\hat{f}(v), \hat{f}(v)) = g(b, b)$ whence $b = 0$ and the kernel of $\hat{f}$ is trivial.

In fact, Galilean spaces are always isomorphic to each other. To show this, we find a Galilean transformation $\psi : A \to B$. Fixing two points $x_\circ \in A$, $y_\circ \in B$, we define $\psi(x_\circ) = y_\circ$ and it suffices to construct $\hat{\psi}$. Let $\{u_i\}_{i=1,2,3}$ and $\{v_j\}_{j=1,2,3}$ be orthonormal bases in $(\ker \tau, g)$ and $(\ker T, h)$ respectively. We extend them to bases $\{u_i\}_{i=0,1,2,3}$ and $\{v_j\}_{j=0,1,2,3}$ of the whole spaces E and F. Define $\tilde{\psi}$ by

$$\tilde{\psi}(u_\circ) = \frac{\tau(u_\circ)}{T(v_\circ)} v_\circ, \quad \tilde{\psi}(u_i) = v_i (i = 1, 2, 3) \tag{3.1}$$

which is clearly Galilean.

In particular, every Galilean spacetime is isomorphic to the space $(\mathbb{R}^4, \mathrm{pr}_1, g_\mathrm{e})$, where pr_1 is the projection on the first factor and g_e is the standard product in $\{0\} \times \mathbb{R}^3$. We call it *Galilean coordinate space*.

It is easy to see that the Galilean transformations of Galilean coordinate space are uniquely written as the composition of three maps:

- A translation:
$$(t, \mathbf{r}) \mapsto (t + u_\circ, \mathbf{r} + \mathbf{u}), \text{ where } (u_\circ, \mathbf{u}) \in \mathbb{R}^4$$
- A *uniform motion* with velocity $\mathbf{v}$:
$$(t, \mathbf{r}) \mapsto (t, \mathbf{r} + t\mathbf{v}), \text{ where } \mathbf{v} \in \mathbb{R}^3$$
- A spatial rotation:
$$(t, \mathbf{r}) \mapsto (t, R\mathbf{r}), \text{ where } R \in O(3, \mathbb{R})$$

In particular, we see that the Galilean group is a Lie group of dimension 10 diffeomorphic to $\mathbb{R}^4 \times \mathbb{R}^3 \times O(3, \mathbb{R})$.

3.1.2 Motion and Dynamics

Let $C \subset A$ a connected curve in Galilean spacetime A. We recall that affine spaces can be regarded as smooth manifolds, and we may consider the tangent space of C in A, which is isomorphic to the vector space E. We say that C is a *world line* when the tangent space at each point of the curve $\mathrm{T}_x C \subset \mathrm{T}_x A$ is transversal to the spatial translations, where we see $\ker \tau$ as a subspace of $\mathrm{T}_x A \cong E$.

The transversality condition suggests the use of world lines to represent the history of a particle that moves through space. The evolution of a particle in space can be described in coordinates by maps of the form

$$\gamma : I \to \mathbb{R}^3, \quad t \mapsto \mathbf{r}(t), \tag{3.2}$$

while in Galilean spacetime the whole trajectory reads

$$\tilde{\gamma} : I \to \mathbb{R}^4, \quad t \mapsto (t, \mathbf{r}(t)), \tag{3.3}$$

and the image of $\tilde{\gamma}$ is clearly a world line.

Conversely, if a world line C has an appropriate degree of differentiability, there exists a parametrization $\xi : J \to A$ with the same regularity and we can write

$$\tau \circ \mathrm{T}_s \xi \neq 0 \tag{3.4}$$

for all $s \in J$. Introducing Galilean coordinates, we write $\hat{\xi}(s) = (\xi_1(s), \boldsymbol{\xi}_2(s))$, and define $t(s) = \xi_1(s)$. If $'$ denotes differentiation with respect to s, we have

$$t'(s) = \xi_1'(s) = \widehat{\tau \circ \mathrm{T}_s \xi} \neq 0 \tag{3.5}$$

and by the inverse function theorem we can write s locally as a function of t, $s = s(t)$. Thus,

$$\hat{\xi}(s(t)) = (t, \boldsymbol{\xi}_2(s(t))) \tag{3.6}$$

and ξ is always the graph of a path $\gamma : I \to \mathbb{R}^3$ in the Galilean coordinate system. Since the coordinate system was arbitrary, we find that the property does not depend on the choice of coordinates.

Assume now that the world line C describes the history of a particle as it moves in space. The path $\gamma : I \to \mathbb{R}^3$ of the previous property is the *motion* or *dynamical trajectory* of the particle. The *speed* of the particle at time t is the derivative of the motion $\dot{\gamma}(t) = D\gamma(t) \in \mathbb{R}^3$ and the *acceleration* is the derivative of the speed $\ddot{\gamma}(t) = D^2\gamma(t) \in \mathbb{R}^3$, where we have used dots to denote differentiation with respect to t.

To describe a mechanical system composed of N particles, we consider the motion of each particle

$$\mathbf{r}_i : I \to \mathbb{R}^3, \quad (i = 1, \ldots, N) \tag{3.7}$$

and construct the map

$$\mathbf{r} = \mathbf{r}_1 \times \cdots \times \mathbf{r}_N : I \to \mathbb{R}^n \tag{3.8}$$

where $n = 3N$. We say that the mechanical system has n *degrees of freedom*, and we call $\mathbb{R}^n$ the *configuration space.*

We now examine Newton's principle of determinacy to study the dynamics of a general system of particles. According to the principle, the motion of a system of particles is completely determined by the position and velocity of the system at an arbitrary time $(t_\circ, \mathbf{r}(t_\circ), \dot{\mathbf{r}}(t_\circ))$. Consequently, there must exists a map $\mathbf{f} : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\ddot{\mathbf{r}} = \mathbf{f}(t, \mathbf{r}, \dot{\mathbf{r}}) \tag{3.9}$$

Equation 3.9 is classically known as *Newton's equation.* Function $\mathbf{f}$ represents the physical *forces* that act upon the system and would be determined by experiment. Conversely, under certain smoothness conditions over $\mathbf{f}$, the theorem of existence and uniqueness of ordinary differential equations shows that an initial position and velocity determine a motion of the system.

Now that we have the equations of motion, which should be regarded as a law of nature, we can apply Galileo's relativity principle and examine the constraints that the principle imposes on the form of $\mathbf{f}$.

Newton's equations should take the same form in every inertial coordinate system. Since Galilean transformations are the maps that take us from one inertial system to another, by the principle we have that they should transform world lines of the system into world lines. In other words, if γ is a solution of Newton's equations, then $\varphi \circ \gamma$ is also a solution, where φ is a Galilean transformation.

Applying this principle to time translations, we see that $\mathbf{f}$ cannot depend explicitly on time. It is easy to see that invariance with respect to spatial translations implies that $\mathbf{f}$ must depend on the relative spatial coordinates $\mathbf{r}_i - \mathbf{r}_j$, while invariance with respect to uniform motion gives a dependence with respect to the relative velocities $\dot{\mathbf{r}}_i - \dot{\mathbf{r}}_j$. Finally, rotational invariance shows that $\mathbf{f}$ is such that $\mathbf{f}(R\mathbf{r}, R\dot{\mathbf{r}}) = R\mathbf{f}(\mathbf{r}, \dot{\mathbf{r}})$, for $R \in O(3, \mathbb{R})$, where we have written $R\mathbf{r} = (R\mathbf{r}_1, \ldots, R\mathbf{r}_N)$.

One should keep in mind that the forces have this form only for a system that comprehends the whole universe. They will have an approximate form when the subsystem in question does not interact strongly with the surroundings.

As an example, consider a system which is divided into two subsystems I and II. The equations of motion determine the trajectories of the whole system. In particular, the motion in subsystem II can be integrated, which gives the coordinates in II as a function of time. If we substitute this dependence into the equations of motion for the coordinates in I, we see that the forces acting on the subsystem depend explicitly on time.

It is experimentally verified that the acceleration of particles affected by forces *of the same kind* is the same modulo a constant factor which is an intrinsic of each particle, the *mass* m. If we write the force acting on particle i as $\mathbf{f}_i = \dfrac{\mathbf{F}_i}{m_i}$, where m_i is the mass of the particle, Newton's equations take their usual form

$$m_i\ddot{\mathbf{r}}_i = \mathbf{F}_i. \tag{3.10}$$

Defining the *linear momentum* of particle i as $\mathbf{p}_i = m_i\dot{\mathbf{r}}_i$, the equation is also written as

$$\dot{\mathbf{p}}_i = \mathbf{F}_i \tag{3.11}$$

which is *Newton's second law.*

From the previous observation it follows immediately that when no force is acting on a particle, the latter must move at a constant velocity, giving a world line which is exactly a straight line. This fact is known as the *principle of inertia* or *Newton's first law.*

3.1.3 Newtonian mechanics in a manifold

At this point, and in order to present the classical functions in mechanics in the most adequate way, it is convenient to consider spaces which are more general than the Euclidean. Instead of $\mathbb{R}^n$, we consider a configuration space that is a smooth manifold M. The differentiable structure of M allows for a description of the motion of the system in terms of coordinates which are not necessarily Cartesian.

Recalling the Galilean principles, we know that space is Euclidean, which means that our generalized configuration space must be Riemannian, i.e., there is a metric g defined on M. The metric defines a natural covariant derivative on the manifold, the *Levi-Civita connection* ∇, which is characterised by being torsion free and Riemannian. Explicitly, we have

$$\nabla_X Y - \nabla_Y X = [X, Y], \quad \nabla g = 0. \tag{3.12}$$

We perform our calculations with this covariant derivative, which gives a natural way to differentiate vector fields in the manifold. Because a Riemannian metric is nondegenerate, we have the isomorphism $g^\flat : \mathrm{T}M \to \mathrm{T}^*M$ and its inverse $g^\sharp : \mathrm{T}^*M \to \mathrm{T}M$.

The forces of Newtonian mechanics can be thought of as maps that place a vector at each point of the configuration space, consequently the corresponding objects in the manifold formulation are vector fields on M, or *force fields.* By $g^\flat$, we can turn force fields into 1-forms, and vice versa. We make the following definition:

Definition 3.1.3. A *Newtonian dynamical system* is a triple $(M, g, \mathcal{F})$ where

- M is a differentiable manifold, the *configuration space.*
- g is a Riemannian metric on M, the *kinetic energy.*
- $\mathcal{F} : M \to \mathrm{T}^*M$ is a 1-form on M, the *force.*

The *force field* of a Newtonian system is defined as $F = g^\sharp \circ \mathcal{F} \in \mathfrak{X}(M)$.

Let $\gamma : I \to M$ be a path in M. The velocity of γ, $\gamma' = \mathrm{T}\gamma \circ \dfrac{d}{dt} \in \mathfrak{X}(M)$ is a vector field on M which defines a 1-form $p_\gamma = g^\flat \circ \gamma' \in \Omega^1(M)$. We refer to p_γ as the *momentum* associated with γ.

We rewrite Newton's equations regarding the covariant derivative of γ' as the acceleration of the particle:

$$\nabla_t \gamma' = F \circ \gamma \tag{3.13}$$

where ∇_t is the covariant derivative along γ. Observe that

$$\langle g^\flat \nabla_t \gamma', X\rangle = g(\nabla_t \gamma', X) = \nabla_t g(\gamma', X) - g(\gamma', \nabla_t X) = \nabla_t \langle p_\gamma, X\rangle - \langle p_\gamma, \nabla_t X\rangle = \langle \nabla_t p_\gamma, X\rangle$$

and in terms of forms 3.13 reads

$$\nabla_t p_\gamma = \mathcal{F} \circ \gamma. \tag{3.14}$$

As before, the solutions of Newton's equation are the *dynamical trajectories* of the system. Observe that in the new language, the principle of inertia is now a differential geometric statement: when no force acts on a particle, the corresponding trajectory is a geodesic of the metric. Thus, geodesics correspond to non accelerated motions inside the manifold.

Introducing coordinates (q^i) on M, we can write $g = g_{ij}\, dq^i \otimes dq^j$. We let Γ^k_{ij} denote the Christoffel symbols of the connection. If the path γ is expressed in coordinates as $(q^i(t))$, then γ' is written $(q^i(t); \dot{q}^i(t))$ in the natural coordinates in $\mathrm{T}M$. We also have p_γ as $(q^i(t); p_i(t)) = (q^i(t); g_{ij}(q(t))\dot{q}^j(t))$ in the natural coordinates of T^*M. Writing $\mathcal{F} = F_i dq^i$ and $F = F^i \frac{\partial}{\partial q^i}$, we obtain the equations of motion in coordinates

$$\ddot{q}^k + \Gamma^k_{ij}(q)\dot{q}^i\dot{q}^j = F^k(q), \quad \dot{p}_k - \Gamma^j_{ik}(q)\dot{q}^i p_j = F_k(q). \tag{3.15}$$

We now define several objects that have been known in mechanics for a long time. The *work* of $\mathcal{F}$ along the path γ is the line integral of $\mathcal{F}$ along γ, i.e.

$$\int_I \gamma^*(\mathcal{F}) = \int_I \langle \mathcal{F} \circ \gamma, \gamma' \rangle dt. \tag{3.16}$$

The *kinetic energy* of the system is the quadratic form associated with g. We see it as a function on the tangent bundle $T : \mathrm{T}M \to \mathbb{R}$ defined by $T(v) = \dfrac{1}{2} g(v, v)$.

Note that for a trajectory of the system, we have

$$\frac{d}{dt}(T \circ \gamma')(t) = \frac{1}{2}\nabla_t(g(\gamma', \gamma')) = g(\nabla_t \gamma', \gamma') = g(F \circ \gamma, \gamma') = \langle \mathcal{F} \circ \gamma, \gamma' \rangle \tag{3.17}$$

which shows that the work of $\mathcal{F}$ is simply the difference in the kinetic energy, that is,

$$\int_{t_1}^{t_2} \gamma^*(\mathcal{F}) = T(\gamma'(t_2)) - T(\gamma'(t_1)). \tag{3.18}$$

In nature, there exist special forces that can be derived from a scalar function or *potential*. Examples of such forces include gravity, the force of a spring or the electrostatic force. We say that $\mathcal{F}$ is *conservative* when $\mathcal{F} = -dV$ for some function $V : M \to \mathbb{R}$. V is the *potential energy*. In a connected manifold, the potential energy is determined up to a constant. For the force field, the condition is $F = -\mathrm{grad} V$.

If $\mathcal{F} = -dV$, we have that

$$\int_I \gamma^*(\mathcal{F}) = \int_I \gamma^*(-dV) = -\int_I d\gamma^*(V) = -\int_{\partial I} \gamma^*(V) = V(\gamma(t_1)) - V(\gamma(t_2)) \tag{3.19}$$

where in the last equality we have used Stokes' theorem. Thus, the work of a conservative force does not depend on the path of integration, but only on the difference of the potential energy between the boundaries.

The variation of the potential energy along a path is given by:

$$\frac{d}{dt}(V \circ \gamma)(t) = \mathcal{L}_{\gamma'(t)}V = \langle dV \circ \gamma(t), \gamma'(t) \rangle = -\langle \mathcal{F} \circ \gamma(t), \gamma'(t) \rangle = -g(F \circ \gamma(t), \gamma'(t)). \quad (3.20)$$

When a system is conservative, we define the *mechanical energy* of the system as the function $E : \mathrm{T}M \to \mathbb{R}$, $E = T + V \circ \tau_M$. Let γ be a path in M, not necessarily a trajectory of the system. By the above considerations,

$$\frac{d}{dt}(E \circ \gamma') = \frac{d}{dt}(T \circ \gamma') + \frac{d}{dt}(V \circ \gamma) = g(\nabla_t \gamma' - F \circ \gamma, \gamma'). \quad (3.21)$$

which shows that E is constant along γ' for the trajectories γ of a conservative system. We say that the mechanical energy of a conservative system is *conserved*, or that it is a *conserved quantity*.

Note that our formulation has only included forces that depend on the spatial coordinates. In general, one may consider objects along the projection τ_M, which would account for forces that depend both on the position and velocity of the particles. Similarly, the time dependence is obtained via the projection $\mathrm{pr}_2 : \mathbb{R} \times M \to M$. The corresponding generalizations are straightforward and one arrives to Newton's equations in the natural form $\nabla_t p_\gamma = \mathcal{F}(t, \gamma'(t))$ or $\nabla_t \gamma' = F(t, \gamma'(t))$.

3.2 Lagrangian Formalism

In the Lagrangian formalism, the motion is described by means of the *configuration space* $\mathrm{T}Q$, which is the tangent bundle of the configuration manifold Q, which accounts for the region where the motion of a system takes place. One assumes that the dynamics of the original system are characterised by a function on the tangent bundle, the Lagrangian of the system, from which the equations of motion are derived. The equations determine a path on the tangent bundle which can be projected onto Q to give a trajectory of the original system.

We begin by investigating the equations of motion of a Newtonian system whose trajectories are constrained to lie inside a submanifold S. The way in which the equations are obtained leads us to consider more general systems that include conservative Newtonian systems as a particular case. The generalization allows solving some classical problems in mechanics.

3.2.1 Constrained Systems

In mechanics, we often deal with situations where the motion of a system is subject to certain constraints. An example of such systems would be a rigid pendulum, where a particle is kept at a constant distance from a point, the pivot of the pendulum. The motion of the particle is thus restricted to move on a 2-sphere centered at the pivot.

In general, let us consider a Newtonian system $(M, g, \mathcal{F})$ whose motion takes place inside a submanifold $S \subset M$. For this to happen, there must be an additional force $\mathcal{R}$ that obliges the system to stay in S. The problem of finding the constraint force $\mathcal{R}$ and solving Newton's equations is hard and impractical, but a reasonable hypothesis on the nature of the constraint will show that the motion is more easily obtained.

Definition 3.2.1. A *Newtonian dynamical system with holonomic constraints* is given by

- A Newtonian dynamical system $(M, g, \mathcal{F})$.

- A submanifold $S \overset{j}{\hookrightarrow} M$, the *constraint submanifold*.
- The possibility to associate to each path $\gamma : I \to M$ contained in S a 1-form $\mathcal{R} \in \Omega^1(\gamma)$ along γ: the *force of constraint*.

Constraints are said to be *holonomic* when they put restrictions on the configuration space. In a more general case, the constraints could be defined in the tangent bundle of the manifold, or even depend on time. An example of a nonholonomic system would be that of a rigid ball, which we imagine as being composed of many particles, rolling on a surface. If the ball rolls without slipping, then the point of the ball that is touching the surface must have a null velocity, which is a constraint on the velocity of the system. Throughout the text, we work only with holonomic constraints, which are much easier to describe.

As before, the force can be seen as a vector field along γ, $R \in \mathfrak{X}(\gamma)$, with $R = g^\sharp \circ \mathcal{R}$. The total force acting upon the system is $\mathcal{F} \circ \gamma + \mathcal{R}$. If γ is contained in S, we may write $\gamma = j \circ \gamma_\circ$ where $\gamma_\circ$ is a path in S.

The constraint force $\mathcal{R}$ is said to be *ideal* when it has the following property:

D'Alembert's Principle. The constraint force $\mathcal{R}$ is orthogonal to the constraint manifold S.

Here, orthogonal means with respect to the metric g on M. One way to express this fact is by the relation $j^*(\mathcal{R}) = 0$, where $j^*(\mathcal{R})(t) = {}^t\mathrm{T}_{\gamma_\circ(t)} j \cdot \mathcal{R}(t)$ and ${}^t\mathrm{T}_x j : \mathrm{T}^*_{j(x)} M \to \mathrm{T}^*_x S$ is the transpose of $\mathrm{T}_x j$. Indeed, one simply observes that $j^*\mathcal{R}(t) = g(R(t), \cdot)$ on $\mathrm{T}_{\gamma_\circ(t)} S$. For γ inside S its velocity γ' will be orthogonal to R, so the work of the force $\mathcal{R}$ along any path in S is zero. It turns out that these conditions are the ones that simplify the problem of motion on S.

We define the dynamics of a constrained system by assuming the validity of the principle of d'Alembert. Thus, the motion of a system with holonomic constraints is determined by the equation

$$\nabla_t p_\gamma = \mathcal{F} \circ \gamma + \mathcal{R} \tag{3.22}$$

where $\mathcal{R}$ is ideal and the image of γ is in $S \subset M$.

We can decompose the tangent space of M at x as $\mathrm{T}_x M = \mathrm{T}_x S \oplus (\mathrm{T}_x S)^\perp$. Globally, we have the orthogonal projections $\mathrm{pr}_{\mathrm{T}S} : \mathrm{T}M \to \mathrm{T}S$ and $\mathrm{pr}_{(\mathrm{T}S)^\perp} : \mathrm{T}M \to (\mathrm{T}S)^\perp$, which are defined in each fiber in the obvious way.

Recall that the map $S \overset{j}{\hookrightarrow} M$ induces a metric g^S on S by means of the pullback $g^S = j^*(g)$. We know from Riemannian geometry that the Levi-Civita connection associated with the induced metric ∇^S is simply the projection of the Levi-Civita connection on the bigger manifold onto S. Thus, for fields $X, Y \in \mathfrak{X}(S)$ we have $\nabla^S_X Y = \mathrm{pr}_{\mathrm{T}S} \circ \nabla_{\mathrm{T}j \circ X}(\mathrm{T}j \circ Y)$.

Let F^S be the projection of the force field F on the tangent bundle of the constraint submanifold, $F^S = \mathrm{pr}_{\mathrm{T}S} \circ F \circ j$ and let $\mathcal{F}^S = j^*(g)^\flat F^S$. If $X \in \mathfrak{X}(S)$ we have

$$\begin{aligned}
\langle j^*(g)^\flat F^S, X \rangle &= j^*(g)(F^S, X) \\
&= (g \circ j)(\mathrm{T}j \circ \mathrm{pr}_{\mathrm{T}S} \circ F \circ j, \mathrm{T}j \circ X) \\
&= (g \circ j)(F \circ j, \mathrm{T}j \circ X) \\
&= \langle \mathcal{F} \circ j, \mathrm{T}j \circ X \rangle \\
&= \langle j^*(\mathcal{F}), X \rangle
\end{aligned}$$

where in the third equality we have used that $F = \mathrm{T}j \circ \mathrm{pr}_{\mathrm{T}S} \circ F + \mathrm{pr}_{(\mathrm{T}S)^\perp} \circ F$ and we conclude that

$$\mathcal{F}^S = j^*(\mathcal{F}). \tag{3.23}$$

We now investigate the relationship between the original system $(M, g, \mathcal{F})$ subject to the constraint force $\mathcal{R}$ and the system $(S, g^S, \mathcal{F}^S)$, which can be thought of as a system induced by the constraints. Let γ be a path $\gamma : I \to M$ whose image lies in S and write $\gamma = j \circ \eta$ where $\eta : I \to S$. If γ is a trajectory of the system,

$$\nabla_t \gamma' = F \circ \gamma + R \tag{3.24}$$

whence it follows that

$$\mathrm{pr}_{\mathrm{T}S} \circ \nabla_t \gamma' = \mathrm{pr}_{\mathrm{T}S} \circ F \circ \gamma, \quad \mathrm{pr}_{(\mathrm{T}S)^\perp} \circ \nabla_t \gamma' = \mathrm{pr}_{(\mathrm{T}S)^\perp} \circ F \circ \gamma + R \tag{3.25}$$

and thus

$$\nabla_t^S \eta' = \mathrm{pr}_{\mathrm{T}S} \circ \nabla_t (\mathrm{T}j \circ \eta') = \mathrm{pr}_{\mathrm{T}S} \circ \nabla_t (j \circ \eta)' = \mathrm{pr}_{\mathrm{T}S} \circ \nabla_t \gamma' = \mathrm{pr}_{\mathrm{T}S} \circ F \circ \gamma = \mathrm{pr}_{\mathrm{T}S} \circ F \circ j \circ \eta = F^S \circ \eta,$$

which shows that the dynamics of the constrained system coincide with the dynamics of the induced system. The constraint force is the normal component of $\nabla_t \gamma' - F \circ \gamma$ with respect to S.

Note that if the original system is conservative with $\mathcal{F} = -dV$ then the induced system is also conservative, since $\mathcal{F}^S = j^*(\mathcal{F}) = j^*(-dV) = -dj^*(V)$ and $j^*(V)$ is the induced potential energy.

Let $(M, g, \mathcal{F})$ be a Newtonian system without constraints. We introduce coordinates (q^i) on M, and the corresponding natural coordinates $(q^i; v^i)$ on $\mathrm{T}M$. Writing $g = g_{ij} dq^i \otimes dq^j$, the kinetic energy is expressed as $\widehat{T}(q; v) = \frac{1}{2} g_{ij}(q) v^i v^j$ whence we have $\frac{\partial \widehat{T}}{\partial v^i}(q; v) = g_{ij} v^j$.

We carry out a few calculations. Let us consider a path γ in M, which we express in coordinates as $(q^i) = (q^i(t))$, then

$$\frac{d}{dt}\left(\frac{\partial T}{\partial v^i} \circ \gamma'\right)(t) = \frac{\partial g_{ij}}{\partial q^k}(q) \dot{q}^j \dot{q}^k + g_{ij}(q) \ddot{q}^j \tag{3.26}$$

$$\frac{\partial T}{\partial q^i} \circ \gamma'(t) = \frac{1}{2} \frac{\partial g_{jk}}{\partial q^i}(q) \dot{q}^j \dot{q}^k \tag{3.27}$$

So that

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\partial T}{\partial v^i} \circ \gamma'\right) - \frac{\partial T}{\partial q^i} \circ \gamma' &= g_{ij} \ddot{q}^j + \left(\frac{\partial g_{ij}}{\partial q^k} - \frac{1}{2} \frac{\partial g_{jk}}{\partial q^i}\right) \dot{q}^j \dot{q}^k \\
&= g_{il} \ddot{q}^l + \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial q^k} + \frac{\partial g_{ik}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^i}\right) \dot{q}^j \dot{q}^k \\
&= g_{il}\left(\ddot{q}^l + \Gamma^l_{jk} \dot{q}^j \dot{q}^k\right)
\end{aligned}$$

and from the formula we see that

$$\frac{d}{dt}\left(\frac{\partial T}{\partial v^i} \circ \gamma'\right) - \frac{\partial T}{\partial q^i} \circ \gamma' = \langle g^\flat \circ \nabla_t \gamma', \frac{\partial}{\partial q^i} \rangle \tag{3.28}$$

Hence, the trajectories of the system satisfy the relation

$$\frac{d}{dt}\left(\frac{\partial T}{\partial v^i} \circ \gamma'\right) - \frac{\partial T}{\partial q^i} \circ \gamma' = \langle \mathcal{F}, \frac{\partial}{\partial q^i} \rangle \tag{3.29}$$

and conversely, the paths satisfying 3.29 are also trajectories of the system. It is worth noting that the term on the right involves a geometric object, the 1-form $\mathcal{F}$, which is coordinate independent. The coordinate expression that appears on the left hand side comes from contracting the force with a coordinate vector. The equation has the same form in every coordinate system.

If the system happens to be conservative with $\mathcal{F} = -dV$, then

$$\langle \mathcal{F}, \frac{\partial}{\partial q^i} \rangle = -\langle dV, \frac{\partial}{\partial q^i} \rangle = -\frac{\partial V}{\partial q^i}. \tag{3.30}$$

Considering the function $V \circ \tau_M$, which does not depend on the velocities, and taking into account the observation just made, we write the expression as

$$\frac{d}{dt}\left(\frac{\partial (T - V \circ \tau_M)}{\partial v^i} \circ \gamma' \right) - \frac{\partial (T - V \circ \tau_M)}{\partial q^i} \circ \gamma' = 0$$

or simply as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v^i} \circ \gamma' \right) - \frac{\partial L}{\partial q^i} \circ \gamma' = 0, \quad (i = 1, \ldots, m) \tag{3.31}$$

where $L = T - V$ is the *Lagrangian* of the mechanical system and m is the dimension of M. This way, we have found that Newton's equations can be derived from a scalar function $L \in \mathcal{F}(\mathrm{T}M)$ defined in the tangent bundle of the configuration space. Equations 3.31 are known as *Lagrange's equations*.

Since this derivation is valid for general differentiable manifolds, the dynamics of a constrained system are found in the same way. If $(S, g^S, \mathcal{F}^S)$ is the system induced by the constraints, we consider the induced kinetic energy $T^S = \mathrm{T}j^*(T)$ and the induced potential energy $V^S = j^*(V)$, if the original system is conservative. The trajectories in S are determined by Lagrange's equations for the Lagrangian $L^S = T^S - V^S \circ \tau_S$. As a function on the tangent bundle, the new Lagrangian is simply the restriction of the original Lagrangian to the tangent space of the constraint manifold, $L^S = \mathrm{T}j^*(L)$.

3.2.2 Calculus of Variations

In the development above, we have found that the dynamics of a mechanical system are determined by Lagrange's equations. These equations appear also in the context of the *calculus of variations*, which gives an interesting interpretation of the behaviour of a classical system of particles. We briefly review the basic theory for this setting, where the equations arise in the most natural way.

The calculus of variations is essentially a generalization of the theory of extrema of ordinary calculus. The latter is reduced to finding stationary points of functions defined over finite-dimensional vector spaces, while the former extends these ideas to spaces of infinite dimension, giving solutions to the so called *variational problems*.

Many problems in physics, and particularly in classical mechanics, can be stated as variational problems. While it is clear that the framework for a general theory of the calculus of variations is the calculus over Banach spaces, the problem of mechanics is of such a particular and simple nature that we shall solve it with standard methods, omitting the development of the general theory. For an introduction to the subject, see [Car67].

We shall be concerned with differentiable maps $\gamma : I \to Q$ where $I = [t_1, t_2]$ is a closed interval in $\mathbb{R}$ and Q is a smooth manifold. Let $W \subseteq \mathrm{T}Q$ be an open subset. We introduce coordinates

(q^i) on Q, and the corresponding natural coordinates $(q^i; v^i)$ in $\mathrm{T}Q$. We let $L : W \to \mathbb{R}$ be a function on W and consider the integral

$$\int_I L(\gamma'(t))dt \tag{3.32}$$

which is defined when $\gamma'(t) \in W$ for all $t \in I$ and does not depend on the coordinates chosen on Q.

The above integral will have a special meaning for certain paths γ. We define Ω the set of paths on M that can be lifted to W

$$\Omega = \{\gamma : I \to Q \,|\, \gamma'(t) \in W \ \forall t \in I\}.$$

We fix two points $q_1, q_2 \in Q$ and consider the paths in Ω joining them

$$\Omega(q_1, q_2) = \{\gamma : I \to Q \,|\, \gamma'(t) \in W \ \forall t \in I, \, \gamma(t_1) = q_1, \gamma(t_2) = q_2\}.$$

Let the map $S : \Omega \to \mathbb{R}$ be defined as

$$S[\gamma] = \int_I L(\gamma'). \tag{3.33}$$

Our variational problem is concerned with finding stationary points of the restriction of S to $\Omega(q_1, q_2)$. The correct way to define a stationary point is to identify Ω with an open set in some Banach space where $\Omega(q_1, q_2)$ can be seen as a closed subspace. In these conditions, S becomes a *functional*, as in the theory of Functional Analysis. For our particular case, a *stationary point* of the functional will be a point such that the "derivative is zero" in some sense.

Let J be a neighbourhood of 0 in $\mathbb{R}$ and consider differentiable maps $\Gamma : J \times I \to Q$ such that

(i) $\Gamma(0, t) = \gamma(t)$.

(ii) $\Gamma(\epsilon, t_1) = q_1, \quad \Gamma(\epsilon, t_2) = q_2$ for all ϵ in J.

Γ is a *variation* of the path γ inside $\Omega(q_1, q_2)$, and gives rise to a vector field h along γ

$$h = \mathrm{T}\Gamma \circ \left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=0} \qquad h(t) = \mathrm{T}\Gamma \circ \frac{\partial}{\partial \epsilon}(0, t) = h^i(t) \left.\frac{\partial}{\partial q^i}\right|_{\gamma(t)}. \tag{3.34}$$

Because of the fixed boundary, we have that $h(t_1) = 0$ and $h(t_2) = 0$. We compute the derivative of $S[\Gamma]$ as a function of ϵ and evaluate at $\epsilon = 0$. In coordinates, we find

$$\begin{aligned}
\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S[\Gamma] &= \int_I \left[\frac{\partial L}{\partial q^i}(q(t), \dot{q}(t)) h^i(t) + \frac{\partial L}{\partial v^i}(q(t), \dot{q}(t)) \dot{h}^i(t)\right] dt && (3.35)\\
&= \int_I \left[\frac{\partial L}{\partial q^i}(q(t), \dot{q}(t)) - \frac{d}{dt}\Big(\frac{\partial L}{\partial v^i}(q(t), \dot{q}(t))\Big)\right] h^i(t) dt + \left[\frac{\partial L}{\partial v^i}(q(t), \dot{q}(t)) h^i(t)\right]_{t_1}^{t_2} && (3.36)\\
&= \int_I \left[\frac{\partial L}{\partial q^i}(q(t), \dot{q}(t)) - \frac{d}{dt}\Big(\frac{\partial L}{\partial v^i}(q(t), \dot{q}(t))\Big)\right] h^i(t) dt && (3.37)\\
& && (3.38)
\end{aligned}$$

where we have derived under the integral sign in the first line, integrated by parts in the second and used the boundary conditions on h in the last. We prove an easy lemma:

Lemma 3.2.2. *Let f be a continuous function defined in I. If f satisfies $\int_I f(t)h(t)dt = 0$ for every continuous function $h(t)$ with $h(t_1) = h(t_2) = 0$, then $f(t) = 0$ for all t in I.*

Proof. Suppose f is different from zero at $t = \tau \in I$. Without loss of generality, assume that $f(\tau)$ is positive (if not, we take $-f$). Then, by continuity there exists a neighbourhood U of τ in which f is positive. Consider a bump function h at $t = \tau$ with support in U. Then $\int_I f(t)h(t)dt > 0$, a contradiction. □

Imposing the derivative in 3.35 to be zero, applying the lemma we arrive at the *Euler-Lagrange equations* for the functional S:

Euler-Lagrange equations.

$$\frac{\partial L}{\partial q^i}(q(t), \dot{q}(t)) - \frac{d}{dt}\Big(\frac{\partial L}{\partial v^i}(q(t), \dot{q}(t))\Big) = 0, \quad (i = 1, \dots, m)$$

where m is the dimension of Q.

Lagrange's equations of a mechanical system are precisely the Euler-Lagrange equations associated with the corresponding Lagrangian L. By definition, a stationary point of S is a path that satisfies the Euler-Lagrange equations. Thus, the motion of a mechanical system is such that it makes the integral S stationary.

3.2.3 Lagrangian Systems

Following the previous line of thought, we define a general Lagrangian system in the natural way:

Definition 3.2.3. A *Lagrangian dynamical system* is a pair (Q, L) where Q is a smooth manifold and L is a function defined on the tangent bundle $\mathrm{T}Q$. Q is the *configuration manifold* and L is the *Lagrangian of the system.*

A very important class of Lagrangian systems are those where the Lagrangian L is of the form $L = T - V \circ \tau_Q$ for some kinetic and potential energies, as defined previously. We say that L is a *mechanical Lagrangian.*

By definition, the dynamics of a Lagrangian system are determined by *Hamilton's principle of least action*:

Hamilton's Principle of Least Action. The motion of the Lagrangian system (Q, L) is given by the stationary points of the functional

$$S[\gamma] = \int_I L(\gamma'). \tag{3.39}$$

S is the *action functional* of the system. We know that the stationary paths of the action must satisfy the Euler-Lagrange equations, which give the dynamics of the Lagrangian system. We wish to find an equivalent formulation of Hamilton's principle that only involves those geometrical objects that may be constructed on the tangent bundle of the configuration manifold. In order to do so, we make some definitions and find a new way to write the equations of motion.

Definition 3.2.4. The *Lagrangian energy* of the system is the function $E_L : \mathrm{T}Q \to \mathbb{R}$ defined as

$$E_L = \mathcal{L}_{\Delta_{\mathrm{T}Q}} L - L, \tag{3.40}$$

where $\Delta_{\mathrm{T}Q}$ is the Liouville field on $\mathrm{T}Q$.

Definition 3.2.5. The *Lagrange 1-form* is the differential form $\theta_L \in \Omega^1(TQ)$ defined as

$$\theta_L = {}^tJ \circ dL, \tag{3.41}$$

where J is the vertical endomorphism in $\mathrm{T}(\mathrm{T}Q)$, and ${}^tJ$ denotes its transpose. The differential of θ_L gives a 2-form, the *Lagrange 2-form*, which is defined by

$$\omega_L = -d\theta_L. \tag{3.42}$$

The Lagrange 2-form always defines a morphism

$$\omega_L^\flat : \mathrm{T}(\mathrm{T}Q) \to \mathrm{T}^*(\mathrm{T}Q), \quad \xi_v \mapsto i_{\xi_v}\omega_L(v) \tag{3.43}$$

Introducing coordinates (q^i) on Q, and considering the natural coordinates $(q^i; v^i)$ in the bundle $\mathrm{T}Q$, the objects we have just defined take the form:

$$E_L = v^i \frac{\partial L}{\partial v^i} - L, \quad \theta_L = \frac{\partial L}{\partial v^i} dq^i, \quad \omega_L = \frac{\partial^2 L}{\partial v^i \partial q^j} dq^i \wedge dq^j + \frac{\partial^2 L}{\partial v^i \partial v^j} dq^i \wedge dv^j \tag{3.44}$$

Given a path $\sigma : I \to \mathrm{T}Q$ written in coordinates as $\sigma(t) = (q(t); v(t))$ we have $\sigma'(t) = (q(t), v(t); \dot{q}(t), \dot{v}(t))$, and then

$$\omega_L^\flat \circ \sigma' - dE_L \circ \sigma = \left[\frac{\partial L}{\partial q^i} \circ \sigma - \frac{d}{dt}\Big(\frac{\partial L}{\partial v^i} \circ \sigma \Big) \right] dq^i \circ \sigma + (\dot{q}^i - v^i)\, d\left(\frac{\partial L}{\partial v^i} \right) \circ \sigma. \tag{3.45}$$

When $\sigma = \gamma'$ for some path γ in Q, the last term drops, and we see that γ satisfies the Euler-Lagrange equations if and only if

$$i_{\gamma''}\omega_L = dE_L \circ \gamma' \tag{3.46}$$

This way, we have rewritten the Euler-Lagrange equations in a way that only involves the Lagrange 2-form and the Lagrangian energy. Both objects have been defined via the vertical endomorphism J, which is the canonical object of the tangent bundle, and the operations of the calculus on manifolds.

Under certain conditions, the solutions of 3.46 are the integral curves of a field in $\mathrm{T}Q$. Let $X \in \mathfrak{X}(\mathrm{T}Q)$ and let σ be an integral curve of X. For σ to be the velocity of a path in Q, X must be a second order differential equation, or equivalently, we must add the condition $T(\tau_Q) \circ X = \mathrm{Id}_{\mathrm{T}Q}$.

Assume $\sigma = \gamma'$ for a path γ in Q, then

$$(i_X\omega_L) \circ \sigma = i_{(X \circ \sigma)}\omega_L = i_{\sigma'}\omega_L = i_{\gamma''}\omega_L, \quad dE_L \circ \sigma = dE_L \circ \gamma' \tag{3.47}$$

and it is clear that γ will satisfy Euler-Lagrange if and only if $i_X\omega_L = dE_L$.

The fields that satisfy equations

$$\mathrm{T}\tau_Q \circ X = \mathrm{Id}_{\mathrm{T}Q}, \quad i_X\omega_L = dE_L \tag{3.48}$$

will be called *Lagrangian fields*. Projecting the integral curves of a Lagrangian field onto Q gives the trajectories of the Lagrangian system.

We conclude this section by solving some classical problems in mechanics in the Lagrangian formalism.

Example 3.2.6. A *mechanical system with one degree of freedom* is a one-dimensional system described by a Lagrangian of the form

$$L(q,v) = \frac{1}{2}v^2 - U(q). \tag{3.49}$$

The Lagrange's equations give the relations

$$\dot{q}(t) = v(t), \quad \dot{v}(t) = -U'(q(t)) \tag{3.50}$$

which can be integrated using the conservation of the mechanical energy of the system

$$E = \frac{1}{2}v^2 + U(q). \tag{3.51}$$

Solving for $v = \dot{q}$, we obtain the separable ODE

$$\dot{q} = \sqrt{2(E - U(q))} \tag{3.52}$$

and the motion is given implicitly by

$$\int_{q_\circ}^{q} \frac{dq}{\sqrt{2(E-U(q))}} = t - t_\circ. \tag{3.53}$$

Example 3.2.7. The one dimensional *harmonic oscillator* is an example of a mechanical system with one degree of freedom. Taking $\mathbb{R}$ as the configuration manifold, the potential energy of the harmonic oscillator is

$$V(q) = \frac{1}{2}\omega^2 q^2 \tag{3.54}$$

where ω is a real number. Applying the general formula 3.53, we have

$$t - t_\circ = \int_{q_\circ}^{q} \frac{dq}{\sqrt{2E - \omega^2 q^2}} = \frac{1}{\omega}\arcsin\left(\frac{\omega q}{\sqrt{2E}}\right) + \text{constant} \tag{3.55}$$

and we obtain the solution

$$q(t) = \frac{\sqrt{2E}}{\omega}\sin(\omega(t - t_\circ) + \varphi_1), \quad v(t) = \sqrt{2E}\cos(\omega(t - t_\circ) + \varphi_1). \tag{3.56}$$

We can also consider multidimensional harmonic oscillators, which are treated similarly. An interesting example is that of the two dimensional anisotropic oscillator, whose Lagrangian is

$$L(q^1, q^2, v^1, v^2) = \frac{1}{2}[(v^1)^2 + (v^2)^2] - \frac{1}{2}[\omega_1^2(q^1)^2 + \omega_2^2(q^2)^2] \tag{3.57}$$

for some real numbers ω_1, ω_2. The system can be treated as two separate subsystems, where the partial energies

$$E_1 = \frac{1}{2}(v^1)^2 + \frac{1}{2}\omega_1^2(q^1)^2, \quad E_2 = \frac{1}{2}(v^2)^2 + \frac{1}{2}\omega_2^2(q^2)^2 \tag{3.58}$$

are conserved. Thus, we arrive at a solution of the form

$$q^1(t) = \frac{\sqrt{2E_1}}{\omega_1}\sin(\omega_1 t + \varphi_1), \quad v^1(t) = \frac{\sqrt{2E_1}}{\omega_1}\omega_1\cos(\omega_1 t + \varphi_1), \tag{3.59}$$

$$q^2(t) = \frac{\sqrt{2E_2}}{\omega_2}\sin(\omega_2 t + \varphi_2), \quad v^2(t) = \frac{\sqrt{2E_2}}{\omega_2}\omega_2\cos(\omega_2 t + \varphi_2). \tag{3.60}$$

When $\omega_1 = \omega_2$, the symmetry of the problem gives other constants of the motion

$$\ell = q^1 v^2 - q^2 v^1, \quad f = v^1 v^2 + q^1 q^2 \tag{3.61}$$

and the motion of the system can be written as a function of q^1 only.

Example 3.2.8. The *two body problem* is concerned with the evolution of a closed system composed of two particles. We describe the particles as points in $\mathbb{R}^3$, where $0 \in \mathbb{R}^3$ is the origin of the physical coordinate system. The configuration manifold is $\mathbb{R}^3 \times \mathbb{R}^3$ and in the standard coordinates we write $(\mathbf{r}_1, \mathbf{r}_2)$, where $\mathbf{r}_1$ and $\mathbf{r}_2$ give the position of each particle. The configuration space is $\mathrm{T}(\mathbb{R}^3 \times \mathbb{R}^3) \cong \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$ and we write $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{v}_1, \mathbf{v}_2)$ for the points in it.

Because the system is closed, the Galilean postulates on the symmetry of spacetime are valid and the most general Lagrangian for the system is of the form

$$L(\mathbf{r}_1, \mathbf{r}_2, \mathbf{v}_1, \mathbf{v}_2) = \frac{1}{2} m_1 |\mathbf{v}_1|^2 + \frac{1}{2} m_2 |\mathbf{v}_2|^2 - U(|\mathbf{r}_1 - \mathbf{r}_2|)$$

where $|\cdot|$ is the usual norm on $\mathbb{R}^3$.

Writing $M = m_1 + m_2$ for the total mass, the change of variables

$$\mathbf{R} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{M}, \quad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \tag{3.62}$$

gives the Lagrangian in the form

$$L(\mathbf{R}, \mathbf{r}, \mathbf{V}, \mathbf{v}) = \frac{1}{2} M |\mathbf{V}|^2 + \frac{1}{2} \mu |\mathbf{v}|^2 - U(|\mathbf{r}|) \tag{3.63}$$

where μ is the *reduced mass* and is defined by $\dfrac{1}{\mu} = \dfrac{1}{m_1} + \dfrac{1}{m_2}$. $\mathbf{R}$ gives the position of the *center of mass* of the system in $\mathbb{R}^3$, and its motion is trivially integrated. Indeed, by Lagrange's equations $\dot{\mathbf{V}} = 0$ and thus

$$\mathbf{R}(t) = \mathbf{R}^0 + \frac{1}{M} \mathbf{P} t \tag{3.64}$$

where $\mathbf{P} = m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2$ is a constant, the *total linear momentum* of the system.

For the other part of the Lagrangian, the symmetry of the problem suggests the use of spherical coordinates. The second part of Lagrangian is then written as

$$L(r, \theta, \phi, v_r, v_\theta, v_\phi) = \frac{1}{2} \mu \big(v_r^2 + r^2 v_\theta^2 + r^2 \sin^2 \theta v_\phi^2 \big) - U(r) \tag{3.65}$$

and does not depend explicitly on ϕ, so the quantity $\dfrac{\partial L}{\partial v_\phi} = \mu r^2 \sin^2 \theta v_\phi$ is conserved. We can consider different spherical coordinates depending on the choice of the axis of rotation for ϕ. For the z axis, we find that

$$\ell_z = \frac{\partial L}{\partial v_\phi} = \mu (x v_y - y v_x) \tag{3.66}$$

in Cartesian coordinates. Similarly for the x and the y axis, we find the conserved quantities

$$\ell_x = \mu (y v_z - z v_y), \quad \ell_y = \mu (z v_x - x v_z). \tag{3.67}$$

We can think of these functions as the components of a vector in $\mathbb{R}^3$, the *total angular momentum* of the system $\boldsymbol{\ell} = (\ell_x, \ell_y, \ell_z)$ which can be written as $\boldsymbol{\ell} = \mu \mathbf{r} \times \mathbf{v}$ where $\times$ now denotes the cross-product in $\mathbb{R}^3$, and we are identifying the vectors $\mathbf{r}$ and $\mathbf{v}$ as elements of the same space.

We conclude that $\mathbf{r}$ is always perpendicular to a constant vector $\boldsymbol{\ell}$ and consequently the trajectory of the system lies entirely inside a plane. Choosing the direction of $\boldsymbol{\ell}$ as the axis to construct spherical coordinates, we get polar coordinates on the plane and the Lagrangian reads

$$L(r, \phi, v_r, v_\phi) = \frac{1}{2} \mu (v_r^2 + r^2 v_\phi^2) - U(r) \tag{3.68}$$

where modulus of the angular momentum is conserved

$$\ell = \frac{\partial L}{\partial v_\phi} = \mu r^2 v_\phi. \tag{3.69}$$

In order to integrate the equations of motion, we observe that the total energy of the subsystem is also conserved, so we write

$$\frac{1}{2}\mu\Big(v_r^2 + \frac{\ell^2}{\mu^2 r^2}\Big) + U(r) = E \tag{3.70}$$

and solving for v_r, we find

$$\dot{r} = \sqrt{\frac{2}{\mu}(E - U(r)) - \frac{\ell^2}{\mu r^2}} \tag{3.71}$$

which can be integrated

$$\int_{r_\circ}^{r} \frac{dr}{\sqrt{\frac{2}{\mu}(E - U(r)) - \frac{\ell^2}{\mu^2 r^2}}} = t - t_\circ. \tag{3.72}$$

The motion of ϕ can be found by substituting r as a function of time in $\phi = \dfrac{\ell}{\mu r^2}$ and then integrating

$$\phi - \phi_\circ = \frac{\ell}{\mu}\int_{t_\circ}^{t} \frac{dt}{r(t)^2} \tag{3.73}$$

it is more interesting however to find the shape of the path of motion, without any mention to time. If we can write $r = r(\phi)$ in some region, then

$$\frac{dr}{d\phi} = \frac{dr}{dt}\frac{dt}{d\phi} = \frac{dr/dt}{d\phi/dt} = \frac{\dot{r}}{\dot{\phi}} \tag{3.74}$$

which gives

$$\phi - \phi_\circ = \ell \int_{r_\circ}^{r} \frac{dr}{r^2\sqrt{2\mu(E - U(r)) - \ell^2/r^2}}. \tag{3.75}$$

Example 3.2.9. An important particular case of the two body problem is that of a potential which is inversely proportional to the distance

$$U(r) = -\frac{\kappa}{r} \tag{3.76}$$

and is historically known as *Kepler's problem.* The force derived from the potential is in the radial direction and is inversely proportional to the distance to the origin. It is found for instance in Newton's theory of gravity.

The trajectories in Kepler's problem can be explicitly obtained. By integrating the equation of the trajectory with the change of variables $u = \dfrac{1}{r}$, one arrives at:

$$\phi = \arccos\left(\frac{(\ell/r) - (\mu\kappa/\ell)}{\sqrt{2\mu E + \mu^2\kappa^2/\ell^2}}\right) + \text{constant}. \tag{3.77}$$

If we take the origin of ϕ so that the constant is zero, we arrive at the formula

$$r = \frac{p}{1 + e\cos\phi} \tag{3.78}$$

where we have written the parameters

$$p = \frac{\ell^2}{\mu\kappa}, \quad e = \sqrt{1 + \frac{2E\ell^2}{\mu\kappa^2}}. \tag{3.79}$$

This can be seen to be the general formula of a conic section with one focus at the origin. The *eccentricity* e determines the type of conic, which can be a circle ($e = 0$), an ellipse ($e < 1$), a parabola ($e = 1$) or a hyperbola ($e > 1$).

3.3 Hamiltonian Formalism

As discussed earlier, Lagrangian mechanics is entirely done in $\mathrm{T}Q$. While we are interested in the physical trajectories on Q, it is more convenient to describe this motion in a bigger manifold where the paths do not cross. One such manifold is the tangent bundle, which we call the *carrier manifold* of the motion, but we can in principle find other examples with the same property.

In the Hamiltonian formalism, the carrier manifold is the cotangent bundle or *phase space* T^*Q. The dynamics of the system are contained in a function in the cotangent bundle, the *Hamiltonian* of the system. The Hamiltonian formulation of mechanics has several advantages over the Lagrangian. In the latter, the equations of motion are in general buried inside the Euler-Lagrange equations, which are a second order system of differential equations, while the former gives a first order system that possesses a remarkable symmetry. The symmetry comes from the canonical symplectic structure of the cotangent bundle. Thus, we are led to develop a formalism for symplectic manifolds in general, which includes Lagrangian systems as a particular case.

3.3.1 Hamiltonian Systems

We give the most general definition of a Hamiltonian system, in order to apply it to the problem of mechanics.

Definition 3.3.1. A Hamiltonian system is a triple (M, ω, H) where M is a smooth manifold, ω is a symplectic form on M and $H : M \to \mathbb{R}$ is a function on M, known as the *Hamiltonian* of the system.

The dynamical information of the system is contained in the Hamiltonian H. We use the symplectic structure to define a vector field Z_H on M, whose integral curves are the trajectories of the system. By definition, Z_H is the symplectic gradient of dH

$$Z_H = \omega^\sharp \circ dH \tag{3.80}$$

or equivalently, it is the only vector field such that

$$i_{Z_H}\omega = dH. \tag{3.81}$$

Thus, the motion is given by the paths $\xi : I \to M$ such that

$$\xi' = Z_H \circ \xi, \tag{3.82}$$

which is known as *Hamilton's equation.*

Introducing Darboux coordinates (q^i, p_i) in M, if $\hat{\xi}(t) = (q^i(t), p_i(t))$ is a trajectory of the system, it satisfies Hamilton's equation, which takes the form

$$\frac{dq^i}{dt} = \quad \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}. \tag{3.83}$$

These particularly beautiful relations are known as *Hamilton's canonical equations*.

Let $f : M \to \mathbb{R}$ be a function on M, and assume $\xi : I \to M$ to be a trajectory of the Hamiltonian system (M, ω, H). According to the previous mathematical development,

$$\frac{d(f \circ \xi)}{dt} = \mathcal{L}_{\xi'} f = \mathcal{L}_{Z_H \circ \xi} f = \{f, H\} \circ \xi \tag{3.84}$$

and we can write Hamilton's equations as

$$\dot{q}^i = \{q^i, H\}, \quad \dot{p}_i = \{p_i, H\}. \tag{3.85}$$

We say that the function f is a *constant of motion* or that f is *conserved* when it is constant along the trajectories of the system. Equivalently, f is conserved if and only if $\{f, H\} = 0$. In particular, H is always a constant of the motion.

Two conserved quantities always give a third constant of the motion by means of the Poisson bracket. Indeed, let f, g be constants of the motion with respect to the Hamiltonian H. From the Jacobi identity, it follows that

$$\{\{f, g\}, H\} = \{f, \{g, H\}\} + \{\{f, H\}, g\} = 0 \tag{3.86}$$

and $\{f, g\}$ is also conserved.

In some cases, it is possible to produce new constants of the motion in this manner, as in the next example:

Example 3.3.2. Consider three dimensional Euclidean space $\mathbb{R}^3$, and its tangent bundle $\mathrm{T}\mathbb{R}^3 \cong \mathbb{R}^{\nvDash} \times \mathbb{R}^3$ with coordinates (x, y, z, v_x, v_y, v_z). Recall that the functions

$$\ell_x = yv_z - zv_y, \quad \ell_y = zv_x - xv_z, \quad \ell_z = xv_y - yv_x \tag{3.87}$$

are the components of the angular momentum $\boldsymbol{\ell} = (\ell_x, \ell_y, \ell_z)$ in $\mathbb{R}^3$. It is straightforward to show that

$$\{\ell_x, \ell_y\} = \ell_z, \quad \{\ell_y, \ell_z\} = \ell_x, \quad \{\ell_z, \ell_x\} = \ell_y \tag{3.88}$$

whence we see that two conserved components of the angular momentum imply the conservation of the third.

3.3.2 Legendre Transform and Hamiltonian Mechanics

Let (Q, L) be a Lagrangian dynamical system.

Definition 3.3.3. The *Legendre map* of L is the fiber derivative of L, $\mathcal{F}L : \mathrm{T}Q \to \mathrm{T}^*Q$. In the natural coordinates, the map is written $(q^i; v^i) \mapsto (q^i; \dfrac{\partial L}{\partial q^i})$.

As already noted in the previous chapter, the Legendre map does not have any distinctive property, but under certain conditions it can be a local diffeomorphism. More precisely, we have:

Proposition 3.3.4. *The following are equivalent:*

(i) $\mathcal{F}L$ is a local diffeomorphism.

(ii) For every $v \in TQ$, $\mathcal{F}^2L(v)$ is nondegenerate.

(iii) The Lagrange 2-form ω_L is nondegenerate.

Proof. Let us write $W_{ij} = \dfrac{\partial^2 L}{\partial v^i \partial v^j}$, which defines a matrix W. (i) and (ii) are equivalent by the inverse function theorem, since $\det \mathrm{T}\mathscr{F}L(v) = \det W(v)$ for $v \in \mathrm{T}Q$, as can easily be seen. It is also easy to see that $\det \omega^\flat_{Lv} = \det W(v)$ for $v \in \mathrm{T}Q$, which gives the equivalence of (ii) and (iii). □

A Lagrangian is said to be *regular* when any of the previous conditions holds. When the Lagrangian is regular, by nondegeneracy of ω_L there exists a unique $X_L \in \mathfrak{X}(\mathrm{T}Q)$ such that

$$i_{X_L}\omega_L = dE_L. \tag{3.89}$$

Furthermore, X_L satisfies the second order condition and it is easy to find its expression in coordinates

$$X_L = v^i \frac{\partial}{\partial q^i} + W^{ik}\left(\frac{\partial L}{\partial q^k} - \frac{\partial^2 L}{\partial v^k \partial q^j} v^j\right)\frac{\partial}{\partial v^i} \tag{3.90}$$

where W^{ij} are the coefficients of the matrix W^{-1}. Thus, a regular Lagrangian gives rise to a Lagrangian field X_L.

We use the Legendre map to investigate the relation between the geometrical objects in the Lagrangian formalism and the ones in the cotangent bundle. Let θ and ω be the canonical forms in T^*Q. We observe that

$$\theta_L = \mathcal{F}L^*(\theta) \quad \omega_L = \mathcal{F}L^*(\omega) \tag{3.91}$$

which are trivially obtained in coordinates by substitution of $\dfrac{\partial L}{\partial v^i}$ in p_i. Thus, the Lagrange forms are actually the canonical objects of the cotangent bundle in $\mathrm{T}Q$.

If L is regular, let $v \in \mathrm{T}_u\mathrm{T}Q$ such that $\omega_L(v, w) = 0$ for all $w \in \mathrm{T}_u\mathrm{T}Q$, then

$$\omega_{Lu}(v, w) = \omega_{\mathcal{F}L(u)}(\mathrm{T}_u\mathcal{F}L \cdot v, \mathrm{T}_u\mathcal{F}L \cdot w) = 0$$

but $\mathrm{T}_u\mathcal{F}L$ is an isomorphism, so we have $\omega_{\mathcal{F}L(u)}(\mathrm{T}_u\mathcal{F}L \cdot v, \xi) = 0$ for all $\xi \in \mathrm{T}_{\mathcal{F}L(u)}\mathrm{T}^*Q$ which by nondegeneracy of ω implies $\mathrm{T}_u\mathcal{F}L \cdot v = 0$, and by injectivity of $\mathrm{T}_u\mathcal{F}L$ we get $v = 0$, which shows that ω_L is nondegenerate.

Thus, a regular Lagrangian induces a symplectic form ω_L in $\mathrm{T}Q$, and $(\mathrm{T}Q, \omega_L)$ is a symplectic manifold. Lagrange's equations are written

$$X_L = \omega_L^\sharp \circ dE_L \tag{3.92}$$

which means that X_L is a Hamiltonian field with Hamiltonian E_L, via the symplectic form ω_L. In other words, $(\mathrm{T}Q, \omega_L, E_L)$ is a Hamiltonian dynamical system.

In order to translate the dynamics in configuration space to equivalent ones in phase space, we need a way to transform the Lagrangian energy E_L into a global function in T^*Q. If the Lagrangian is regular, we can pushforward the energy onto T^*Q locally, via the Legendre map.

We say that L is *hyperregular* when the associated Legendre map is a (global) diffeomorphism. In these conditions, $\mathcal{F}L$ is called the *Legendre transform*. We thus obtain a function $H = \mathcal{F}L_*(E_L)$ defined in T^*Q, the *Hamiltonian* of the system. We make the cotangent bundle into a Hamiltonian system by defining its trajectories as the integral curves of the field

$$Z_H = \omega^\sharp \circ dH \tag{3.93}$$

which is the symplectic gradient of dH. We thus get the Hamiltonian system $(\mathrm{T}^*Q, \omega, H)$, which is equivalent to $(\mathrm{T}Q, \omega_L, E_L)$ via the Legendre transform.

We have

$$dH = d(\mathcal{F}L_*(E_L)) = \mathcal{F}L_*(dE_L) = \mathcal{F}L_*(i_{X_L}\omega_L) = i_{\mathcal{F}L_*(X_L)}\omega$$

whence

$$Z_H = \mathcal{F}L_*(X_L), \tag{3.94}$$

which means that the fields are $\mathcal{F}L$-related (see chapter 5) $X_L \underset{\mathcal{F}L}{\sim} Z_H$. Equivalently, if $\gamma : I \to Q$ is a solution of Lagrange's equations, $\xi = \mathcal{F}L \circ \gamma'$ is a solution of Hamilton's equations. Conversely, if $\xi : I \to \mathrm{T}^*Q$ is a solution of Hamilton's equations, then $\gamma = \pi_Q \circ \xi$ is a solution of Lagrange's equations.

It is useful to remember the expression of the Hamiltonian for the most simple mechanical systems. Recall that the Lagrangian of a conservative mechanical system is of the form $L = T - V \circ \tau_Q$, where T is the kinetic energy defined by the metric and V is the potential energy, which is a function in Q. In coordinates, the Lagrangian reads

$$L(q, v) = \frac{1}{2} g_{ij}(q) v^i v^j - V(q) \tag{3.95}$$

and the Legendre map is written

$$\mathcal{F}L(q^i; v^i) = \left(q^i; \frac{\partial L}{\partial v^i}\right) = (q^i; g_{ij} v^j) \tag{3.96}$$

which is invertible by the nondegeneracy of the metric. Thus, the Legendre map is a diffeomorphism and we have that mechanical Lagrangians are always hyperregular.

By the previous remarks, $H = \mathcal{F}L_*(E_L) = (\mathcal{F}L^{-1})^*(E_L)$, and we must find the inverse of the Legendre transform $\mathcal{F}L^{-1}$. As a function of velocities, the momenta are written $p_i = g_{ij} v^j$, and the relation is easily inverted as $v^i = g^{ij} p_j$, where (g^{ij}) is the inverse matrix of (g_{ij}). Thus,

$$\mathcal{F}L_*(T) = \frac{1}{2} g_{kl} \mathcal{F}L_*(v^k) \mathcal{F}L_*(v^l) = \frac{1}{2} g_{kl} g^{ki} p_i g^{lj} p_j = \frac{1}{2} \delta^i_l g^{lj} p_i p_j = \frac{1}{2} g^{ij} p_i p_j$$

and in order to compute the Hamiltonian of the system, we must simply invert the metric and change the sign of the potential energy:

$$H(q, p) = \frac{1}{2} g^{ij} p_i p_j + V(q). \tag{3.97}$$

With this last observation in mind, we conclude the chapter by translating the previous problems into the Hamiltonian formalism:

Example 3.3.5. The Hamiltonian of the system with one degree of freedom is simply

$$H(q, p) = \frac{1}{2} p^2 + U(q), \tag{3.98}$$

and we take the canonical symplectic structure in $\mathbb{R}^2$, $\omega = dq \wedge dp$, giving Hamilton's equations

$$\dot{q}(t) = p(t), \quad \dot{p}(t) = -U'(q(t)). \tag{3.99}$$

The system is integrated as in the Lagrangian case, writing p as a function of q by conservation of energy, and solving the resulting ODE for q.

Example 3.3.6. The Hamiltonian of the two body problem is written as

$$H(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2) = \frac{1}{2m_1}|\mathbf{p}_1|^2 + \frac{1}{2m_2}|\mathbf{p}_2|^2 + U(|\mathbf{r}_1 - \mathbf{r}_2|^2). \tag{3.100}$$

We consider the canonical symplectic form ω in $\mathrm{T}^*Q \cong \mathrm{T}^*(\mathbb{R}^3 \times \mathbb{R}^3) \cong \mathbb{R}^12$. Changing the coordinates to

$$\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{M}, \quad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \tag{3.101}$$

as before, we obtain the Hamiltonian

$$H(\mathbf{R}, \mathbf{r}, \mathbf{P}, \mathbf{p}) = \frac{1}{2M}|\mathbf{P}|^2 + \frac{1}{2\mu}|\mathbf{p}|^2 + U(|\mathbf{r}|^2) \tag{3.102}$$

whence we see immediately that the total linear momentum P is conserved

$$\{\mathbf{P}, H\} = 0, \tag{3.103}$$

and the dynamics of $\mathbf{R}$ are obtained from the equation

$$\dot{\mathbf{R}} = \{\mathbf{R}, H\} = \frac{1}{M}\mathbf{P}. \tag{3.104}$$

The second part of the Hamiltonian is again written in spherical coordinates as

$$H(r, \theta, \phi, p_r, p_\theta, p_\phi) = \frac{1}{2\mu}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2\sin^2\theta}\right) + U(r) \tag{3.105}$$

so $\{p_\phi, H\} = 0$ for any spherical coordinates, and the problem can be reduced to motion on a plane as we did before. Thus, the new Hamiltonian is

$$H(r, \phi, p_r, p_\phi) = \frac{1}{2\mu}\left(p_r^2 + \frac{p_\phi^2}{r^2}\right) + U(r) \tag{3.106}$$

and the problem is solved in the same way as in the Lagrangian case.

Chapter 4

Classical Hamilton-Jacobi Theory

In the 1820s Hamilton showed that the problem of geometrical optics could be completely described by means of a scalar function, *Hamilton's characteristic function*, satisfying a first order partial differential equation [CS31]. The ideas that arose from this work led Hamilton to the discovery of a profound relationship between optics and analytical mechanics, which culminated in his celebrated formulation of the latter. As a particular consequence of the analogy, the *action function* of mechanics also satisfied a partial differential equation, which would later be known as the *Hamilton-Jacobi equation.* It was Jacobi who solved the inverse problem of producing solutions of the equations of motions from a special class of solutions of the Hamilton-Jacobi equation, which became and still is the most powerful method for integrating the motion of a dynamical system [Jac84].

In this chapter, we present the Hamilton-Jacobi equation by discussing some of the important ideas of analytical mechanics and geometrical optics that led to its classical formulation [LL77, volume 2, chapter 7], [Arn74, chapter 9], [CH37]. We then formalize some aspects of the original theory by writing the equation in a coordinate-independent manner, explaining Jacobi's method from this point of view [AM78, chapter 5] and applying it to some classical problems. We finish by making a few comments about the complete integrability of mechanical systems [Arn74, chapter 10], [AM78, chapter 5].

4.1 Action as a Function of the Coordinates

In mechanics, the Hamilton-Jacobi equation appears naturally when considering the action functional of the Lagrangian as a function of the spatial coordinates and time. In order to establish this dependence, throughout this section we work in the *extended configuration manifold* $\mathbb{R} \times Q$, which makes the time coordinate explicit as in Galilean spacetime.

Let (Q, L) be a Lagrangian dynamical system and $A = (t_1,\ q_{(1)}), B = (t_2, q_{(2)}) \in U \subset I \times Q \subset \mathbb{R} \times Q$ two points in the extended configuration manifold, where U is open. Consider the expression of the action functional

$$S(t_1, q_{(1)}; t_2, q_{(2)}) = \int_{t_1}^{t_2} L(\gamma'(s))ds \tag{4.1}$$

where the path $\gamma : I \to Q$ is a dynamical trajectory of the Lagrangian system such that

$$\gamma(t_1) = q_{(1)}, \quad \gamma(t_2) = q_{(2)}. \tag{4.2}$$

We assume that U is such that the path in 4.1 is uniquely determined by the points A, B. It can be shown [Arn74] that such a set exists for a sufficiently small $|t_2 - t_1|$, in some neighbourhood of the $q_{(1)}$, $q_{(2)}$ in Q. Thus, we obtain a function S in $U \times U \subset (\mathbb{R} \times \mathrm{T}Q) \times (\mathbb{R} \times \mathrm{T}Q)$, traditionally known as the *action function* of the Lagrangian system.

The paths are smooth functions of the boundary conditions, and we may write the path joining A and B in symplectic coordinates as

$$q^i(s) = f^i(s; t_1, q_{(1)}, t_2, q_{(2)}), \quad p_i(s) = g_i(s; t_1, q_{(1)}, t_2, q_{(2)}). \tag{4.3}$$

We are interested in computing the differential of the action function. Fix $A_\circ, B_\circ \in U$ and consider a variation of the endpoints depending on a parameter ϵ defined in a neighbourhood of zero $J \subset \mathbb{R}$, that is, we have maps $A, B : J \to U$ such that

$$A|_{s=0} = A_\circ, \quad B|_{s=0} = B_\circ. \tag{4.4}$$

The paths now depend on ϵ as

$$q^i(s) = f^i(s; A(\epsilon), B(\epsilon)), \quad p_i(s) = g_i(s; A(\epsilon), B(\epsilon)). \tag{4.5}$$

and we write

$$q^i_{(1)} = f^i(t_1(\epsilon); A(\epsilon), B(\epsilon)), \quad p^{(1)}_i = g_i(t_1(\epsilon); A(\epsilon), B(\epsilon)) \tag{4.6}$$

$$q^i_{(2)} = f^i(t_2(\epsilon); A(\epsilon), B(\epsilon)), \quad p^{(2)}_i = g_i(t_2(\epsilon); A(\epsilon), B(\epsilon)) \tag{4.7}$$

for the coordinates in $A_\circ$ and $B_\circ$.

We denote differentiation with respect to s by a dot, and with respect to ϵ by δ. It is convenient to write the action as

$$S = \int_{t_1}^{t_2} L(q(s), \dot{q}(s)) ds = \int_{t_1}^{t_2} \Big[p_i(s)\dot{q}^i(s) - H(q(s), p(s)) \Big] ds. \tag{4.8}$$

which is a function of ϵ by the above considerations.

Differentiating the action function, we have

$$\begin{aligned} \delta S = & \Big[p_i(t_2)\dot{q}^i(t_2) - H(q(t_2), p(t_2)) \Big] \delta t_2 - \Big[p_i(t_1)\dot{q}^i(t_1) - H(q(t_1), p(t_1)) \Big] \delta t_1 + \\ & + \int_{t_1}^{t_2} \Big[p_i \delta\dot{q}^i + \dot{q}^i \delta p_i - \frac{\partial H}{\partial q^i}\delta q^i - \frac{\partial H}{\partial p_i}\delta p_i \Big] ds \end{aligned}$$

but the paths are solutions of Hamilton's equations, whence

$$\begin{aligned} \delta S = & \Big[p_i(t_2)\dot{q}^i(t_2) - H(q_{(2)}, p^{(2)}) \Big] \delta t_2 - \Big[p_i(t_1)\dot{q}^i(t_1) - H(q_{(1)}, p^{(1)}) \Big] \delta t_1 + \int_{t_1}^{t_2} \Big[p_i \delta\dot{q}^i + p_i \delta q^i \Big] ds \\ = & \Big[p_i(t_2)\dot{q}^i(t_2) - H(q_{(2)}, p^{(2)}) \Big] \delta t_2 - \Big[p_i(t_1)\dot{q}^i(t_1) - H(q_{(1)}, p^{(1)}) \Big] \delta t_1 + \int_{t_1}^{t_2} (p_i \delta q^i)^{\cdot} ds \\ = & \Big[p_i(t_2)\dot{q}^i(t_2) - H(q_{(2)}, p^{(2)}) \Big] \delta t_2 - \Big[p_i(t_1)\dot{q}^i(t_1) - H(q_{(1)}, p^{(1)}) \Big] \delta t_1 + \\ & + \Big[p^{(2)}_i \delta q^i|_{s=t_2} - p^{(1)}_i \delta q^i|_{s=t_1} \Big]. \end{aligned}$$

From 4.3, it follows that

$$\delta[q^i_{(1)}] = \dot{q}^i(t_1)\delta t_1 + \delta q^i|_{s=t_1}, \quad \delta[q^i_{(2)}] = \dot{q}^i(t_2)\delta t_2 + \delta q^i|_{s=t_2}, \tag{4.9}$$

and thus

$$\delta S = p_i^{(2)}\delta q_{(2)}^i - H(q_{(2)}, p^{(2)})\delta t_2 - p_i^{(1)}\delta q_{(1)}^i + H(q_{(1)}, p^{(1)})\delta t_1. \tag{4.10}$$

Taking variations of the form

$$(t_1 + \epsilon,\, q_{(1)}^1, \dots, q_{(1)}^i, \dots, q_{(1)}^m;\, t_2,\, q_{(2)}^1, \dots, q_{(2)}^i, \dots, q_{(2)}^m) \tag{4.11}$$
$$(t_1,\, q_{(1)}^1, \dots, q_{(1)}^i + \epsilon, \dots, q_{(1)}^m;\, t_2,\, q_{(2)}^1, \dots, q_{(2)}^i, \dots, q_{(2)}^m) \tag{4.12}$$
$$(t_1,\, q_{(1)}^1, \dots, q_{(1)}^i, \dots, q_{(1)}^m;\, t_2 + \epsilon,\, q_{(2)}^1, \dots, q_{(2)}^i, \dots, q_{(2)}^m) \tag{4.13}$$
$$(t_1,\, q_{(1)}^1, \dots, q_{(1)}^i, \dots, q_{(1)}^m;\, t_2,\, q_{(2)}^1, \dots, q_{(2)}^i + \epsilon, \dots, q_{(2)}^m) \tag{4.14}$$

we obtain the differential of the action function

$$dS(t_1, q_{(1)}; t_2, q_{(2)}) = p_i^{(2)}dq_{(2)}^i - H(q_{(2)}, p^{(2)})dt_2 - p_i^{(1)}dq_{(1)}^i + H(q_{(1)}, p^{(1)})dt_1, \tag{4.15}$$

which gives the relations

$$\frac{\partial S}{\partial q_{(2)}^i} = p_i^{(2)}, \quad \frac{\partial S}{\partial t_2} = -H(q_{(2)}, p^{(2)}) \tag{4.16}$$
$$\frac{\partial S}{\partial q_{(1)}^i} = -p_i^{(1)}, \quad \frac{\partial S}{\partial t_1} = H(q_{(1)}, p^{(1)}). \tag{4.17}$$

It is common to fix the point $A = A_\circ$ and consider the paths that emanate from it. If we write (t, q) instead of $(t_2, q_{(2)})$, we obtain the expression

$$dS(t, q) = p_i dq^i - H(q, p)dt \tag{4.18}$$

where p_i is the momentum of the path joining q_1 and q at time t. By 4.16, we have

$$\frac{\partial S}{\partial q^i} = p_i, \quad \frac{\partial S}{\partial t} = -H(q, p) \tag{4.19}$$

and we obtain the celebrated *Hamilton-Jacobi equation*

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}\right) = 0. \tag{4.20}$$

The action function can be explicitly calculated in the most simple mechanical systems:

Example 4.1.1. Consider the one dimensional free particle, whose Lagrangian is

$$L(q, v) = \frac{1}{2}v^2.$$

Fixing two points $(t_1, q_{(1)}), (t_2, q_{(2)}) \in \mathbb{R} \times \mathbb{R}$, there exists a unique dynamical trajectory $q(t)$ such that

$$q(t_1) = q_{(1)}, \quad q(t_2) = q_{(2)},$$

given by

$$q(s; t_1, q_{(1)}, t_2, q_{(2)}) = q_{(1)} + \frac{q_{(2)} - q_{(1)}}{t_2 - t_1}s. \tag{4.21}$$

The action function is trivially obtained

$$S(t_1, q_{(1)}; t_2, q_{(2)}) = \int_{t_1}^{t_2} \frac{1}{2}\left(\frac{q_{(2)} - q_{(1)}}{t_2 - t_1}\right)^2 ds = \frac{1}{2}\frac{(q_{(2)} - q_{(1)})^2}{t_2 - t_1} \tag{4.22}$$

and gives

$$\frac{\partial S}{\partial q_{(2)}} = \frac{q_{(2)} - q_{(1)}}{t_2 - t_1} = p^{(2)}, \quad \frac{\partial S}{\partial t_2} = -\frac{1}{2}\left(\frac{q_{(2)} - q_{(1)}}{t_2 - t_1}\right)^2 = -H(q_{(2)}, p^{(2)}) \tag{4.23}$$

$$\frac{\partial S}{\partial q_{(1)}} = -\frac{q_{(2)} - q_{(1)}}{t_2 - t_1} = -p^{(1)}, \quad \frac{\partial S}{\partial t_1} = \frac{1}{2}\left(\frac{q_{(2)} - q_{(1)}}{t_2 - t_1}\right)^2 = H(q_{(1)}, p^{(1)}). \tag{4.24}$$

as expected.

Example 4.1.2. Consider the one dimensional harmonic oscillator with Lagrangian

$$L(q, v) = \frac{1}{2}v^2 - \frac{1}{2}\omega^2 q^2.$$

Fixing two points $(t_1, q_{(1)}), (t_2, q_{(2)}) \in \mathbb{R} \times \mathbb{R}$, the solutions of the equations of motion

$$q(t) = A\cos(\omega t) + B\sin(\omega t) \tag{4.25}$$

joining the two points must satisfy

$$q_{(1)} = A\cos(\omega t_1) + B\sin(\omega t_1), \quad q_{(2)} = A\cos(\omega t_2) + B\sin(\omega t_2) \tag{4.26}$$

which determine a unique path if and only if

$$\begin{vmatrix} \cos(\omega t_1) & \sin(\omega t_1) \\ \cos(\omega t_2) & \sin(\omega t_2) \end{vmatrix} = \sin(\omega(t_2 - t_1)) \neq 0 \tag{4.27}$$

which is equivalent to $t_2 - t_1 \neq \dfrac{\pi}{\omega}n$ for $n \in \mathbb{Z}$.

Under the last assumption, let us write the path joining the points and the corresponding velocity in the form

$$q(s) = q_{(1)}\cos(\omega(s - t_1)) + \frac{v_1}{\omega}\sin(\omega(s - t_1)) \tag{4.28}$$

$$v(s) = v_1\cos(\omega(s - t_1)) - q_{(1)}\omega\sin(\omega(s - t_1)). \tag{4.29}$$

where v_1 is an appropriate function of the endpoints.

We observe that

$$v(s)^2 - \omega^2 q(s)^2 = -2q_{(1)}v_1\omega\sin(2\omega(s - t_1)) + (\frac{v_1^2}{\omega^2} - q_{(1)}^2)\omega^2\cos(2\omega(s - t_1))$$

and a primitive function of $v(s)^2 - \omega^2 q(s)^2$ is

$$q_{(1)}v_1\cos(2\omega(s - t_1)) + \frac{1}{2}(\frac{v_1^2}{\omega^2} - q_{(1)}^2)\omega\sin(2\omega(s - t_1))$$

which can be written as

$$q_{(1)}v_1\left(1 - 2\sin^2(\omega(s - t_1))\right) + \omega\left(\frac{v_1^2}{\omega^2} - q_{(1)}^2\right)\sin(\omega(s - t_1))\cos(\omega(s - t_1))$$

whence

$$S(t_1, q_{(1)}; t_2, q_{(2)}) = -q_{(1)}v_1\sin^2(\omega(t_2 - t_1)) + \frac{\omega}{2}\left(\frac{v_1^2}{\omega^2} - q_{(1)}^2\right)\sin(\omega(s - t_1))\cos(\omega(t_2 - t_1))$$

and for v_1 to give the desired solution with $q(t_2) = q_{(2)}$, we must have

$$v_1 = \omega \frac{q_{(2)} - q_{(1)} \cos(\omega(t_2 - t_1))}{\sin(\omega(t_2 - t_1))}.$$

After some algebra, we arrive at the expression

$$S(t_1, q_{(1)}; t_2, q_{(2)}) = \frac{\omega}{2} \frac{(q_{(1)}^2 + q_{(2)}^2) \cos(\omega(t_2 - t_1)) - 2q_{(1)}q_{(2)}}{\sin(\omega(t_2 - t_1))} \tag{4.30}$$

for the action function, and we find

$$\frac{\partial S}{\partial q_{(2)}} = \frac{\omega(q_{(2)} \cos \omega(t_2 - t_1) - q_{(1)})}{\sin(\omega(t_2 - t_1))} = p^{(2)}, \tag{4.31}$$

$$\frac{\partial S}{\partial t_2} = -\frac{1}{2} \frac{\omega^2(q_{(1)}^2 + q_{(2)}^2 - 2q_{(1)}q_{(2)} \cos(\omega(t_2 - t_1)))}{\sin^2 \omega(t_2 - t_1)} = -H(q_{(2)}, p^{(2)}), \tag{4.32}$$

$$\frac{\partial S}{\partial q_{(1)}} = \frac{\omega(q_{(1)} \cos \omega(t_2 - t_1) - q_{(2)})}{\sin(\omega(t_2 - t_1))} = -p^{(1)}, \tag{4.33}$$

$$\frac{\partial S}{\partial t_1} = \frac{1}{2} \frac{\omega^2(q_{(1)}^2 + q_{(2)}^2 - 2q_{(1)}q_{(2)} \cos(\omega(t_2 - t_1)))}{\sin^2(\omega(t_2 - t_1))} = H(q_{(1)}, p^{(1)}). \tag{4.34}$$

4.2 Hamiltonian Optics

Optics studies the properties of light, the physical nature of which has fascinated and mystified scientists for centuries. Throughout history, experiment has shown that light behaves in two apparently irreconcilable ways. On the one hand, the properties of light are seen to be similar to those of mechanical particles in that, for example, light can be seen to take preferred paths in space to travel from one point to another. On the other hand, light behaves as a perturbation that propagates in space or *wave*, which gives rise to various phenomena that are not typical of a particle, such as interference and diffraction.

The experimental facts led to the development of several theories that tried to explain each behaviour separately. Thus, the particle properties could be deduced from a variational principle of Fermat that gave the trajectories of light particles or *rays*, while light waves found a good description in the *wavefront* approach of Huygens. In his papers in optics [CS31], Hamilton was the first to establish a deep connection between both theories.

We start by investigating the theory of rays. *Geometrical optics* is the study of light without any mention to its wave properties, which still exist but may be ignored under certain assumptions [LL77, volume 2, chapter 7]. In this context, light is described by *rays*, which are the paths travelled by particles of light. We see rays as parametrized curves $\gamma : I \to Q$, where I is a real interval and Q a manifold (the configuration manifold of mechanics).

The behaviour of light in a physical medium is characterised by a function n, the *index of refraction*, which in general depends on the position and orientation of the ray. For simplicity, we assume that the medium is *isotropic*, so that n depends on the position only, and is thus a function in Q. Assume also that Q is Riemannian, so that a metric g is defined and let us write $\|\cdot\|$ for the induced norm. The rays are determined by *Fermat's principle*:

Fermat's principle of least time. Given $q_{(1)}$, $q_{(2)}$ in Q, the light ray γ travelling from $q_{(1)}$ to $q_{(2)}$ is such that it makes the functional

$$S[\gamma] = \int_{t_1}^{t_2} n(\gamma(s))\|\gamma'(s)\|ds$$

stationary under variations fixing the points in the boundary $\gamma(t_1) = q_{(1)}$, $\gamma(t_2) = q_{(2)}$.

Physically, the index of refraction n is the inverse of the speed of light in the medium, and $\|\gamma'(s)\|ds$ is the line element of the curve γ, so the integral is actually a measure of the time needed by light to travel from $q_{(1)}$ to $q_{(2)}$ along its path, and does not depend on the parametrization of the ray. This quantity is known as the *optical length* of the ray, and the problem of ray optics is equivalent to the variational problem of finding the stationary paths of the optical length.

Fermat's principle of least time becomes exactly Hamilton's principle of least action, if we define the Lagrangian function

$$L(q, v) = n(q)\|v\|$$

and consider (Q, L) as a Lagrangian dynamical system. Thus, rays satisfy the Euler-Lagrange equations, which are found to be equivalent to the relation

$$\nabla_t((n \circ \gamma) \cdot \gamma') = \text{grad n} \circ \gamma. \tag{4.35}$$

By analogy with mechanics, we define the *optical momenta* from the Lagrangian as the functions $p_i = \dfrac{\partial L}{\partial q^i} = n^2 \dfrac{g_{ij} v^j}{L}$, but we observe that we do not have a Legendre map with good properties. Indeed, the fact that L is homogeneous of degree one in the velocities implies that $\left| \dfrac{\partial^2 L}{\partial v^i \partial v^j} \right| = 0$ and the Lagrangian energy is identically zero

$$E_L = v^i \frac{\partial L}{\partial v^i} - L = 0 \tag{4.36}$$

whence it follows that the Hamiltonian is also zero.

Regardless of this fact, we can write the optical length as a function $S(q_{(1)}, q_{(2)})$ of the endpoints of the ray, just like we did with the mechanical action in the previous section. Fixing $q_{(1)}$, we find that the relations derived earlier still hold for this particular case, in the form

$$\frac{\partial S}{\partial t} = -H(q, p) = 0, \quad \frac{\partial S}{\partial q^i} = p_i = \frac{\partial L}{\partial q^i}(q, v) \tag{4.37}$$

where we have written t, q, v, p instead of $t_2, q_{(2)}, v_{(2)}, p^{(2)}$. As a function of the coordinates, $S(q)$ is known as *Hamilton's characteristic function*, or more commonly as the *eikonal* of the system of rays. The main idea is to produce the momenta of the trajectories of light from the partial derivatives of the eikonal.

We use Euler's homogeneous function theorem to obtain the relation

$$v^i \frac{\partial L}{\partial v^i} = L \tag{4.38}$$

which will substitute the Hamilton-Jacobi equation of mechanics. If we substitute $\dfrac{\partial L}{\partial v^i}$ by $\dfrac{\partial S}{\partial q^i}$ in 4.38, and divide both sides by L we obtain from the expression of p_i:

$$1 = \frac{v^i}{L} \frac{\partial S}{\partial q^i} = g^{ij} \frac{\partial S}{\partial q^i} \frac{\partial S}{\partial q^j} \frac{1}{n^2} \tag{4.39}$$

where (g^{ij}) is the inverse of the metric, as usual. Note that this last equation can also be written intrinsically as

$$\|\text{grad } S\|^2 = n^2 \tag{4.40}$$

as is readily verified. Equation 4.40 is known in optics as the *eikonal equation*.

We now explore the wave like description of the propagation of light by means of *wavefronts*. Fix a point $q_{(1)}$ in Q and consider the set of points $q_{(2)}$ such that there exists a light ray from $q_{(1)}$ to $q_{(2)}$ that traverses its path in time less than or equal to t. We denote this set by $\Psi(q_{(1)}, t)$. From the previous definitions, it follows that $\Psi(q_{(1)}, t)$ is precisely the set of points $q_{(2)}$ such that $S(q_{(1)}, q_{(2)},) \leq t$. The manifold boundary of $\Psi(q_{(1)}, t)$ is known as the wavefront of the source at $q_{(1)}$ after a time t, and we denote it by $\Phi(q_{(1)}, t) = \partial\Psi(q_{(1)}, t)$. With these new definitions, we can formulate an important theorem of Huygens about the relationship between wavefronts at different times:

Theorem 4.2.1 (Huygens' principle). *Let $\Phi(q_{(1)}, t)$ be the wavefront of $q_{(1)}$ after a time t. For every q in $\Phi(q_{(1)}, t)$, consider the wavefronts $\Phi(q, s)$ after a time s. Then, $\Phi(q_{(1)}, t + s)$ is the envelope of the wavefronts $\Phi(q, s)$.*

To prove the theorem, one shows that given a point $q_{(t+s)}$ in $\Phi(q_{(1)}, t + s)$, there exists $q_{(t)}$ in $\Phi(q_{(1)}, t)$ such that the wavefronts $\Phi(q_{(1)}, t + s)$ and $\Phi(q_{(t)}, s)$ are tangent at $q_{(t+s)}$ [Arn74, chapter 9]. We have that the wavefronts are actually the preimages of the eikonal, that is

$$\Phi(q_{(1)}, t) = S(q_{(1)}, \cdot)^{-1}(t). \tag{4.41}$$

The main point of the wavefront view is to consider the wavefronts at different times, which implicitly gives the evolution of the light rays coming from $q_{(1)}$. We can thus speak of a unique wavefront that propagates in space as a function of time, in the same way as a wave does. Since the problem now is to describe the time evolution of the wavefront, we consider the gradient of the eikonal as a measure of its velocity, grad S. From the eikonal equation 4.40, we have that the velocity of propagation of the wavefront is completely determined by the characteristics of the medium. Moreover, it follows form relations 4.37 that the expression of the gradient in coordinates is

$$\text{grad } S = g^{ij}\frac{\partial S}{\partial q^j}\frac{\partial}{\partial q^i} = g^{ij}\frac{\partial L}{\partial v^j}\frac{\partial}{\partial q^i} = \frac{n^2}{L}v^i\frac{\partial}{\partial q^i} \tag{4.42}$$

where we have written v^i for the velocity of the path γ joining $q_{(1)}$ and $q_{(2)}$, which is a function of the endpoints and depends on the parametrization of the ray. Thus, we find that the rays are parallel to the gradient of the eikonal, which implies that they are orthogonal to the wavefronts. This way, we recover the rays from the integral curves of grad S, which give a particular parametrization of the trajectories of light.

We have thus shown two approaches to geometrical optics. The ray picture focuses on the particle nature of light, which reduces the problem to the study of trajectories in the medium, while the wavefront interpretation considers the level sets of a function that characterises the propagation of light. The first point of view is very similar to the geometric formulation of classical mechanics, where a system of particles is described by its trajectories arising from a variational principle. On the other hand, the eikonal in the wavefront approach is the optical analogue of the action function in mechanics.

This wonderful duality is not surprising, since Hamilton found the principles of mechanics by generalizing the most prominent features of geometrical optics. In particular, the characteristic functions arising in the theory are found to satisfy a partial differential equation, the Hamilton-Jacobi equation in mechanics or the eikonal equation in optics. In the next section, we study whether it is possible to find a converse of these facts, that is, to find solutions to the PDE that produce valid trajectories of the system under consideration. Thus while the analogy between rays and dynamical trajectories is very natural, there is the other, most remarkable, of wavefronts and partial differential equations.

4.3 Jacobi's Method

The previous situations are not the only context in which the Hamilton-Jacobi equation appears. In the Hamiltonian formalism, it is clear that symplectomorphisms play a major part, since they preserve the symplectic structure and transform Hamiltonian systems into new ones. In chapter 2 we saw that a more general class of maps, the canonical transformations, can be described via generating functions 2.1.18. The idea of Jacobi's method is to find a canonical transformation such that the motion of the original mechanical system can be trivially integrated. We show how the desired generating functions satisfy the Hamilton-Jacobi equation.

We start by recalling the results in chapter 2 concerning Lagrangian submanifolds and symplectic maps (particularly 2.1.22). Let (M, ω) and (N, ρ) be symplectic manifoldsn and let $\Omega = \omega \ominus \rho$. Let $f : M \to N$ be a diffeomorphism, and $j_f : \Gamma_f \hookrightarrow M \times N$ its graph.

Ω is symplectic, and in particular closed, so by the Poincaré lemma we can locally write $\Omega = -d\Theta$ where $\Theta \in \Omega^1(M \times N)$ is a 1-form. Then, $j_f^* \Omega = -j_f^*(d\Theta) = -d(j_f^*\Theta)$ and f is symplectic if and only if $j_f^*\Theta \in \Omega^1(\Gamma_f)$ is closed. Using the Poincaré lemma once again, we may write $j_f^*\Theta = -dS$ where $S : \Gamma_f \to \mathbb{R}$ is defined locally. The function S is the *generating function* of the symplectic map f, and depends on the choice of Θ.

If $\omega = -d\theta$, $\rho = -d\vartheta$, we can always pick $\Theta = \mathrm{pr}_1^*\theta - \mathrm{pr}_2^*\vartheta$ and the generating function coincides with the ones found in 2.1.18. This is easily seen by introducing Darboux coordinates (Q^i, P_i) and (q^i, p_i) in M and N respectively, and choosing (Q^i, P_i, q^i, p_i) in the product $M \times N$. (Q^i, P_i) are also coordinates in Γ_f, since $M \to \Gamma_f$, $x \mapsto (x, f(x))$ is a diffeomorphism. In these coordinates, the inclusion is written $\hat{j}_f(Q, P) = (Q^i, P_i, \hat{f}(Q, P)) = (Q, P, q(Q, P), p(Q, P))$ for some functions q, p, and thus $j_f^*\Theta = P_i dQ^i - p(Q, P)d(q(Q, P)) = \theta - f^*\vartheta = -dS$.

There are other possible coordinates on the graph Γ_f. As an example, let us write again $q(Q, P)$, $p(Q, P)$ and assume

$$\left|\frac{\partial(Q^1, \dots, Q^m, q^1, \dots, q^m)}{\partial(Q^1, \dots, Q^m, P_1, \dots, P_m)}\right| = \left|\frac{\partial(q^1, \dots, q^m)}{\partial(P_1, \dots, P_m)}\right| \neq 0 \tag{4.43}$$

holds at some point $(Q_\circ, P_\circ)$. By the inverse function theorem, the coordinate map $\psi(Q, P) = (Q, q(Q, P))$ is invertible at some neighbourhood of $(Q_\circ, q_\circ)$ where $q_\circ = q(Q_\circ, P_\circ)$, and we can write $P = P(Q, q)$. We also have the p as a function of the (Q, q), by substituting the dependence of P in the definition: $p = p(Q, P(Q, q))$. It follows that the functions $(Q^1, \dots, Q^m, q^1, \dots, q^m)$ are local coordinates in Γ_f.

In these coordinates, $P_i(Q, q)dQ^i - p_i(Q, q)dq^i = -dS$ is written as

$$p_i(Q, q) = \frac{\partial S}{\partial q^i}(Q, q), \quad P_i(Q, q) = -\frac{\partial S}{\partial Q^i}(Q, q) \tag{4.44}$$

where S is a function of the (Q, q).

We develop Jacobi's idea a bit more, before formalizing it. Let (M, ω, H) be a Hamiltonian system, (N, ρ) a symplectic manifold and $f : N \to M$ a symplectomorphism. The map f induces a Hamiltonian system (N, ρ, K) with $K = f^*(H)$. Introducing Darboux coordinates

(q^i, p_i) in M, and (Q^i, P_i) in N, let us assume that the new Hamiltonian K depends only on the Q coordinates, that is,

$$\frac{\partial K}{\partial P_i} = 0 \ (i = 1, \ldots, m), \quad K = K(Q).$$

The dynamical paths in the new Hamiltonian system are trivially found, since Hamilton's equations in N take the form:

$$\dot{Q}^i(t) = 0, \quad \dot{P}_i(t) = -\frac{\partial K}{\partial Q^i}(Q(t)) \tag{4.45}$$

and direct integration gives

$$Q^i(t) = Q^i_\circ, \quad P_i = P_i^\circ - \frac{\partial K}{\partial Q^i}(Q_\circ)(t - t_\circ). \tag{4.46}$$

for the initial conditions $Q^i(t_\circ) = Q^i_\circ$, $P_i(t_\circ) = P_i^\circ$.

Now let S be the generating function of f defined from the symplectic potentials, and assume that we can choose (q, Q) as coordinates in Γ_f. From the previous observation, we have the relations

$$p_i = \frac{\partial S}{\partial q^i}(q, Q), \quad P_i = -\frac{\partial S}{\partial Q^i}(q, Q). \tag{4.47}$$

f is an isomorphism of Hamiltonian systems, so it turns dynamical trajectories in N into dynamical trajectories in M. Let γ be a trajectory in M, and $f \circ \gamma$ the corresponding trajectory in N. Then, the path $\gamma \times (f \circ \gamma)$ is in Γ_f and is written in coordinates as $(q(t), Q_\circ)$ where the $Q^i_\circ$ are constants. The q are solutions of the equations of motion, which are written as

$$\dot{q}^i = \frac{\partial H}{\partial p_i}\left(q(t), \frac{\partial S}{\partial q^i}(q(t), Q_\circ)\right), \quad \dot{p}_i(t) = -\frac{\partial H}{\partial q^i}\left(q(t), \frac{\partial S}{\partial q^i}(q(t), Q_\circ)\right) \tag{4.48}$$

and by the coordinate representation, we also have

$$\begin{aligned}
\dot{p}_i(t) &= \frac{d}{dt}\left(\frac{\partial S}{\partial q^i}(q(t), Q_\circ)\right) \\
&= \frac{\partial^2 S}{\partial q^j \partial q^i}(q(t), Q_\circ)\dot{q}^j(t) \\
&= \frac{\partial^2 S}{\partial q^j \partial q^i}(q(t), Q_\circ)\frac{\partial H}{\partial p_j}\left(q(t), \frac{\partial S}{\partial q^i}(q(t), Q_\circ)\right)
\end{aligned}$$

and equating the last expression to the corresponding one in 4.48, we obtain

$$\frac{\partial H}{\partial q^i}\left(q(t), \frac{\partial S}{\partial q^i}(q(t), Q_\circ)\right) + \frac{\partial H}{\partial p_j}\left(q(t), \frac{\partial S}{\partial q^i}(q(t), Q_\circ)\right)\frac{\partial^2 S}{\partial q^j \partial q^i}(q(t), Q_\circ) = 0. \tag{4.49}$$

Assume now that N is the cotangent bundle T^*Q of some manifold Q, and let (q^i, p_i) be the natural coordinates. If we consider the Q^i as constants, then S can be seen as a function on Q that is determined by the parameters Q^i. The last equation is then equivalent to

$$d(H \circ dS) = 0 \tag{4.50}$$

and the function $H \circ dS$ is constant, that is,

$$H \circ dS = E \tag{4.51}$$

for a constant value E. 4.51 is known as the *time independent Hamilton-Jacobi equation*, but since we will mostly work with this expression we refer to it as the *Hamilton-Jacobi equation* throughout the rest of the work. Note that we recover 4.51 from the time dependent Hamilton-Jacobi equation in 4.20 by assuming a solution of the latter of the form

$$S(t,q) = W(q) - Et, \tag{4.52}$$

where E is a constant and W is a function in Q satisfying the time independent equation.

We now formalize the previous ideas in a geometrical manner, by proving the following theorem:

Theorem 4.3.1 (Hamilton-Jacobi)**.** *Let the system (T^*Q, ω, H) be Hamiltonian, where $\omega = -d\theta$ is the canonical symplectic form on the cotangent bundle. Let $S : T^*Q \to \mathbb{R}$ be a function. The following are equivalent:*

(i) For every curve γ in Q such that

$$\gamma' = T\pi_Q \circ Z_H \circ dS \circ \gamma$$

the curve $dS \circ \gamma$ is an integral curve of Z_H.

(ii) S is a solution of the Hamilton-Jacobi equation $H \circ dS = E$, a constant, or in coordinates:

$$H\left(q^i, \frac{\partial S}{\partial q^i}\right) = E$$

We prove an easy lemma first:

Lemma 4.3.2. *Let $v, w \in T(T^*Q)$ and $S : Q \to \mathbb{R}$. We have*

$$\omega(T(dS \circ \pi_Q) \cdot v, w) = \omega(v, w - T(dS \circ \pi_Q) \cdot w)$$

Proof of the lemma. We have, by 2.2.3, that $(dS)^*\omega = -d(dS)^*\theta = -d^2S = 0$ so

$$\omega(\mathrm{T}(dS \circ \pi_Q) \cdot v, w) = \omega(\mathrm{T}(dS \circ \pi_Q) \cdot v, w - \mathrm{T}(dS \circ \pi_Q) \cdot w)$$

and the statement of the lemma is the same as

$$\omega(v - \mathrm{T}(dS \circ \pi_Q) \cdot v, w - \mathrm{T}(dS \circ \pi_Q) \cdot w) = 0$$

which is trivial, since $v - \mathrm{T}(dS \circ \pi_Q) \cdot v$ is vertical, and ω is zero on two vertical vectors (which is obvious in the natural coordinates). $\square$

Proof of the theorem. Let γ be such that $\gamma' = \mathrm{T}\pi_Q \circ Z_H \circ dS \circ \gamma$. We have

$$\begin{aligned}(dS \circ \gamma)' &= \mathrm{T}(dS) \circ \gamma' \\ &= \mathrm{T}(dS) \circ \mathrm{T}\pi_Q \circ Z_H \circ dS \circ \gamma \\ &= \mathrm{T}(dS \circ \pi_Q) \circ Z_H \circ dS \circ \gamma.\end{aligned}$$

Let $v \in \mathrm{T}_{dS(\gamma(t))}\mathrm{T}^*Q$, then

$$\begin{aligned}\omega((dS \circ \gamma)', v) &= \omega(\mathrm{T}(dS \circ \pi_Q) \cdot Z_H(dS(\gamma(t))), v) \\ &= \omega(Z_H(dS(\gamma(t))), v) - \omega(Z_H(dS(\gamma(t))), \mathrm{T}(dS \circ \pi_Q) \cdot v) \\ &= \langle dH \circ (dS \circ \gamma)(t), v\rangle - \langle dH \circ (dS \circ \gamma)(t), \mathrm{T}(dS \circ \pi_Q) \cdot v\rangle\end{aligned}$$

and if $w \in \mathrm{T}_q Q$, it follows that

$$\begin{aligned}\langle dH(dS(q)), \mathrm{T}_q(dS) \cdot w\rangle &= \langle {}^t\mathrm{T}_q(dS) \cdot dH(dS(q)), w\rangle \\ &= \langle d_q(H \circ dS), w\rangle\end{aligned}$$

which means that $dS \circ \gamma$ is an integral curve of Z_H if and only if $d_{\gamma(t)}(H \circ dS) = 0$, where the γ are the integral curves of a certain field in Q. Varying the integral curves, $\gamma(t)$ takes every value in Q, and the last statement is the same as $d(H \circ dS) = 0$. Thus, $dS \circ \gamma$ is an integral curve of Z_H if and only if $d(H \circ dS) = 0$ which means $H \circ dS = E$, for a constant E, and the equivalence follows. □

In the particular case of mechanical Lagrangians ($L = T - U$) we already know that $H(q,p) = \frac{1}{2} g^{ij} p_i p_j + U(q)$ so $\frac{\partial H}{\partial p_i} = g^{ij} p_j$ and the condition $\gamma' = \mathrm{T}\pi_Q \circ Z_H \circ dS \circ \gamma$ is written as

$$\dot{q}^i(t) = \frac{\partial H}{\partial p_i}\left(q(t), \frac{\partial S}{\partial q}(q(t))\right) = g^{ij}(q(t)) \frac{\partial S}{\partial q^j}(q(t)) \tag{4.53}$$

which is equivalent to

$$\gamma' = \mathrm{grad}\ S \circ \gamma, \tag{4.54}$$

where the gradient is taken with respect to the metric defining the kinetic energy. Thus, we see that the condition on the trajectories in Q to produce mechanical paths in T^*Q via dS is that they be integral curves of the gradient, as for the rays in geometrical optics. Equivalently, if the trajectories in Q are orthogonal to the level sets of a solution of the Hamilton-Jacobi equation, then we can lift those trajectories to T^*Q and obtain valid solutions of Hamilton's equations of motion.

We have shown that there are some integral curves of the Hamiltonian system (T^*Q) that can be integrated in a lower dimensional manifold (Q) if we can find an appropriate solution of the Hamilton-Jacobi equation. We investigate this phenomenon in abstract in the next chapter.

Note that a solution of the Hamilton-Jacobi equation gives a family of integral curves of the original system, but by no means does it have to reproduce the whole dynamics of the system.

Example 4.3.3. Consider the Hamiltonian of a free particle with two degrees of freedom $H(q^1, q^2, p^1, p^2) = \frac{1}{2}(p^1)^2 + \frac{1}{2}(p^2)^2$. The function $S(q^1, q^2) = \alpha q^1$ with $\alpha^2 = 2E$ is clearly a solution of the Hamilton-Jacobi equation. If $\gamma(t) = (q^1(t), q^2(t)) \in \mathbb{R}^2$, then the equation $\gamma' = \mathrm{T}\pi_Q \circ Z_H \circ dS \circ \gamma$ is written:

$$\dot{q}^1 = \alpha, \quad \dot{q}^2 = 0$$

whose solutions are of the form

$$q^1(t) = \alpha t + \text{constant}, \quad q^2(t) = \text{constant}$$

and lifting the solution to $\mathrm{T}(\mathbb{R}^2) \cong \mathbb{R}^4$ by $dS \circ \gamma$, we get

$$q^1(t) = \alpha t + \text{constant}, \quad q^2(t) = \text{constant}, \quad p^1 = \alpha, \quad p^2 = 0$$

which reproduces all the possible motions of the particle in the direction determined by q^1, but not the general motion of the particle.

To be able to produce all the admissible trajectories of the original problem, we must find a *complete solution* of the Hamilton-Jacobi equation. The general solution to a partial differential equation depends in general on an arbitrary function of the coordinates. A complete solution is a solution that depends on m parameters, where m is the number of degrees of freedom in the original system. By varying the parameters and solving the corresponding dynamics in Q, we obtain the whole motion in the system T^*Q via the map dS. More precisely:

Definition 4.3.4. A *complete solution* of the Hamilton-Jacobi equation is a function $S(q;\alpha)$ on Q depending on m parameters α^i defined in some Euclidean space, such that

(i) For each value of α^i, $S(q;\alpha)$ is a solution of the Hamilton-Jacobi equation: $H \circ dS(\cdot,\alpha) = E$

(ii) The matrix $\left(\dfrac{\partial^2 S}{\partial q^i \partial \alpha^j}\right)$ is invertible.

The following result justifies the definition:

Theorem 4.3.5. *Let $S(q,\alpha)$ be a function given on a neighbourhood of some point $(q_\circ,\alpha_\circ)$ of the Cartesian product of two m-dimensional real vector spaces $\mathbb{R}^m \times \mathbb{R}^m$. If*

$$\left|\frac{\partial^2 S}{\partial q^i\,\partial \alpha^j}\right|_{q_\circ,\alpha_\circ} \neq 0$$

then S is the generating function of some canonical transformation.

Proof. Consider the m equations

$$\frac{\partial S}{\partial \alpha^i}(q,\alpha) = p_i, \quad (i = 1,\dots,m)$$

where $p_i \in \mathbb{R}$. By hypothesis, we can apply the implicit function theorem, and the α can be written as a function of (q,p) in some neighbourhood of $\left(q^i_\circ, \dfrac{\partial S}{\partial q_i}(q_\circ,\alpha_\circ)\right)$. Let us write these functions as $\alpha^i = Q^i(q,p)$.

Consider now the m functions

$$\beta^i(q,\alpha) = -\frac{\partial S}{\partial \alpha^i}, \quad (i = 1,\dots,m)$$

and set

$$P^i(q,p) = \beta^i(q,Q(q,p)).$$

The map sending (q,p) to $(Q(q,p),P(q,p))$ is a local diffeomorphism, since

$$\left|\frac{\partial(Q,P)}{\partial(q,p)}\right| = \left|\frac{\partial(Q,P)}{\partial(q,Q)}\cdot\frac{\partial(q,Q)}{\partial(q,p)}\right| = \left|\frac{\partial^2 S}{\partial q \partial \alpha}\right| \cdot \left|\frac{\partial^2 S}{\partial q \partial \alpha}\right|^{-1}_{q_\circ,\alpha_\circ} = 1 \neq 0$$

In fact,

$$\left|\frac{\partial(q,Q)}{\partial(q,p)}\right| = \left|\frac{\partial^2 S}{\partial q \partial \alpha}\right|^{-1}_{q_\circ,\alpha_\circ} \neq 0$$

and (q,Q) are local coordinates in the graph of the map. By construction, S is the generating function of the map $(Q,P) \mapsto (q,p)$. □

Example 4.3.6. A complete solution for 4.3.3 would be the function

$$S(q^1, q^2) = \alpha_1 q^1 + \alpha_2 q^2$$

where α_1, α_2 are real numbers such that $\alpha_1^2 + \alpha_2^2 = 2E$. Observe that $\left|\frac{\partial^2 S}{\partial q \partial \alpha}\right| = 1$, and the solution reproduces all the possible trajectories:

$$q^1(t) = \alpha_1 t + \text{constant},\ q^2(t) = \alpha_2 t + \text{constant},\ p_1(t) = \alpha_1,\ p_2(t) = \alpha_2. \tag{4.55}$$

4.4 Solving the Hamilton-Jacobi Equation

It is clear that the problem of finding a complete solution to a partial differential equation is in general much harder than integrating a system of ODEs, but there are many situations where the former can be solved while the solution of the later is unknown. The most popular method to produce complete solutions of the Hamilton-Jacobi equation is that of separation of variables. When the system possesses symmetries that suggest a natural change of variables, one looks for solutions which are a combination of functions of such variables and reduce the problem to an integrable set of ordinary differential equations. We formalize this argument and solve some representative examples with Jacobi's method.

Assume that the coordinate q^1 and the corresponding momentum $p_1 = \frac{\partial S}{\partial q^1}$ appear in the Hamilton-Jacobi equation as a combination $\phi\left(q^1, \frac{\partial S}{\partial q^1}\right)$ which does not involve any other coordinate. That is, the Hamiltonian in the Hamilton-Jacobi equation is of the form:

$$H\left(q^2, \ldots, q^m, \frac{\partial S}{\partial q^2}, \ldots, \frac{\partial S}{\partial q^m}, \phi\left(q^1, \frac{\partial S}{\partial q^1}\right)\right). \tag{4.56}$$

We look for a separable solution $S(q^1, \ldots, q^m) = S_1(q^1) + S_2(q^2, \ldots, q^m)$. The Hamilton-Jacobi is then

$$H\left(q^2, \ldots, q^m, \frac{\partial S_2}{\partial q^2}, \ldots, \frac{\partial S_2}{\partial q^m}, \phi\big(q^1, S_1'(q^1)\big)\right) = E. \tag{4.57}$$

If we can solve for $\phi\big(q^1, S_1'(q^1)\big)$ in the last equation (a sufficient condition would be $\frac{\partial H}{\partial \phi} \neq 0$), then we get an identity of the form

$$\phi\big(q^1, S_1'(q^1)\big) = \Phi\left(q^2, \ldots, q^m, \frac{\partial S_2}{\partial q^2}, \ldots, \frac{\partial S_2}{\partial q^m}\right) \tag{4.58}$$

where each side has variables that do not appear in the other side. We conclude that the solution must satisfy

$$\phi\big(q^1, S_1'(q^1)\big) = \alpha_1 \tag{4.59}$$

where α_1 is a constant. Substituting α_1 in the original equation as a parameter, we get

$$H\left(q^2, \ldots, q^m, \frac{\partial S}{\partial q^2}, \ldots, \frac{\partial S}{\partial q^m}, \alpha_1\right) = E. \tag{4.60}$$

The first equation is an ODE and can be easily integrated, while the second is the Hamilton-Jacobi equation for a function in $m - 1$ coordinates, and we reduce the problem of finding a complete solution.

As a particular case of this situation, consider a Hamiltonian where the coordinate q^1 is cyclic, that is, such that it does not appear explicitly in the Hamiltonian $\frac{\partial H}{\partial q^1} = 0$. We know that the associated momentum p^1 is a constant of the motion, and we can take S of the form $S(q^1, \dots, q^m) = p_1 q^1 + S_2(q^2, \dots, q^m)$, where the conserved quantity is now a parameter.

Ideally, we would like to be able to apply this procedure repeatedly until we obtain a separable solution dependent on m parameters $\alpha_1, \dots, \alpha_m$

$$S(q^1, \dots, q^m; \alpha) = S_1(q^1; \alpha) + \cdots + S_m(q^m; \alpha). \tag{4.61}$$

While this method does not always work, it already gives solutions to an important class of problems.

Example 4.4.1. As we saw earlier, the two dimensional anisotropic Harmonic oscillator is described by the Hamiltonian

$$H = \frac{1}{2}(p^1)^2 + \frac{1}{2}(p^2)^2 + \frac{1}{2}\omega_1^2(q^1)^2 + \frac{1}{2}\omega_2^2(q^2)^2.$$

Assuming a solution of the Hamilton-Jacobi equation of the form $S = S_1(q^1) + S_2(q^2)$, we find that

$$E_1 = \frac{1}{2}(p^1)^2 + \frac{1}{2}\omega_1^2(q^1)^2, \quad E_2 = \frac{1}{2}(p^2)^2 + \frac{1}{2}\omega_2^2(q^2)^2 \tag{4.62}$$

where E_1 and E_2 are parameters such that $E_1 + E_2 = E$. It follows that

$$S = \int \pm\sqrt{2E_1 - \omega_1^2(q^1)^2}dq^1 + \int \pm\sqrt{2E_2 - \omega_2^2(q^2)^2}dq^2 \tag{4.63}$$

which is defined up to a constant. The sign in front of the square roots depends on the initial condition of the path under consideration. For simplicity, we will always assume that the initial momentum is positive (i.e. the sign of the root is positive).

By 4.3.1, we must solve the system

$$\dot{q}^1 = \{q^1, H\} = \sqrt{2E_1 - \omega_1^2(q^1)^2}, \quad \dot{q}^2 = \{q^2, H\} = \sqrt{2E_2 - \omega_2^2(q^2)^2} \tag{4.64}$$

which is trivially integrated

$$\int \frac{dq^1}{\sqrt{2E_1 - \omega_1^2(q^1)^2}} = t + \text{const}, \quad \int \frac{dq^2}{\sqrt{2E_2 - \omega_2^2(q^2)^2}} = t + \text{constant}. \tag{4.65}$$

and explicitly

$$\arcsin\left(\frac{\omega_1 q^1}{\sqrt{2E_1}}\right) = \omega_1 t + \text{constant}, \quad \arcsin\left(\frac{\omega_2 q^2}{\sqrt{2E_2}}\right) = \omega_2 t + \text{constant}. \tag{4.66}$$

Solving for q^1, q^2, we obtain the paths

$$q^1(t) = \frac{\sqrt{2E_1}}{\omega_1}\sin(\omega_1 t + \text{constant}), \quad q^2(t) = \frac{\sqrt{2E_2}}{\omega_2}\sin(\omega_2 t + \text{constant}) \tag{4.67}$$

which correspond to the general motion in the system, and we conclude that S is a complete solution.

Example 4.4.2. In spherical coordinates (r,θ,ϕ), the Hamiltonian of a particle in a conservative field is written as

$$H(r,\theta,\phi) = \frac{1}{2}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2\sin^2\theta}\right) + U(r,\theta,\phi)$$

and the variables can be separated if the potential is of the form

$$U = a(r) + \frac{b(\theta)}{r^2} + \frac{c(\phi)}{r^2\sin^2\theta}. \tag{4.68}$$

Assuming a separable solution to the Hamilton-Jacobi equation $S = S_r(r) + S_\theta(\theta) + S_\phi(\phi)$ we find the constant expressions

$$\alpha = \frac{1}{2}p_\phi^2 + c(\phi), \quad \beta = \frac{1}{2}p_\theta^2 + b(\theta) + \frac{\alpha}{\sin^2\theta}, \quad E = \frac{1}{2}p_r^2 + a(r) + \frac{\beta}{r^2} \tag{4.69}$$

whence we arrive at

$$S = \int \sqrt{2(\alpha - c(\phi))}d\phi + \int \sqrt{2(\beta - b(\theta) - \frac{\alpha}{\sin\theta^2})}d\theta + \int \sqrt{2(E - a(r) - \frac{\beta}{r^2})}dr \tag{4.70}$$

depending on the three parameters E, α and β. S is then easily seen to be a complete solution.

Example 4.4.3. The Kepler problem can also be integrated with Jacobi's method. Since the motion takes place in a plane, in polar coordinates the Hamilton-Jacobi equation reads

$$\frac{1}{2}\left(\left(\frac{\partial S}{\partial r}\right)^2 + \left(\frac{\partial S}{\partial \phi}\right)^2 \frac{1}{r^2}\right) - \frac{k}{r} = E \tag{4.71}$$

and ϕ is cyclic. Hence, we assume a solution of the form $S(r,\phi) = S_r(r) + p_\phi\phi$ and clearly

$$S_r(r) = \int \sqrt{2E + 2\frac{k}{r} - \frac{p_\phi}{r^2}}dr. \tag{4.72}$$

The motion in (r,ϕ) is found from

$$\dot{r} = \sqrt{2E + 2\frac{k}{r} - \frac{p_\phi^2}{r^2}}, \quad \dot{\phi} = \frac{p_\phi}{r^2} \tag{4.73}$$

integrating:

$$t - t_\circ = \int_{r_\circ}^{r} \frac{dr}{\sqrt{2E + \frac{2k}{r} - \frac{p_\phi^2}{r^2}}}, \quad \phi - \phi_\circ = p_\phi \int_{r_\circ}^{r} \frac{dr}{r^2\sqrt{2E + \frac{2k}{r} - \frac{p_\phi^2}{r^2}}} \tag{4.74}$$

where the second equation comes from writing ϕ as a function of r and considering $\dfrac{d\phi}{dr} = \dfrac{\dot{\phi}}{\dot{r}}$, as before.

Example 4.4.4. The geometrical problem of finding the geodesics on a *triaxial ellipsoid* S (i.e. an ellipsoid whose axes have different lengths) can be solved with Jacobi's method in the following manner. We already showed in chapter 3 that the problem is equivalent to integrating the equations of motion for a particle which is constrained to move on the surface of the ellipsoid.

The ostensible symmetry of the constraint manifold suggests the use of *ellipsoidal coordinates* [CH37, vol. 1, ch. IV], which are defined at the point (x, y, z) as the three roots (ρ, σ, τ) of the equation in s

$$\frac{x^2}{s+a^2} + \frac{y^2}{s+b^2} + \frac{z^2}{s+c^2} = 1, \tag{4.75}$$

where $a > b > c$ and the solutions are ordered decreasingly as

$$\rho \geq -c^2 \geq \sigma \geq -b^2 \geq \tau \geq -a^2. \tag{4.76}$$

If we see S as being embedded in $\mathbb{R}^3$ and choose a, b, c as the lengths of its axes, (ρ, σ, τ) are then coordinates adapted to the submanifold S. More explicitly, $S = \{(\rho, \sigma, \tau) \,|\, \rho = 0\}$ and we work with the coordinates (σ, τ) on S.

If $j : S \hookrightarrow \mathbb{R}^3$ is the obvious embedding, elementary calculations show that in coordinates

$$x(\sigma,\tau) = a\sqrt{\frac{(\sigma+a^2)(\tau+a^2)}{(b^2-a^2)(c^2-a^2)}}, \quad y(\sigma,\tau) = b\sqrt{\frac{(\sigma+b^2)(\tau+b^2)}{(c^2-b^2)(a^2-b^2)}}, \tag{4.77}$$

$$z(\sigma,\tau) = c\sqrt{\frac{(\sigma+c^2)(\tau+c^2)}{(a^2-c^2)(b^2-c^2)}}. \tag{4.78}$$

Using the classical notation of the theory of surfaces, the coefficients of the first fundamental form are

$$E = \frac{\partial^2 x}{\partial\sigma^2} + \frac{\partial^2 y}{\partial\sigma^2} + \frac{\partial^2 z}{\partial\sigma^2}, \quad F = \frac{\partial^2 x}{\partial\sigma\partial\tau} + \frac{\partial^2 y}{\partial\sigma\partial\tau} + \frac{\partial^2 z}{\partial\sigma\partial\tau}, \quad G = \frac{\partial^2 x}{\partial\tau^2} + \frac{\partial^2 y}{\partial\tau^2} + \frac{\partial^2 z}{\partial\tau^2} \tag{4.79}$$

and the Lagrangian of the mechanical system is written as

$$L(\sigma,\tau,v^\sigma,v^\tau) = \frac{1}{2}(E(v^\sigma)^2 + 2Fv^\sigma v^\tau + G(v^\tau)^2). \tag{4.80}$$

Inverting the metric, the Hamiltonian reads

$$H(\sigma,\tau,p_\sigma,p_\tau) = \frac{1}{2}\frac{1}{EG-F^2}(Gp_\sigma^2 - 2Fp_\sigma p_\tau + Ep_\tau^2). \tag{4.81}$$

After some calculations, one arrives at the expressions

$$E = (\sigma-\tau)A(\sigma), \quad F = 0, \quad G = (\tau-\sigma)A(\tau) \tag{4.82}$$

where

$$A(t) = \frac{1}{4}\frac{t}{(t+a^2)(t+b^2)(t+c^2)} \tag{4.83}$$

and the Hamiltonian takes the simple form

$$H = \frac{1}{2(\sigma-\tau)}\Big(-\frac{p_\tau^2}{A(\tau)} + \frac{p_\sigma^2}{A(\sigma)}\Big). \tag{4.84}$$

Writing the Hamilton-Jacobi equation $H \circ dS = \alpha$ and looking for a solution of the form $S(\sigma,\tau) = S_\sigma(\sigma) + S_\tau(\tau)$ we see that

$$2\alpha\sigma - \frac{(S_\sigma')^2}{A(\sigma)} = 2\alpha\tau - \frac{(S_\tau')^2}{A(\tau)} = -\beta \tag{4.85}$$

for some constant β. These yield the solution

$$S = \int \sqrt{(2\alpha\tau + \beta)A(\tau)}d\tau + \int \sqrt{(2\alpha\sigma + \beta)A(\sigma)}d\sigma. \tag{4.86}$$

The equations of motion are

$$\dot{\sigma} = \frac{\partial H}{\partial p_\sigma} = \frac{1}{(\sigma - \tau)}\frac{\sqrt{2\alpha\sigma + \beta}}{\sqrt{A(\sigma)}}, \quad \dot{\tau} = \frac{\partial H}{\partial p_\tau} = -\frac{1}{(\sigma - \tau)}\frac{\sqrt{2\alpha\tau + \beta}}{\sqrt{A(\tau)}} \tag{4.87}$$

and assuming that we can express σ as a function of τ, we get

$$\frac{d\sigma}{d\tau} = \frac{\dot{\sigma}}{\dot{\tau}} = -\frac{\sqrt{2\alpha\sigma + \beta}}{\sqrt{2\alpha\tau + \beta}}\frac{\sqrt{A(\tau)}}{\sqrt{A(\sigma)}} \tag{4.88}$$

which is a separable ODE and can be integrated, giving the trajectory implicitly as

$$\int \sqrt{\frac{A(\sigma)}{2\alpha\sigma + \beta}}d\sigma + \int \sqrt{\frac{A(\tau)}{2\alpha\tau + \beta}}d\tau = \gamma \tag{4.89}$$

where γ is a constant.

4.5 Action-Angle Coordinates

In order to integrate a system of $2m$ ordinary differential equations, we must find in general $2m$ constants of the motion. If the system is Hamiltonian, in many cases it suffices to find only m conserved quantities. When it comes to integrability, each constant reduces the degrees of freedom of the system by two.

In this section we study the general phenomenon. It is clear that the constants of the motion that completely reduce the problem must satisfy some *independence* condition between them. We make the following suggestive definition:

Definition 4.5.1. Let (M, ω, H) be Hamiltonian system with $\dim M = 2m$. We say that the system is *completely integrable* if there exist m functions F_i such that

(i) $\{F_i, H\} = 0$ for all i (i.e. the F_i are constants of the motion).

(ii) $\{F_i, F_j\}$ for all i, j (i.e. the F_i are in involution).

(iii) $dF_1 \wedge \ldots \wedge dF_m \neq 0$ everywhere (i.e. the F_i are linearly independent)

The m parameters $Q^1, \ldots, Q^m$ in a complete solution of the Hamilton-Jacobi equation satisfy the requirements of the definition. They are indeed constants of the motion if we write them as a combination of the position and momenta. Since S can be viewed as the generating function of a symplectic map f, we have

$$\{Q^i, Q^j\} = \{f^*(q^i), f^*(q^j)\} = f^*\{q^i, q^j\} = 0$$

and they are in involution. Finally, independence follows from the completeness of the solution.

Completely integrable systems will be found to be completely integrable in the sense of the equations of motion, which highlights the importance of the complete solutions of the Hamilton-Jacobi equation.

Among the systems we have studied so far, there are the ones where every coordinate is ignorable. In these cases, the Hamiltonian is a function of the momenta only, and the motion is trivially integrated. We have also found systems whose trajectories are closed curves and correspond to a periodic motion, as in the simple harmonic oscillator. With this example in mind, we define the *action-angle coordinates*:

Definition 4.5.2. A Hamiltonian system (M, ω, H) admits *action-angle coordinates* (φ, I) in some open set $U \subset M$, if:

(i) There exists a symplectomorphism $\psi : U \to \mathbb{T}^m \times B^m$, between the m-torus $\mathbb{T}^m = \mathbb{S}^1 \times \cdots \mathbb{S}^1$ and a ballB^m in $\mathbb{R}^m$.

(ii) The induced Hamiltonian $\psi_*(H)$ does not depend on the φ coordinates.

The torus accounts for the periodic motion of the system, and the equations of motion are trivially integrated.

If $F_1, \ldots, F_m$ are constants of the motion, we write their cartesian product as $F = F_1 \times \ldots \times F_m : M \to \mathbb{R}^m$. Given a point c in $\mathbb{R}^m$, we denote its preimage by F as $\Sigma_c = F^{-1}(c)$. Σ_c is an *energy surface.*

We conclude the chapter presenting an important theorem, which shows that completely integrable systems admit action-angle coordinates by imposing a topological condition over the energy surfaces.

Theorem 4.5.3. *Let (M, ω, H) be a completely integrable Hamiltonian system, and let the sets Σ_c be compact in M. Then the Σ_c are diffeomorphic to the torus $\mathbb{T}^m$ and the system admits action-angle coordinates.*

We give the main ideas of the proof, without proving some delicate intermediate steps. We follow [Arn74] and [AM78].

Proof. The constants of motion define m Hamiltonian fields $Z_{F_i} = \omega^\sharp dF_i$, which are tangent to the Σ_c, since

$$\mathscr{L}_{Z_{F_j}} F_i = \langle dF_i, Z_{F_j} \rangle = \omega(Z_{F_i}, Z_{F_j}) = \{F_i, F_j\} = 0$$

and the F_i are invariant by the flow of the fields. The Z_{F_i} are linearly independent by the independence of the differentials, which implies that the Σ_c are regular submanifolds of M and the fields form a base of the tangent space. Moreover, the fields commute with respect to the Lie bracket, since we know that

$$[Z_{F_i}, Z_{F_j}] = Z_{\{F_i, F_j\}} = 0.$$

One then uses the commuting flows of the m fields to construct a diffeomorphism from Σ_c to the torus $\mathbb{T}^m$ [Arn74, chapter 10].

Assume that the Darboux coordinates in M are written (q, p) and take $\theta = p_i dq^i$ as the symplectic potential. Let c be in some open set $U \subset \mathbb{R}^m$ and consider the fundamental cycles $\gamma_1(c), \ldots, \gamma_m(c)$ in $\Sigma_c \cong \mathbb{T}^m$, that is, cycles that form a basis of the first homology group $H_1(\mathbb{T}^m) \cong \mathbb{R}^m$. Let $\lambda = \lambda_1 \times \ldots \times \lambda_m : U \to \mathbb{R}^m$ where each λ_i is defined by the integral

$$\lambda_i(c) = \oint_{\lambda_i(c)} j_c^*(\theta),$$

where we have written $j_c : \Sigma_c \hookrightarrow M$ for the inclusion of Σ_c. We assume that λ is a diffeomorphism onto its image. Then, taking a smaller U we can assume $\lambda(U) = B^m$, obtaining a map $\lambda \circ F$:

$F^{-1}(U) \to B^m$ which is half of the diffeomorphism ψ. We must now find a map Γ in $F^{-1}(U)$ such that $\psi = \Gamma \times (\lambda \circ F)$, and so gives the angle coordinates.

We observe that $j_c^*(\theta)$ is a closed one form in Σ_c, since the Z_{F_i} form a basis of the tangent space of Σ_c and

$$d(j_c^*(\theta))(Z_{F_i}, Z_{F_j}) = -j_c^*\omega(Z_{F_i}, Z_{F_j}) = -j_c^*\{F_i, F_j\} = 0.$$

Because the dF_i are independent, we have that $\det \dfrac{\partial F_i}{\partial p_j} \neq 0$ and the equation $F(q,p) - \lambda^{-1}(I) = 0$ for fixed I can be solved for p in a neighbourhood of $q_\circ$. We write $p = p(q, I)$.

We define

$$S(q, I) = \int_{(q_\circ, p_\circ)}^{(q,p)} j^*_{\lambda^{-1}(I)}(\theta)$$

where the integral is taken over any path in $\Sigma_{\lambda^{-1}(I)}$ joining $(q_\circ, p_\circ)$ and (q, p). Since $j^*_{\lambda^{-1}(I)}(\theta)$ is closed, Stokes' theorem shows that the integral does not depend on the path, if (q, p) is sufficiently close to $(q_\circ, p_\circ)$. Globally the function does depend on the path of integration. We now define $\Gamma = \Gamma^1 \times \ldots \times \Gamma^m : F^{-1}(U) \to \mathbb{T}^m$ where

$$\Gamma^i(q, p) = \frac{\partial S(q, I)}{\partial I_i}\Big|_{I=(\lambda \circ F)(q,p)}$$

are multivalued functions. The variation of Γ^i along the cycle $\gamma_j(\lambda^{-1}(I))$ is easily shown to be

$$\oint_{\gamma_j(\lambda^{-1}(I))} d(\Gamma^i \circ j_{\gamma^{-1}(I)}) = \delta_i^j$$

which means that the functions are well defined *modulo* 1 and determine angle coordinates in $\mathbb{T}^m$.

We finally define $\psi = \Gamma \times (\lambda \circ F) : F^{-1}(U) \to \mathbb{T}^m \times B^m$. It can be shown that

$$\frac{\partial S(q, I)}{\partial q^i} = p_i(q, I)$$

which, together with

$$\frac{\partial S(q, I)}{\partial I_i} = \Gamma^i(q, p)$$

proves that S is the generating function of $\psi : (q, p) \mapsto (\varphi = \Gamma, I)$. Thus, the map ψ is symplectic and consequently a local diffeomorphism. It can be shown that ψ is bijective, and thus gives a global diffeomorphism. It is easy to see that the Hamiltonian is independent of the angle coordinates, as

$$\frac{\partial \psi_*(H)}{\partial \varphi^i} = dI_i(\psi_*(Z_H)) = d(\lambda_i \circ F)(Z_H) \circ \psi^{-1} = (d\lambda_i \circ \mathrm{T}F)(Z_H) \circ \psi^{-1}$$

but

$$\mathrm{T}F(Z_H) = (dF_1(Z_H), \ldots, dF_m(Z_H)) = (\{F_1, H\}, \ldots, \{F_m, H\}) = 0.$$

$\square$

Chapter 5

A Generalized Geometric Setting for Hamilton-Jacobi Theory

The classical Hamilton-Jacobi theory that we have reviewed in the previous chapter relies heavily on the symplectic structure of the cotangent bundle. In this chapter, we are interested in one particular aspect of the theory: the description of the dynamics of a system by means of the dynamics in a lower dimensional manifold. We investigate the analogue to the Hamilton-Jacobi equation for general dynamical systems, and study the role of the symplectic structure in the general framework. We follow the articles [CGM$^+$06] and [CGM$^+$16].

5.1 Dynamical Systems and Slicing Vector Fields

We have several structures on the carrier manifolds of mechanical systems, such as vector bundles and symplectic manifolds, which are crucial to define the dynamics. Nevertheless, the trajectories of most of these systems are determined by a vector field, whose construction differs from one formalism to the other. Forgetting about the structures present in these constructions, we are left with a general situation. For the sake of generality, we make the following definition:

Definition 5.1.1. A *dynamical system* is a pair (P, Z) where P is a manifold and Z a vector field on P.

The integral curves of Z in P are the trajectories of the dynamical system. It is clear how the Lagrangian and Hamiltonian dynamical systems are particular cases of this definition. Motivated by Jacobi's method, we are interested in relating the dynamics of a system with the dynamics of a smaller system. Let (P, Z) and (M, X) be dynamical systems and $\alpha : M \to P$ a differentiable map. We have the following result:

Proposition 5.1.2. *The following are equivalent:*

(i) For every integral curve ξ of X, $\gamma = \alpha \circ \xi$ is an integral curve of Z.

(ii) The fields X and Z are α-related *($X \underset{\alpha}{\sim} Z$), i.e.,*

$$T\alpha \circ X = Z \circ \alpha. \tag{5.1}$$

Proof. Assuming (i), by hypothesis we have

$$Z \circ \alpha \circ \xi = Z \circ \gamma = \gamma' = (\alpha \circ \xi)' = \mathrm{T}\alpha \circ \xi' = \mathrm{T}\alpha \circ X \circ \xi$$

and since the integral curves of X are defined locally at every point in M, we get (ii).

Conversely, assuming (ii), we have

$$\gamma' = (\alpha \circ \xi)' = \mathrm{T}\alpha \circ \xi' = \mathrm{T}\alpha \circ X \circ \xi = Z \circ \alpha \circ \xi = Z \circ \gamma$$

and the equivalence follows. $\square$

Clearly, condition $X \underset{\alpha}{\sim} Z$ is equivalent to the commutativity of the diagram

$$\begin{tikzcd} \mathrm{T}M \arrow[r, "\mathrm{T}\alpha"] \arrow[d, shift left] & \mathrm{T}P \arrow[d, shift left] \\ M \arrow[r, "\alpha"] \arrow[u, bend left, "X"] & P \arrow[u, bend right, "Z"'] \end{tikzcd}$$

The trajectories starting at $\alpha(M) \subset P$ are found by solving a system of differential equations in M, and lifting the solution to P via α. This way, (M, X) contains partial information about the dynamics of (P, Z), and can be thought of as a "slice" of the original system. We turn this terminology into a definition:

Definition 5.1.3. The triple (M, α, X) is a *slicing* of the dynamical system (P, Z) if it satisfies the *slicing equation* 5.1.

Observe that, in some cases, the field X in the slicing equation is completely determined by the map α, such as when α is an immersion (i.e. $\mathrm{T}\alpha$ injective) as can be readily seen from the definition. In particular, if α is an embedding, then $X = \alpha^*(Z)$ is simply the pullback of Z by α. In these cases, we might speak of (M, α) as the solution of the slicing equation.

If (M, α, X) is a solution of the slicing equation, we can construct other solutions by means of diffeomorphisms $\varphi : M \to \varphi(M)$. Thus, φ produces the slicing $(\varphi(M), \varphi_*(\alpha), \varphi_*(X))$, which is equivalent to the original. In particular, when α is an embedding, $\alpha(M) \overset{j}{\hookrightarrow} P$ is a regular submanifold and the slicing equation expresses the fact that Z is tangent to $\alpha(M)$

$$\begin{tikzcd} \mathrm{T}\alpha(M) \arrow[r, hook, "\mathrm{T}j"] \arrow[d, shift left] & \mathrm{T}P \arrow[d, shift left] \\ \alpha(M) \arrow[r, hook, "j"] \arrow[u, bend left, "Z_\circ"] & P \arrow[u, bend right, "Z"'] \end{tikzcd}$$

where $Z_\circ$ is the restriction of Z to $\alpha(M)$.

Later, we will show that the slicing equation is actually a generalization of the Hamilton-Jacobi equation, and we will investigate necessary conditions allowing to pass from one form of the equation to the other. As in the classical Hamilton-Jacobi theory, in order to reproduce the whole dynamics of (P, Z), we need a *complete solution* of the slicing equation. This amounts to a family of vector fields X_c indexed by some parameter space A and satisfying the slicing equation.

Definition 5.1.4. A *complete slicing* of (P, Z) is given by

(i) A surjective map $\overline{\alpha} : M \times A \to P$.

(ii) A vector field $\overline{X} : M \times A \to \mathrm{T}M$ along the projection $\mathrm{pr}_1 : M \times A \to M$.

(iii) For each $c \in A$, the map $\alpha_c \equiv \overline{\alpha}(\cdot, c) : M \to P$ and the vector field $X_c \equiv \overline{X}(\cdot, c) : M \to \mathrm{T}M$ constitute a slicing of Z:

$$\begin{tikzcd} \mathrm{T}M \times A \arrow[r, "\mathrm{T}_1\overline{\alpha}"] \arrow[d, shift left] & \mathrm{T}P \arrow[d, shift right] \\ M \times A \arrow[r, "\overline{\alpha}"] \arrow[u, bend left, "\overline{X}"] & P \arrow[u, bend right, "Z"'] \end{tikzcd}$$

In particular, the integral curves of Z constitute a slicing of the dynamics:

Example 5.1.5. Let $\alpha : I \to P$ be an integral curve of Z. Then, the triple $(I, \alpha, \frac{d}{dt})$ is a slicing of (P, Z), as can be seen from the definition of α':

$$\begin{tikzcd} \mathrm{T}I \arrow[r, "\mathrm{T}\alpha"] \arrow[d, shift left] & \mathrm{T}P \arrow[d, shift right] \\ I \arrow[r, "\alpha"] \arrow[u, bend left, "\frac{d}{dt}"] \arrow[ur, "\alpha'"] & P \arrow[u, bend right, "Z"'] \end{tikzcd}$$

If $z_\circ \in P$ is a *noncritical point* of Z (i.e. $Z(z_\circ) \neq 0$, the zero vector in $\mathrm{T}_{z_\circ P}$), we can construct a hypersurface $A \subset P$ transversal to Z in a neighbourhood of $z_\circ$. Restricting the flow of Z, $F : I \times P \to P$. to a smaller product $I_\circ \times A_\circ$ with $A_\circ \subset A$, we get a diffeomorphism $F_\circ : I_\circ \times A_\circ \to P_\circ$ with some neighbourhood $P_\circ$ of $z_\circ$.

By construction, $\frac{\partial}{\partial t} \underset{F_\circ}{\sim} Z_\circ$ where $Z_\circ$ is the restriction of Z to $P_\circ$, and thus $F_\circ$, $\frac{\partial}{\partial t}$ give a complete slicing of $(P_\circ, Z_\circ)$. This is the well known procedure to *straighten out* vector fields.

The previous example alerady gives a way to prove the existence of local complete slicings:

Theorem 5.1.6. *Let (P, Z) be a dynamical system, and $z_\circ \in P$ a noncritical point of Z. Let (M, α, X) be a solution of the slicing equation for Z, with $z_\circ = \alpha(x_\circ)$, and such that α is an immersion at $x_\circ$.*

There exists an open neighbourhood $M_\circ$ of $x_\circ$, an open neighbourhood $A_\circ$ of 0 in $\mathbb{R}^k$ (where $k = \dim P - \dim M$), and a diffeomorphism $\overline{\alpha} : M_\circ \times A_\circ \to P_\circ$, with an open neighbourhood $P_\circ$ of $z_\circ$, such that $\overline{\alpha}$ is a complete slicing for $Z|_{P_\circ}$, and $\overline{\alpha}(\cdot, 0) = \alpha|_{M_\circ}$.

Proof. Since every immersion is a local embedding, and the result is local, we can assume that M is a regular submanifold of P and take α as the inclusion. Considering coordinates $(z^1, \ldots, z^m, z^{m+1}, \ldots, z^p)$ in P adapted to M, we can describe the submanifold as $\{z^{m+1} = \cdots = z_p = 0\}$.

By hypothesis, Z is tangent to M, and straightening its restriction $Z_\circ$ in M, we can further take the $(z^1, \ldots, z^m)$ such that $Z = \frac{\partial}{\partial z^1}$. In a small product $M_\circ \times A_\circ$ we define in coordinates the map $\overline{\alpha}(x; s^1, \ldots, s^k) = (z^1(x), \ldots, z^m(x), s^1, \ldots, s^k)$, which is a diffeomorphism between a neighbourhood of $(x_\circ, 0)$ and a neighbourhood $P_\circ$ of $z_\circ$ in P. Z is clearly tangent to the submanifolds $\alpha_s(M_\circ)$, and $\overline{\alpha}$ is a complete slicing of $(P_\circ, Z|_{P_\circ})$. $\square$

We can extend the notion of a conserved quantity by considering maps $F : P \to A$ between manifolds instead of functions. Thus, F will be a (generalized) constant of motion if $F \circ \gamma$ is constant for the trajectories γ of (P, Z). There is an obvious relationship between complete slicings and constants of motion:

Proposition 5.1.7. *Let (P, Z) be a dynamical system, $\overline{\alpha} : M \times A \to P$ a diffeomorphism and consider the natural projection $pr_2 : M \times A \to A$. $\overline{\alpha}$ is a complete slicing for Z if and only if $F = pr_2 \circ \overline{\alpha}^{-1} : P \to A$ is a constant of motion for Z.*

Proof. It is clear. □

The constants of motion in the proposition satisfy a strong regularity condition, as their preimages are all diffeomorphic to M via $\alpha_c : M \to F^{-1}(c)$ and together give a global diffeomorphism $\overline{\alpha} : M \times A \to P$. We have already encountered these type of constants throughout the work.

Example 5.1.8. Consider the simple harmonic oscillator with phase space $\mathbb{R}^2$ and Hamiltonian $H = \frac{1}{2}p^2 + \frac{1}{2}q^2$. The Hamiltonian field defining the dynamics is

$$Z_H = p\frac{\partial}{\partial q} - q\frac{\partial}{\partial p} = -\frac{\partial}{\partial \theta},$$

whose integral curves are the equilibrium at the origin and concentric circumferences traversed clockwise, where the second expression for the field is written in polar coordinates (r, θ). We exclude the origin, so that phase space is now $\mathbb{R}^2 \setminus \{0\}$. The Hamiltonian is a constant of the motion whose level sets are circumferences of different radii, $H^{-1}(E) = \{p^2 + q^2 = 2E\} \cong \mathbb{S}^1$. Consider the diffeomorphism

$$\overline{\alpha} : \mathbb{S}^1 \times \mathbb{R}^+ \to \mathbb{R}^2 \setminus \{0\}, \quad (u, E) \mapsto \sqrt{2E}\, u$$

where we see $\mathbb{S}^1$ as being embedded in $\mathbb{R}^2$. The $\alpha_E = \alpha(\cdot, E)$ give the level sets of H. Taking the natural coordinates $\theta \mod 2\pi$ in $\mathbb{S}^1$, we consider the field $X = -\dfrac{\partial}{\partial \theta}$ in $\mathbb{S}^1$, which gives a complete slicing $(\mathbb{S}^1, \overline{X} = X \times \mathrm{Id}_{\mathbb{R}^+}, \overline{\alpha})$.

5.2 Slicings of Hamiltonian Systems

The symplectic structure of phase space is ubiquitous in the classical Hamilton-Jacobi theory. In this section we investigate the actual role of the symplectic structure, by considering the Hamiltonian problem in the general framework discussed above. In order to relate both approaches, we will make use of the following lemma:

Lemma 5.2.1. *Let (P, ω, H) be a Hamiltonian dynamical system. Let $\alpha : M \to P$ be a map, and X an arbitrary vector field on M. We have the relations:*

$$\begin{aligned} {}^t(T\alpha) \circ \omega^\flat \circ T\alpha \circ X &= i_X \alpha^*(\omega), \\ {}^t(T\alpha) \circ \omega^\flat \circ Z \circ \alpha &= d\alpha^*(H). \end{aligned}$$

Proof. Let $z = \alpha(x)$, $v, w \in \mathrm{T}_x M$. We have

$$\begin{aligned} \langle [{}^t(\mathrm{T}\alpha) \circ \omega^\flat \circ \mathrm{T}\alpha](v), w\rangle &= \langle \omega^\flat \circ \mathrm{T}\alpha(v), \mathrm{T}\alpha(w)\rangle \\ &= \omega(\mathrm{T}\alpha(v), \mathrm{T}\alpha(w)) \\ &= \alpha^*(\omega)(v, w) \\ &= \langle [\alpha^*(\omega)]^\flat(v), w\rangle \end{aligned}$$

whence ${}^t(\mathrm{T}\alpha) \circ \omega^\flat \circ \mathrm{T}\alpha = [\alpha^*(\omega)]^\flat$, and the first equality follows.

Since $\omega^\flat \circ Z = dH$, we have ${}^t(\mathrm{T}\alpha) \circ \omega^\flat \circ Z \circ \alpha = {}^t(\mathrm{T}\alpha) \circ dH \circ \alpha = \alpha^*(dH) = d\alpha^*(H)$, and we have the second relation. □

Proposition 5.2.2. *If (M,α,X) is a solution of the slicing equation for (P,Z), $\mathrm{T}\alpha\circ X = Z\circ\alpha$, then*

$$i_X\alpha^*(\omega) = d\alpha^*(H). \tag{5.2}$$

Proof. Composing ${}^t(\mathrm{T}\alpha)\circ\omega^\flat$ with the slicing equation, the proposition is a direct consequence of 5.2.1. $\square$

The converse is not true, as ${}^t(\mathrm{T}\alpha)$ is not invertible in general.

Example 5.2.3. Consider the simple harmonic oscillator as before, with $H = \dfrac{1}{2}(p^2+q^2)$, $Z_H = p\dfrac{\partial}{\partial q} - q\dfrac{\partial}{\partial p}$. Take $\alpha : I\subset\mathbb{R}\to\mathbb{R}^2$ defined as $\alpha(t) = (r_\circ\cos(t), r_\circ\sin(t))$ where $r_\circ$ is a constant. The 2-form $\alpha^*(\omega)$ is identically zero, since it lies in a space of dimension one. We also have that the function $\alpha^*(H) = \dfrac{1}{2}r_\circ^2$ is a constant, whence $i_X\alpha^*(\omega) = d\alpha^*(H) = 0$ for every field X on I. Let $X = \dfrac{d}{dt}\in\mathfrak{X}(I)$, then

$$\begin{aligned}
\mathrm{T}\alpha\circ X(t) &= -r_\circ\sin(t)\frac{\partial}{\partial q}\bigg|_{\alpha(t)} + r_\circ\cos(t)\frac{\partial}{\partial p}\bigg|_{\alpha(t)}\\
Z\circ\alpha(t) &= r_\circ\sin(t)\frac{\partial}{\partial q}\bigg|_{\alpha(t)} - r_\circ\cos(t)\frac{\partial}{\partial p}\bigg|_{\alpha(t)}
\end{aligned}$$

and they do not coincide.

Condition 5.2 is already the Hamilton-Jacobi equation in 4.51. In the cotangent bundle, we looked for $\alpha = dS$ such that $d[dS^*(H))] = 0$, and by the canonical symplectic structure we had $dS^*(\omega) = -d[(dS)^*\theta] = -d^2S = 0$.

Of course, we still need to revert the implication in order to recover the result in theorem 4.3.1. We show two ways to achieve this: by imposing an isotropy condition on the image of α, and by considering an additional fiber bundle structure on the manifold P. We now study the first condition; the second one will be treated on the next section.

In analogy with the classical Hamilton-Jacobi equation, an important class of solutions of the slicing equation are those satisfying

$$\alpha^*(\omega) = 0. \tag{5.3}$$

When the image $\alpha(M)$ is locally a submanifold of P, for instance if α is a map of constant rank, the condition means that $\alpha(M)\subset P$ is an isotropic submanifold with respect to the symplectic form ω. Indeed, the vectors in $\mathrm{T}\alpha(M)$ are of the form $\mathrm{T}\alpha\cdot v$ for $v\in\mathrm{T}M$, whence $\omega(\mathrm{T}\alpha\cdot v, \mathrm{T}\alpha\cdot v) = \alpha^*(\omega)(v,w) = 0$ and $\mathrm{T}\alpha(M)\subset[\mathrm{T}\alpha(M)]^\perp$. The converse is also true. Thus, we refer to 5.3 as the *isotropy condition*.

When α is an immersion, $\dim\alpha(M) = \dim M$ and recalling 2.1.20(ii), we must have $2\dim M\leq \dim P$ as a necessary condition. Although the isotropy condition simplifies 5.2, we cannot revert the implication yet: we will need an extra assumption.

If E is a vector space and $F\subset E$ a subspace, we write $F^\circ\subset E^*$ for the *annihilator* of F, that is, the subspace $F^\circ = \{\phi\in E^*|\phi(v) = 0\ \forall v\in F\}$. We globalize the definition for submanifolds $M\subset N$ by letting $[\mathrm{T}M]^\circ = \bigcup\limits_{x\in M}[\mathrm{T}_xM]^\circ\subset\bigcup\limits_{x\in M}\mathrm{T}^*_xP$.

For the sake of clarity, it is convenient to consider product bundles of the form

$$M\times_\alpha E = \{(x,v)\in M\times E|\ \alpha(x) = \pi(v)\},\quad \mathrm{pr}_1 : M\times_\alpha E\to M$$

where $\pi : E \to P$ is a bundle over P. $M \times_\alpha E$ is known as the *pullback* of E by f. When α is an embedding, we understand $M \times_\alpha E$ as the restriction of E to the submanifold $\alpha(M)$.

Lemma 5.2.4. *If α is an embedding, then* $\ker {}^t(\mathrm{T}\alpha) = [M \times_\alpha \mathrm{T}\alpha(M)]^\circ$.

Proof. We note that ${}^t(\mathrm{T}\alpha) : M \times_\alpha \mathrm{T}^*P \to \mathrm{T}^*M$. Let $\phi \in \ker {}^t(\mathrm{T}\alpha)$. By nondegeneracy, there exists a unique $v = \omega^\sharp \circ \phi \in M \times_\alpha \mathrm{T}P$ such that $\phi = \omega^\flat \circ v$, and letting $w \in \mathrm{T}M$ we have

$$0 = \langle {}^t(\mathrm{T}\alpha) \circ \omega^\flat \circ v, w\rangle = \langle \omega^\flat \circ v, \mathrm{T}\alpha(w)\rangle = \omega(v, \mathrm{T}\alpha(w))$$

and $\phi \in [M \times_\alpha T\alpha(M)]^\circ$. The converse is trivial. □

Theorem 5.2.5. *Let (P, ω, H) be a Hamiltonian dynamical system, with Hamiltonian vector field $Z = Z_H$, and let $\alpha : M \to P$ be an embedding.*

(i) If α is a solution of the slicing equation 5.1 and satisfies the isotropy condition 5.3, then

$$d\alpha^*(H) = 0. \tag{5.4}$$

(ii) Conversely, if α satisfies this equation and the Lagrangianity condition *(i.e. the isotropy condition and* $\dim P = 2 \dim M$*, see 2.1.21) then it is a solution of the slicing equation.*

Proof. We have already proved (i).

For (ii), if $\alpha^*(dH) = {}^t(\mathrm{T}\alpha) \circ dH \circ \alpha = 0$ then $dH \circ \alpha$ takes values in $\ker {}^t(\mathrm{T}\alpha)$, which is equal to $[M \times_\alpha \mathrm{T}\alpha(M)]^\circ$ by the previous lemma. The Lagrangianity condition implies $[\mathrm{T}\alpha(M)]^\circ = \omega^\flat[\mathrm{T}\alpha(M)]$ (2.1.21), and there exists $Z_\circ \in \mathrm{T}\alpha(M) \subset \mathrm{T}P$ such that $dH \circ \alpha = \omega^\flat \circ Z_\circ \circ \alpha$. The unique field with this property is the Hamiltonian field Z, and it follows that $Z = Z_\circ$ in $\alpha(M)$. In particular, Z is tangent to $\alpha(M)$, and the tangency is equivalent to (M, α) being a solution of the slicing equation. □

The embeddings α satisfying the Lagrangianity condition which are solutions of the slicing equation will be called *Lagrangian slicings*. The isotropy condition is not sufficient for the equivalence:

Example 5.2.6. Consider the free particle with two degrees of freedom $H = \frac{1}{2}p_1^2 + \frac{1}{2}p_2^2$, and let $\alpha : I \to \mathrm{T}^*Q \equiv \mathbb{R}^4$ be defined in some real interval as $t \mapsto (0, 0, \cos(t), \sin(t))$, where the expression is in the natural coordinates in T^*Q, as usual. Let $X = \frac{d}{dt} \in \mathfrak{X}(I)$, then

$$\mathrm{T}\alpha \circ X = -\sin(t)\frac{\partial}{\partial p_1}\bigg|_{\alpha(t)} + \cos(t)\frac{\partial}{\partial p_2}\bigg|_{\alpha(t)}$$

$$Z_H \circ \alpha = \cos(t)\frac{\partial}{\partial q^1}\bigg|_{\alpha(t)} + \sin(t)\frac{\partial}{\partial q^2}\bigg|_{\alpha(t)}$$

so $Z_H \circ \alpha \neq \mathrm{T}\alpha \circ X$ but $d[\alpha^*(H)] = 0$ and $\alpha^*(\omega) = 0$.

One way to produce Lagrangian submanifolds invariant by the dynamics Z_H is with a sufficient number of special constants of the motion. Let $F : P \to \mathbb{R}^k$ be a submersion. It is easy to show that the level sets $P_c = F^{-1}(c)$ are coisotropic submanifolds if and only if the functions $F_i = \mathrm{pr}_i \circ F$ are in involution. Thus, $\frac{1}{2}\dim P$ linearly independent constants of the motion in involution give Lagrangian submanifolds P_c with Z_H tangent to them. In the context of canonical transformations, the arbitrary parameters in a complete solution of the Hamilton-Jacobi equation already satisfy these properties.

We have an analogue to the local existence theorem 5.1.6 with the Lagrangian condition.

Theorem 5.2.7. *Let (P, ω, H) be a Hamiltonian system and $z_\circ \in P$ a* noncritical *point of H (i.e. $dH(z_\circ) \neq 0$). There exists a Lagrangian slicing of $Z = Z_H$ passing through $z_\circ$, which is contained in a local complete Lagrangian slicing.*

Proof. By a generalization of Darboux's theorem, the Jacobi-Liouville theorem (see [AM78, example 4.3.4(i)]), H can be included in a set of local symplectic coordinates $(q^1, \dots, q^m, H = p_1, \dots, p_m)$ centered at $z_\circ$. But then $Z = \partial/\partial q^1$ is tangent to the Lagrangian submanifolds defined by $p_1 = c_1, \dots, p_m = c_m$, which constitute the complete slicing. ☐

5.3 Slicings of Fiber Bundles

We now consider the situation where the dynamical system (P, Z) is fibered over M. More precisely, we let $\pi : P \to M$ be a fiber bundle, which can be included in the diagram of the slicing equation as

$$\begin{tikzcd} \mathrm{T}M \arrow[r, shift right, "\mathrm{T}\alpha"'] \arrow[d, shift right] & \mathrm{T}P \arrow[l, shift right, "\mathrm{T}\pi"'] \arrow[d, shift left] \\ M \arrow[u, shift right, bend left, "X"] \arrow[r, shift left, "\alpha"] & P \arrow[l, shift left, "\pi"] \arrow[u, shift left, bend right, "Z"'] \end{tikzcd}$$

With this structure, it becomes natural to look for solutions of the slicing equation that are also sections of the projection π. That is, we consider $\alpha : M \to P$ such that $\pi \circ \alpha = \mathrm{Id}_M$. This relation implies that α is injective, and taking the tangent maps we see that it is also an immersion.

The section α determines the partial dynamics completely. If X satisfies the slicing equation $\mathrm{T}\alpha \circ X = Z \circ \alpha$, composing with $\mathrm{T}\pi$ from the left we have

$$\mathrm{T}\pi \circ \mathrm{T}\alpha \circ X = \mathrm{T}(\pi \circ \alpha) \circ X = \mathrm{Id}_M \circ X = X$$

and we have an explicit expression for the field

$$X = \mathrm{T}\pi \circ Z \circ \alpha. \tag{5.5}$$

Rewriting 5.1 in this connection, we have

Proposition 5.3.1. *A section α of $\pi : P \to M$ is a solution of the slicing equation if and only if*

$$T(\alpha \circ \pi) \circ Z \circ \alpha = Z \circ \alpha. \tag{5.6}$$

In other words, $T(\alpha \circ \pi) \circ Z$ agrees with Z on $\alpha(M)$.

Consider α a section of π. If $z = \alpha(x) \in P$, then $\mathrm{T}_z(\alpha \circ \pi) : \mathrm{T}_z P \to \mathrm{T}_z P$ is an endomorphism such that

$$\mathrm{T}_z(\alpha \circ \pi) \circ \mathrm{T}_z(\alpha \circ \pi) = \mathrm{T}_x\alpha \circ \mathrm{T}_x(\pi \circ \alpha) \circ \mathrm{T}_z\pi = \mathrm{T}_z(\alpha \circ \pi),$$

a projector in $\mathrm{T}_z P$.

By injectivity of α, we have

$$\ker \mathrm{T}_z(\alpha \circ \pi) = \ker \mathrm{T}_z\pi = V_z P$$

where $V_z P$ are the vertical vectors in $\mathrm{T}_z P$. Also, by surjectivity of π,

$$\mathrm{Im}\ \mathrm{T}_z(\alpha \circ \pi) = \mathrm{T}_z\alpha(M).$$

For a projector, we have the decomposition $\mathrm{T}_zP = \mathrm{Im}\ \mathrm{T}_z(\alpha\circ\pi) \oplus \ker \mathrm{T}_z(\alpha\circ\pi)$, and hence

$$\mathrm{T}_{\alpha(x)}P = V_{\alpha(x)}P \oplus \mathrm{T}_{\alpha(x)}\alpha(M). \tag{5.7}$$

If we have a family α_c of nonoverlapping sections covering P, which are described by a diffeomorphism $\overline{\alpha} : M\times A \to P$ with $\alpha_c = \overline{\alpha}(\cdot, c)$ as before, then $\overline{\alpha}$ defines a *horizontal subbundle* of $\mathrm{T}P$. With this we mean a subbundle $H \subset \mathrm{T}P$ which is complementary to $VP \subset \mathrm{T}P$.

H is an integrable subbundle whose integral manifolds are given by the embeddings α_c. Conversely, it can be shown that the integral manifolds of an integrable horizontal subbundle are locally the images of sections of the bundle. For $\overline{\alpha}$ to be a complete solution, Z should also be tangent to the submanifolds $\alpha_c(M)$. Thus, complete slicings for sections are equivalent to integrable horizontal subbundles such that Z is a section of the subbundle.

Finally, we can consider both the symplectic and fibered structure in the same dynamical system. Let (P,ω,H) be a Hamiltonian system that is also a fiber bundle $\pi : P\to M$. We look for sections α of π that are solutions of the slicing equation for $Z = Z_H$, and thus $X = \mathrm{T}\pi\circ Z\circ\alpha$.

Recall that 5.2 cannot be inverted, since ${}^t(\mathrm{T}\alpha)$ is not injective in general. But we know from 5.2.4 that $\ker {}^t(\mathrm{T}\alpha) = [M\times_\alpha \mathrm{T}\alpha(M)]^\circ$. Since $\pi\circ\alpha = \mathrm{Id}_M$,

$$\mathrm{T}\pi(\mathrm{T}\alpha\circ X - Z\circ\alpha) = \mathrm{T}(\pi\circ\alpha)\circ(\mathrm{T}\pi\circ Z\circ\alpha) - \mathrm{T}\pi\circ Z\circ\alpha = 0$$

and the field $\mathrm{T}\alpha\circ X - Z\circ\alpha$ takes values in VP. Consequently, in order to revert the direct implication we only need to impose injectivity of ${}^t(\mathrm{T}\alpha)\circ\omega^\flat$ in the subbundle $M\times_\alpha VP$. This is of course equivalent to the condition

$$\omega^\flat(M\times_\alpha VP)\cap[M\times_\alpha \mathrm{T}\alpha(M)]^\circ = \{0\}. \tag{5.8}$$

We recall that the the tangent spaces of the fibers P_c of π are given by $\ker \mathrm{T}\pi = VP$, the vertical vectors. If the fibers are isotropic submanifolds, by definition we have $V_zP \subset (V_zP)^\perp$, which is equivalent to $\omega^\flat(V_zP)\subset (V_zP)^\circ$.

We also know that $\mathrm{T}_zP = \mathrm{T}_z\alpha(M)\oplus V_zP$ for $z = \alpha(x)$, and every vector $u\in \mathrm{T}_zP$ is written uniquely as $u = v + w$, where $v\in \mathrm{T}_z\alpha(M)$ and $w \in V_zP$. If we take $\phi\in[\mathrm{T}_z\alpha(M)]^\circ\cap(V_zP)^\circ$, then $\phi(u) = \phi(v)+\phi(w) = 0$ by hypothesis, whence $\phi = 0$ and thus

$$\omega^\flat(V_zP)\cap(\mathrm{T}_z\alpha(M))^\circ \subset (V_zP)^\circ\cap(\mathrm{T}_z\alpha(M))^\circ = \{0\}$$

which gives 5.8. These observations lead to the following result:

Proposition 5.3.2. *If α is a section of π and the fibers are isotropic, then*

$$\omega^\flat(M\times_\alpha VP)\cap[M\times_\alpha T\alpha(M)]^\circ = \{0\}.$$

As a consequence, we get:

Theorem 5.3.3. *Let (P,ω,H) be a Hamiltonian dynamical system on a bundle $\pi : P\to M$. Let $\alpha : M\to P$ be a section of π, and define $X = T\pi\circ Z\circ\alpha$.*

If the fibers of π are isotropic, then α is a solution of the slicing equation if and only if

$$i_X\alpha^*(\omega) = d\alpha^*(H).$$

Proof. By the isotropy condition, ${}^t(\mathrm{T}\alpha)\circ\omega^\flat$ is injective when applied to vertical vectors. Therefore $\mathrm{T}\alpha\circ X - Z\circ\alpha = 0$ (which is vertical) if and only if $i_X\alpha^*(\omega) - d\alpha^*(H) = 0$. □

Combining theorems 5.2.5 and 5.3.3, we get the following corollary:

Corollary 5.3.4. *Let (P, ω, H) be a Hamiltonian dynamical system fibered over M, with isotropic fibers. Let $\alpha : M \to P$ be a section with isotropic image. α is a solution of the slicing equation if and only if*

$$d\alpha^*(H) = 0.$$

Proof. Since sections are immersions, we have that α is a local embedding.

By isotropy of $\alpha(M)$, we have $\dim P \leq 2 \dim M$. Now, since $\dim P = \dim \ker \mathrm{T}\pi + \dim \mathrm{Im}\ \mathrm{T}\pi = \dim \ker \mathrm{T}\pi + \dim(\ker \mathrm{T}\pi)^{\perp}$ we have $\dim \mathrm{Im}\ \mathrm{T}\pi = \dim(\ker \mathrm{T}\pi)^{\perp}$ and by isotropy of $VP = \ker \mathrm{T}\pi$ we deduce $\dim P \leq 2 \dim \mathrm{Im}\ \mathrm{T}\pi = 2 \dim M$. Thus, $\alpha(M)$ is locally a Lagrangian embedding.

The equivalence is then clear by 5.2.5. □

The isotropy of the fibers in the theorem is necessary:

Example 5.3.5. In the free particle with two degrees of freedom, consider the map $\pi : \mathrm{T}^*Q \equiv \mathbb{R}^4 \to \mathbb{R}$ defined as $(q^1, q^2, p_1, p_2) \mapsto q^1$, which gives a fibered structure to $\mathbb{R}^4$. Consider the local section $\alpha : (-1, 1) \subset \mathbb{R} \to \mathbb{R}^4$, $t \mapsto (t, t, t, \sqrt{1 - t^2})$ and define $X = \mathrm{T}\pi \circ Z_H \circ \alpha = t\frac{\partial}{\partial t}$. While the relation $i_X \alpha^*(\omega) = d[\alpha^*(H)]$ holds, we have

$$\begin{aligned}
\mathrm{T}\alpha \circ X(t) &= t\frac{\partial}{\partial q^1}\bigg|_{\alpha(t)} + t\frac{\partial}{\partial q^2}\bigg|_{\alpha(t)} + t\frac{\partial}{\partial p_1}\bigg|_{\alpha(t)} - \frac{t}{\sqrt{1-t^2}}\frac{\partial}{\partial p_2}\bigg|_{\alpha(t)} \\
Z \circ \alpha(t) &= t\frac{\partial}{\partial q^1}\bigg|_{\alpha(t)} + \sqrt{1-t^2}\frac{\partial}{\partial q^2}\bigg|_{\alpha(t)}
\end{aligned}$$

and α is not a solution of the slicing equation.

With the development in these last sections, we can finally recover the classical result of Hamilton-Jacobi theory. Let Q be a configuration manifold and consider the cotangent bundle T^*Q with the canonical symplectic structure $\omega = -d\theta$, where θ is the canonical 1-form. Consider the Hamiltonian system $(\mathrm{T}^*Q, \omega, H)$ with its vector bundle structure $\pi_Q : \mathrm{T}^*Q \to Q$.

If we look for sections of π_Q, i.e. differential 1-forms $\alpha : Q \to \mathrm{T}^*Q$, satisfying the slicing equation, the field carrying the partial dynamics of the system is $X = \mathrm{T}\pi_Q \circ Z_H \circ \alpha$. In the natural coordinates, it is written $X = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} \circ \alpha$, or equivalently $X = \mathcal{F}H \circ \alpha$, where $\mathcal{F}H : \mathrm{T}^*Q \to \mathrm{T}Q$ is the fiber derivative of the Hamiltonian.

The fibers of T^*Q, i.e. the cotangent spaces T^*_qQ, are isotropic submanifolds of T^*Q, as $j_q^*(\omega) = 0$ for the inclusion $j_q : \mathrm{T}^*_q \hookrightarrow \mathrm{T}^*Q$. If we further assume that the image $\alpha(Q)$ is an isotropic submanifold, then $\alpha^*(\omega) = 0$ and we can apply corollary 5.3.4. But by proposition 2.2.3 $0 = \alpha^*(\omega) = -d\alpha^*(\theta) = -d\alpha$ and α must be closed. Applying the Poincaré lemma, α is locally exact, and we can write $\alpha = dS$ for some function $S : Q \to \mathbb{R}$ defined locally. Under these hypotheses, by 5.3.4 we find that the slicing equation is equivalent to $d[dS^*(H)] = 0$, which is the same as

$$H \circ dS = E$$

for some constant E. This is precisely the classical Hamilton-Jacobi equation.

We now look back at some classical problems from our new point of view.

Example 5.3.6. In the simple harmonic oscillator with two degrees of freedom, the movement is restricted to level surfaces of the Hamiltonian $H = \frac{1}{2}([q^1]^2 + [p_1]^2 + [q^2]^2 + [p_2]^2)$, which are diffeomorphic to $\mathbb{S}^3$. For the time being, let us identify $\mathbb{S}^3$ with the complex numbers $(z, w) \in \mathbb{C}^2$ such that $|z|^2 + |w|^2 = 1$, via the map $(q^1, q^2, p_1, p_2) \mapsto (q^1 + ip_1, q^2 + ip_2)$. Consider the 2-sphere $\mathbb{S}^2$ as a subset of $\mathbb{C} \times \mathbb{R}$ defined by the pairs $(z, x) \in \mathbb{C} \times \mathbb{R}$ such that $|z|^2 + x^2 = 1$.

Let $\tilde{f} : \mathbb{C}^2 \to \mathbb{C} \times \mathbb{R}$ be a map with $\tilde{f}(z, w) = (2z \cdot \overline{w}, |z|^2 - |w|^2)$. With the standard norms in $\mathbb{C}^2 \cong \mathbb{R}^4$ and $\mathbb{C} \times \mathbb{R} \cong \mathbb{R}^3$, we have $|\tilde{f}(z, w)|^2 = (|z|^2 + |w|^2)^2 = |(z, w)|^4$ and the map restricts to $f : \mathbb{S}^3 \to \mathbb{S}^2$, which is surjective. If $(z_1, w_1), (z_2, w_2) \in \mathbb{S}^3$ are such that $f(z_1, w_1) = f(z_2, w_2)$, it is easy to show that $(z_2, w_2) = (z_1 \cdot u, w_1 \cdot u)$ for some $u \in \mathbb{S}^2 \subset \mathbb{C}$. Thus, given $p \in \mathbb{S}^2$, the preimage $f^{-1}(p)$ is diffeomorphic to $\mathbb{S}^1$. Moreover, letting $U = \mathbb{S}^2 \setminus \{(0, 0, 1)\}$, $V = \mathbb{S}^2 \setminus \{(0, 0, -1)\}$ we have the coverings

$$\begin{aligned} g &: f^{-1}(U) \to \mathbb{S}^1 \times U, \quad (z, w) \mapsto (w/|w|, f(z, w)) \\ h &: f^{-1}(V) \to \mathbb{S}^1 \times V, \quad (z, w) \mapsto (z/|z|, f(z, w)) \end{aligned}$$

with inverses

$$\begin{aligned} g^{-1} &: \mathbb{S}^1 \times U \to f^{-1}(U), \quad (u, (c, x)) \mapsto \left(\frac{c \cdot u}{\sqrt{2(1-x)}}, \sqrt{\frac{1-x}{2}} u \right) \\ h^{-1} &: \mathbb{S}^1 \times V \to f^{-1}(V), \quad (u, (c, x)) \mapsto \left(\sqrt{\frac{1+x}{2}} u, \frac{\overline{c} \cdot u}{\sqrt{2(1+x)}} \right) \end{aligned}$$

whence it follows that $f : \mathbb{S}^3 \to \mathbb{S}^2$ gives $\mathbb{S}^3$ a fibered structure with fibers diffeomorphic to $\mathbb{S}^1$. f is known as the *Hopf fibration*.

Fixing $E \in \mathbb{R}^+$, the injection

$$j_E : \mathbb{S}^3 \to \mathbb{R}^4, \quad v \mapsto \sqrt{2E} v$$

identifying $\mathbb{S}^3$ with the 3-sphere of radius $\sqrt{2E}$ in $\mathbb{R}^4$ gives a slicing of $(\mathbb{R}^4, Z_H)$ with the vector field $Y = Z_H|_{\mathbb{S}^3} \in \mathfrak{X}(\mathbb{S}^3)$, the restriction of Z_H to the sphere of radius 1.

The Hopf fibration is important for this particular problem, since it is a constant of motion of the system $(\mathbb{S}^3, Y)$, as is readily checked. Moreover, taking $\theta \mod 2\pi$ as the natural coordinates in $\mathbb{S}^1$, we observe that the field $X = \frac{\partial}{\partial \theta}$ is a slicing of $(\mathbb{S}^3, Y)$, via the maps g^{-1}, h^{-1}. We thus obtain local complete solutions of the slicing problem for Y and Z,

$$\begin{tikzcd} \mathrm{T}\mathbb{S}^1 \times U \arrow[r, "\mathrm{T}_1 f^{-1}"] \arrow[d, shift left] & \mathrm{T}\mathbb{S}^3 \arrow[d, shift left] \\ \mathbb{S}^1 \times U \arrow[r, "f^{-1}"'] \arrow[u, bend left, "\overline{X}"] & \mathbb{S}^3 \arrow[u, bend right, "Y"'] \end{tikzcd} \qquad \begin{tikzcd} \mathrm{T}\mathbb{S}^3 \times \mathbb{R}^+ \arrow[r, "\mathrm{T}_1 j"] \arrow[d, shift left] & \mathrm{T}\mathbb{R}^4 \arrow[d, shift left] \\ \mathbb{S}^3 \times \mathbb{R}^+ \arrow[r, "j"'] \arrow[u, bend left, "\overline{Y}"] & \mathbb{R}^4 \arrow[u, bend right, "Z"'] \end{tikzcd}$$

which in turn produce a local complete solution of the original slicing problem by composing the maps

$$\begin{tikzcd} \mathrm{T}\mathbb{S}^1 \times U \times \mathbb{R}^+ \arrow[r, "\mathrm{T}_1 j \circ \mathrm{T}_1 f^{-1}"] \arrow[d, shift left] & \mathrm{T}\mathbb{R}^4 \arrow[d, shift left] \\ \mathbb{S}^1 \times U \times \mathbb{R}^+ \arrow[r, "j \circ f^{-1}"'] \arrow[u, bend left, "\overline{\overline{X}}"] & \mathbb{R}^4 \arrow[u, bend right, "Z"'] \end{tikzcd} .$$

Example 5.3.7. In the two body problem, we first decouple the dynamics of the system using the center of mass and the relative coordinates, so that the slicing problem in $\mathrm{T}^*Q \cong \mathbb{R}^6 \times \mathbb{R}^6$ is reduced to two smaller slicing problems, one of which (corresponding to the center of mass) is trivially integrated. For the other subsystem (corresponding to the relative coordinates), which is diffeomorphic to $\mathbb{R}^6$, we can reduce the dimension by constructing spherical coordinates around the direction of the angular momentum. The motion takes place in the submanifold $\Pi = \{\theta = \pi/2,\, p_\theta = 0\} \subset \mathbb{R}^6$ of dimension 4. The problem is finally brought to its simplest form by using conservation of the functions H and $p_\phi = \ell$. The situation can be summed up in the following diagrams:

$$
\begin{tikzcd}
\mathrm{T}(\mathbb{R}^2 \setminus \{0\}) \arrow[r, "\mathrm{T}\alpha_{E,\ell}"] \arrow[d, shift left] & \mathrm{T}\mathbb{R}^4 \arrow[d, shift right] \\
\mathbb{R}^2 \setminus \{0\} \arrow[r, "\alpha_{E,\ell}"'] \arrow[u, bend left, "X"] & \mathbb{R}^4 \arrow[u, bend right, "Y"']
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{T}\mathbb{R}^4 \arrow[r, "\mathrm{T}\beta"] \arrow[d, shift left] & \mathrm{T}\mathbb{R}^6 \arrow[d, shift right] \\
\mathbb{R}^4 \arrow[r, "\beta"'] \arrow[u, bend left, "Y"] & \mathbb{R}^6 \arrow[u, bend right, "Z"']
\end{tikzcd}
$$

where the maps in spherical coordinates are written $\alpha(r,\phi) = (r, \phi, \pm\sqrt{2(E - \frac{k}{r}) - \frac{p_\ell^2}{r^2}}, \ell)$ and $\beta(r,\phi,p_r,p_\phi) = (r, \pi/2, \phi, p_r, 0, p_\phi)$, while the vector fields are $X = \pm\sqrt{2(E - \frac{k}{r}) - \frac{p_\phi^2}{r^2}}\frac{\partial}{\partial r} + \frac{p_\phi}{r^2}\frac{\partial}{\partial \phi}$ and Y is the restriction of Z_H to Π.

The previous examples suggest that, in practice, one must reduce the problem by successively applying the slicing equation. Consecutive slicings give rise to less complex systems, which are easier to solve. The general scheme can be summed up in the following commutative diagrams

$$
\begin{tikzcd}
\mathrm{T}M \arrow[r, "\mathrm{T}\alpha"] \arrow[d, shift left] & \mathrm{T}N \arrow[d, shift right] \\
M \arrow[r, "\alpha"'] \arrow[u, bend left, "X"] & N \arrow[u, bend right, "Y"']
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{T}N \arrow[r, "\mathrm{T}\beta"] \arrow[d, shift left] & \mathrm{T}P \arrow[d, shift right] \\
N \arrow[r, "\beta"'] \arrow[u, bend left, "Y"] & P \arrow[u, bend right, "Z"']
\end{tikzcd}
$$

where (N, Y) is a slicing of (P, Z), and (M, X) is a slicing of (N, Y). Composing the maps in the diagrams, we obtain a new slicing of (P, Z)

$$
\begin{tikzcd}
\mathrm{T}M \arrow[r, "\mathrm{T}(\beta\circ\alpha)"] \arrow[d, shift left] & \mathrm{T}P \arrow[d, shift right] \\
M \arrow[r, "\beta\circ\alpha"'] \arrow[u, bend left, "X"] & P \arrow[u, bend right, "Z"']
\end{tikzcd}
$$

If the previous slicings are complete, i.e. if we have local diffeomorphisms $\overline{\alpha}$, $\overline{\beta}$ with

$$
\begin{tikzcd}
\mathrm{T}M \times A \arrow[r, "\mathrm{T}_1\overline{\alpha}"] \arrow[d, shift left] & \mathrm{T}N \arrow[d, shift right] \\
M \times A \arrow[r, "\overline{\alpha}"'] \arrow[u, bend left, "\overline{X}"] & N \arrow[u, bend right, "Y"']
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{T}N \times B \arrow[r, "\mathrm{T}\overline{\beta}"] \arrow[d, shift left] & \mathrm{T}P \arrow[d, shift right] \\
N \times B \arrow[r, "\overline{\beta}"'] \arrow[u, bend left, "\overline{Y}"] & P \arrow[u, bend right, "Z"']
\end{tikzcd}
$$

then we also have the complete slicing

$$\begin{tikzcd}
\mathrm{T}M \times A \times B \arrow[r, "\mathrm{T}_1\overline{\beta}\circ(\mathrm{T}_1\overline{\alpha}\times \mathrm{Id}_B)"] \arrow[d, shift left] & \mathrm{T}P \arrow[d, shift right] \\
M \times A \times B \arrow[u, bend left, "\overline{X}\times \mathrm{Id}_B"] \arrow[r, "\overline{\beta}\circ(\overline{\alpha}\times \mathrm{Id}_B)"'] & P \arrow[u, bend right, "Z"']
\end{tikzcd}$$

for the dynamics of (P, Z).

Conclusions

In this work, we have elaborated the necessary theory for a rigorous mathematical formulation of analytical mechanics, writing its fundamental objects and relations in the language of differential geometry. As a particular aspect of the general theory, we have studied the Hamilton-Jacobi equation from the classical point of view, obtaining the equation in different situations, explaining ways to solve it and stating the relationship between the equations of motion and the PDE. Furthermore, we have investigated this association from a new perspective, starting from a slicing of the dynamics of a general dynamical system and studying the PDEs arising from the additional structures on the carrier manifold. More precisely, we have found the analogues to the Hamilton-Jacobi equation for Hamiltonian systems and fibered manifolds, recovering the classical Hamilton-Jacobi theory as a particular case. We have successfully applied the theory to the classical two body problem and the simple two dimensional harmonic oscillator, showing the importance of consecutive applications of the slicing equation in the problem of the integration of the motion of a system.

Future work

The next step should be to produce further examples where the theory can be effectively applied. These should show whether there are additional features of the framework left to consider. In particular, the new thoughts on the consecutive application of the slicing equation should be reflected upon and expanded.

It would also be interesting to study the generalized Hamilton-Jacobi equation in the context of mechanical systems with symmetry, where it is possible to find invariant submanifolds of the dynamics in other ways.

Bibliography

[AM78] R. Abraham and J. E. Marsden. *Foundations of mechanics.* 2nd edn., Addison-Wesley, 1978.

[Arn74] V. I. Arnold. *Mathematical methods of classical mechanics.* Springer-Verlag, 1974.

[Car67] H. Cartan. *Formes différentielles. Applications élémentaires au calcul des variations et à la théorie des courbes et des surfaces.* Hermann, 1967.

[CGM+06] J. F. Cariñena, X. Gràcia, G. Marmo, E. Martínez, M. C. Muñoz-Lecanda, and N. Román-Roy. Geometric Hamilton-Jacobi theory. *Int. J. Geom. Methods Mod. Phys.*, **3**(7):1417–1458, (2006).

[CGM+16] J. F. Cariñena, X. Gràcia, G. Marmo, E. Martínez, M. C. Muñoz-Lecanda, and N. Román-Roy. Structural aspects of Hamilton-Jacobi theory. *Int. J. Geom. Methods Mod. Phys.*, **13**((2)):1650017, 29, 2016.

[CH37] R. Courant and D. Hilbert. *Methods of mathematical physics.* volumes 1 and 2, Wiley, 1989 (1937).

[CS31] A.W. Conway and J.L. Synge. *The Mathematical Papers of Sir William Rowan Hamilton*, volume I. Geometrical Optics. Cambridge, 1931.

[Gra15] X. Gracia. Ampliació de models matemàtics de la física. lecture notes, UPC, `https://mat-web.upc.edu/people/xavier.gracia/ammf/files/ammfresA5.pdf`, 2015. [Online; accessed 9-January-2017].

[Jac84] C. G. J. Jacobi. *Jacobi's lectures on dynamics.* ed. A. Clebsch (Hindustan Book Agency, New Delhi), 2nd rev. edn. edition, 2009 (1884).

[JS98] J.V. José and E.J. Saletan. *Classical dynamics: a contemporary approach.* Cambridge, 1998.

[Lan99] S. Lang. *Fundamentals of differential geometry.* Springer-Verlag, 1999.

[Lee47] Hwa-Chung Lee. The universal integral invariants of hamiltonian systems and application to the theory of canonical transformations. In *Proc. Roy. Soc. Edinburgh. Sect. A.*, volume **62**, 1947. 237-246.

[Lee03] J. M. Lee. *Introduction to smooth manifolds.* Springer-Verlag, 2003.

[LL77] L.D. Landau and E.M. Lifshitz. *Course of theoretical physics.* volumes 1,2 and 3, 2nd eng. ed., 3rd eng. ed. and 3rd eng. ed., Pergamon, 1969, 1971, 1977.

[Spi65] M. Spivak. *Calculus on manifolds.* Addison-Wesley, 1965.