Geometric and numerical analysis of nonholonomic systems

Alexandre Anahory Simoes

PhD thesis supervised by:
Juan Carlos Marrero and David Martín de Diego

Instituto de Ciencias Matemáticas
Universidad Autónoma de Madrid
2021

This page was intentionally left empty.

Abstract

Geometric mechanics is a fairly recent field of mathematics lying in the intersection of at least four different scientific fields: differential geometry, physics, numerical analysis and dynamical systems. Its starting point is to shed light on the underlying geometry behind mechanics and use it to obtain new results which frequently reach a variety of different mathematical fields. One of the practical applications that was made possible by using geometric techniques was the ability to construct *variational integrators*, which are numerical methods reproducing the geometry of the original mechanical system such as symplecticity, conservation of momentum and energy. These methods are often computationally cheaper than standard ones while demonstrating an adequate qualitative behaviour even at low order.

However, not all mechanical systems may be approximated using variational integrators. Nonholonomic mechanics is one of such cases, where we lack a variational principle, symplecticity and conservation of momentum, in general. Hence, the investigation of the geometric structure of nonholonomic mechanics must be carried out having into account its non-symplectic and non-variational nature.

In this thesis, we will deduce new geometric and analytical properties of nonholonomic systems which hopefully will provide a new insight to the subject. Our main definition, which we will meet across all sections, is the *nonholonomic exponential map*. This map is a generalization of the well-known Riemannian exponential map and we will see that it plays a role in the description of nonholonomic trajectories as well as on the applications to numerical analysis. After introducing this new object, the thesis may be divided into two parts. In the first part, we take advantage of the nonholonomic exponential map to present new geometric properties of mechanical nonholonomic systems such as the existence of a constrained Riemannian manifold containing radial nonholonomic trajectories with fixed starting point and

on which they are geodesics. This is a new and surprising result because it opens the possibility of applying variational techniques to nonholonomic dynamics, which is commonly seen to be non-variational in nature. Also, introduce the notion of a nonholonomic Jacobi field and provide a nonholonomic Jacobi equation. In the second part, which is more applied, we use the nonholonomic exponential map to characterize the exact discrete trajectory of nonholonomic systems. Then we propose a numerical method which is able to generate the exact trajectory. On the last chapter, we discuss contact systems and apply the nonholonomic exponential map to construct an exact discrete Lagrangian function for these systems.

Resumen

La mecánica geométrica es un campo de trabajo bastante reciente de las matemáticas que se encuentra en la intersección de al menos cuatro campos científicos diferentes: geometría diferencial, física, análisis numérico y sistemas dinámicos. Su punto de partida es arrojar luz sobre la geometría subyacente a la mecánica y utilizarla para obtener nuevos resultados que, con frecuencia, llegan a diversos campos matemáticos. Una de las aplicaciones prácticas que se hizo posible mediante el uso de técnicas geométricas fue la capacidad de construir *integradores variacionales*, que son métodos numéricos que reproducen la geometría del sistema mecánico original como la simplecticidad y la conservación del momento y de la energía. Estos métodos son a menudo más baratos computacionalmente que los estándar, a la vez que demuestran un comportamiento cualitativo adecuado, incluso a bajo orden.

Sin embargo, no todos los sistemas mecánicos pueden aproximarse mediante integradores variacionales. La mecánica no holónoma es uno de esos casos, en los que carecemos de un principio variacional, de simplecticidad y de la conservación del momento en general. Por lo tanto, la investigación de la estructura geométrica de la mecánica no holónoma debe realizarse teniendo en cuenta su naturaleza no simpléctica y no variacional.

En esta tesis, deduciremos nuevas propiedades geométricas y analíticas de los sistemas no holónomos que esperamos proporcionen una nueva visión para tratar los mismos. Nuestra definición principal, que encontraremos en todas las secciones, es la de *aplicación exponencial no holónoma*. Esta aplicación es una generalización de la conocida aplicación exponencial riemanniana y veremos que desempeña un papel en la descripción de trayectorias no holónomas, así como en aplicaciones al análisis numérico. Tras introducir este nuevo objeto, la tesis puede dividirse en dos partes. En la primera parte, usamos la aplicación exponencial no holónoma para presentar nuevas propiedades geométricas de los sistemas mecánicos no holónomos, como la existencia de

una variedad riemanniana restringida que contiene a las trayectorias radiales no holónomas con punto de partida fijo y en la que las mismas son geodesicas. Se trata de un resultado nuevo y sorprendente porque abre la posibilidad de aplicar técnicas variacionales a la dinámica no holónoma, que comúnmente se considera no variacional por naturaleza. Además, damos una nueva definición de campos de Jacobi no holónomos y encontramos una ecuación de Jacobi no holónoma. En la segunda parte, más aplicada, utilizamos la aplicación exponencial no holónoma para caracterizar la trayectoria discreta exacta de los sistemas no holónomos. A continuación, proponemos un método numérico capaz de generar la trayectoria exacta.En el último capítulo, discutimos los sistemas de contacto y empleamos la aplicación exponencial no holónoma para construir una función Lagrangiana discreta exacta para sistemas de contacto discretos.

Acknowledgements

This section is devoted to all of those who joined me during the elaboration of the thesis. An acknowledgement is entirely deserved by the people who supported me or led me here today writing the last words of my thesis, but also to those who every day work make the life of a researcher easier. After all, no deed is accomplished by one person alone.

My first acknowledgement is to my family. My parents who were always supporting and never doubted the path I have chosen though it is not free of obstacles or *saudades*. Then, to my grandmother who believed with the certainty that only grandparents are capable of, that I would surely succeed. My aunt, sisters and brother-in-law for supporting me and visiting me whenever they could and my two nephews who have the strength and the talent to change the world and inspire me to do at least half of what they will eventually do.

If the family was the background support to enroll the PhD at ICMAT, my fellow colleagues were the daily basis support. It was a pleasure to share this journey with them. I thank them for all coffee breaks, lunches and also many fruitful discussions. Among them, I must give a special thank to my office mates Rodrigo Sato, Daniel Lear, Víctor Jiménez and Manuel Laínz. But I can not also forget lunch time with many others such as Jared, Adrián, Stefano Maró, Mari Ángeles, Dahyana, María Barbero, Ángel Castro, Cristina Sardón, among many other who occasionally passed by. Thank you all for making the task much more easy to be carried out.

I wish to thank all those incredible researchers at ICMAT who have shared their knowledge with me. A special thank is due to Manuel de León, with whom we collaborated and who is always ready to introduce new students in the research group and share all his experience and expertise, and to Leonardo Colombo, who generously helped me whenever he thought necessary revealing the traces of the great advisor he already is.

I am also grateful to ICMAT and CSIC, both at the institutional level as well as to the people who make that ICMAT functions every single day as Miguel Anchuelo, Esther Ruiz, Esther Fuentes, Laura Rojas, Teresa, María Jesus, Juan Carlos, Carlos, among many others.

Another special thanks is due to the Geometric, Mechanics and Control (GMC) network, which is responsible for organizing annual conferences, schools and workshops every year. These are great opportunities for young PhD students as myself to network with other people working on the related subjects. Within the network, I want to give a special thank to Miguel Muñoz-Lecanda, whose help turned out to be essential to be able to finish my thesis, and to fellow PhD colleagues who carried their doctoral studies slightly at the same time as myself such as Iván, Jordi and Marcin. And, last but not the least, a special thank to Edith Padrón for inviting me twice to La Laguna. Surely, the two research stays were by far the most productive weeks during my thesis.

I would also like to mention three guests we had the pleasure to welcome at ICMAT and whose visits were highly valuable to the development of this thesis: Luís García-Naranjo, Knut Hüper and Fátima Silva Leite.

Also, thanks to David Iglesias for reading this manuscript and for his many useful suggestions; and thanks to Rafael Orivé for accepting to be the tutor of my thesis.

I would like to thank the other two minds behind this project: my thesis directors David Martín de Diego and Juan Carlos Marrero. Both are widely recognized to be among the best researchers on geometric mechanics and so it was an honour to have the opportunity to be directed by them. I must thank them for initiating me in the research career, for kindly sharing their knowledge and ideas with me and for giving me full access to the GMC network. I finish my Ph.D. better prepared to eventually become a professional researcher thanks to them.

Also, a special thank to Juan Carlos for receiving me in his office and his home at La Laguna. I must say that the discussions with Juan Carlos are always inspiring. And to David, I am thankful for his attentive daily presence in the office right on the next door. Thank you for all the hours spent on discussing maths, for giving me the opportunity to attend several conferences and to network with fellow researchers, in so many different places.

At last, I am eternally grateful to Nuria, for supporting me unconditionally in this long journey. Without her help, the obstacles in the way of a young researcher, could be unbearable to face. If I overcome them all, she will have a huge share of responsibility. And also, I am thankful for all the help given by Nuria's family, both from Madrid and Plasencia, who always made me fell at home. And thank to Residencia Felipenses: without them we would not be able to find a place at Madrid.

I acknowledge the support by the *Fundação para a Ciência e Tecnologia* (Portugal) under the research fellowship SFRH/BD/129882/2017 partially funded by the European Union through the European Social Fund. I am also thankful to the Spanish Ministry of Science and Innovation under the grants PID2019-106715GB-C21, MTM2016-76702-P, "Severo Ochoa Programme for Centres of Excellence in R&D" (SEV-2015-0554) and to the Spanish National Research Council, through the "Ayuda extraordinaria a Centros de Excelencia Severo Ochoa" (20205CEX001), which financed several travel expenses and my accommodation at the ICMAT.

FCT Fundação para a Ciência e a Tecnologia

Notation

All the manifolds and maps in the text are smooth except we say otherwise. Einstein summation convention is used.

We use intensively the following notation without explaining it from now further:

$C^\infty(Q)$ - the set of smooth functions on the smooth manifold Q.

$\mathfrak{X}(Q)$ - the set of smooth vector fields on the manifold Q.

$\Gamma(\mathcal{D})$ - the set of smooth sections of a vector bundle, where $\mathcal{D}$ is the total space of the bundle.

$\Omega^k(Q)$ - the set of smooth k-forms on the manifold Q.

Contents

Concern for man and his fate must always form the chief interest of all technical endeavours. Never forget this in the midst of your diagrams and equations.

by Albert Einstein

Simplicity is the highest goal, achievable when you have overcome all difficulties. After one has played a vast quantity of notes and more notes, it is simplicity that emerges as the crowning reward of art.

by Frédéric Chopin

Mechanics is the paradise of the mathematical sciences because by means of it one comes to the fruits of mathematics.

by Leonardo Da Vinci

Chapter 1

Introduction

This thesis is intended to be understandable by any graduate student with some background in differential geometry and some basic notions from mechanics. We will assume that the reader is familiar with smooth manifolds and has a general knowledge of other standard topics in mathematics such as Lie groups, ordinary differential equation (ODE) theory and dynamical systems.

This is a thesis on geometric mechanics, a subject of applied mathematics that uses techniques of differential geometry to study mechanical systems. Of course, the realm of Mechanics is extremely wide and includes both classical particle mechanics as well as continuum mechanics and field theories. It deals with systems from a classical to a relativistic and a quantum perspectives.

Mechanics has been one of the main branches of Physics for at least two millennia. Indeed, one could date the first systematic descriptions of body and particle motions, composition and properties back to the age of Ancient Greece, by the hand of philosophers such as Aristotle or Archimedes. However, a true scientific revolution had to wait until the XVI and XVII centuries through the works of Nicolaus Copernicus, Galileo Galilei, Johannes Kepler and Sir Isaac Newton, who may be considered the founders of modern mechanics. In fact, almost every first undergraduate course in Mechanics begins with Newton's second law, relating the force F with the mass m and the acceleration a of the system:

$$F = ma,$$

which encodes the equations of motion for any classical system in an inertial reference frame.

It was not until the XVIII century, that Leonhard Euler and Joseph-Louis Lagrange introduced one of the core concepts of this thesis: Euler-Lagrange equations. They appeared while the authors were studying the tautochrone problem, that is, determining the curve for which the time taken by a particle under the influence of gravity to travel to a fixed final point is independent of the initial point. More generally, Euler-Lagrange equations constitute the main ingredient of the fundamental theorem of Calculus of Variations, a branch of mathematics developed precisely by Euler and Lagrange, and according to which a twice differentiable curve $x : I \to M$ in a manifold M is a critical point of the functional

$$\mathcal{S}(x) = \int_{t_0}^{t_1} F(x(t), \dot{x}(t))\; dt$$

if and only if the curve x satisfies the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) - \frac{\partial F}{\partial x} = 0.$$

These foundational works on Calculus of Variations, led to the development of Lagrangian mechanics based on the *Principle of Least Action.* In this framework, there is a functional $\mathcal{S}$ called the *action* that is minimized by the mechanical trajectories. Usually, the action is expressed as an integral of a function L, called the Lagrangian function, depending on position q and velocities $\dot{q}$ of the system which is given by the *kinetic energy* K minus the *potential energy* function V, i.e., $L = K - V$. Hence, Euler-Lagrange equations become necessary and sufficient conditions to find the motion of mechanical systems, that is, they become the equations of motion, thereby replacing Newton's second law.

Almost a century later, Sir William Rowan Hamilton introduced an alternative description of mechanics that, nowadays, carries his own name: Hamiltonian mechanics. Starting with nothing more than a function, called the Hamiltonian function, depending on positions q and momenta p of the system and typically given by $H = K + V$, one extracts the Hamiltonian equations

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} \\ \dot{p} = -\frac{\partial H}{\partial q}. \end{cases}$$

There are two immediate differences between Euler-Lagrange and Hamiltonian equations: the former are a system of n second order differential equations, while the latter is a system of $2n$ first order differential equations; the former depend on positions and velocities of the system while the latter depend on positions and momenta of the system. In spite of their differences, they are equivalent whenever we have a smooth correspondence between velocities and momenta, allowing us to pass from one framework to the other. Hence, Hamiltonian equations might also replace Newton's second law as equations of motion. We emphasize that this is the usual situation in classical Newtonian mechanics, where the momenta are simply given by

$$p_i = m\dot{q}^i.$$

At this point, Mechanics was already so close to differential geometry that it could almost grasp it. Indeed, the solutions of Hamiltonian equations are just the integral curves of a vector field X_H in the phase space, that is, the joint space of positions and momenta on which H depends. If we see it as a $2n$-dimensional manifold with coordinates $(q^1, ..., q^n, p_1, ..., p_n)$, then after introducing a *symplectic form* on this space, given by

$$\omega_Q = dq^i \wedge dp_i,$$

we may characterize X_H as the unique vector field satisfying the geometric equation

$$i_{X_H}\omega_Q = dH.$$

This overlap between mechanics and differential geometry opened the door to the introduction of several geometric techniques to solve problems in mechanics, but also the other way around: mechanics provided inspiration to develop new theories in differential geometry like symplectic geometry or Poisson geometry.

The pioneer of geometric mechanics may be considered to be Henri Poincaré (see the introduction in [AM78]). Though he lacked some tools from differential geometry that had not been introduced by then, he realized at the beginning of the XX century that some qualitative questions on the dynamics of mechanical systems could not be answered without a global geometric description, as those related with stability. The intrinsic global geometric approach suggested by Poincaré was made possible through the use of exterior calculus, introduced by Elie Cartan. Some prominent mathematicians which successfully applied differential geometric and topological techniques obtaining new insights into mechanics were George Birkhoff, Andrei Kolmogorov, Vladimir Arnold, Jürgen Moser or Jerrold Marsden, among many others.

More or less simultaneously to the works of Poincaré, Albert Einstein introduced the theory of general relativity, which casts some light on the relation between mechanics and (semi-)Riemannian geometry.

In the recent few years and mainly from the second half of the XX century, built on the pillars of modern day differential geometry, the subject of Geometric Mechanics arose as an independent branch of mathematics in the intersection of differential geometry, physics, analysis and numerical analysis.

Next, we are going to briefly introduce the main topics on which we develop our research.

1.1 Nonholonomic mechanics

Nonholonomic systems are, so to speak, mechanical systems with a prescribed restriction on the velocities. The constraints on velocities might arise due to different physical causes and may appear under different forms: a restriction in the directions or in the norm; may be time-dependent or not; may vary with position of the system on the configuration manifold or even may vary according to the state on the phase space.

For example, in $\mathbb{R}^n$, a velocity constraint could be forcing a particle's velocity to take the values of a curve $f:\mathbb{R}\to\mathbb{R}^n$, in the sense that $\dot{q}(t)=f(t)$, with q denoting the position of the particle. Or else, we could constrain the velocity to remain constant $\|\dot{q}(t)\|=1$.

In this thesis, we are interested in linear nonholonomic constraints. As the name suggests, these are locally given by an expression which is linear on the velocities, though it might change smoothly from point to point. In the n-dimensional euclidean space $\mathbb{R}^n$, a curve q satisfies a linear nonholonomic constraint if and only if it satisfies a system of k equations $k<n$:

$$\sum_{i=0}^{n} \mu_i^a(q)\dot{q}^i = 0, \quad a\in\{1,...,k\}. \tag{1.1.1}$$

Geometrically, an expression like this defines a *distribution*, which is the smooth assignment to every point on a manifold of a subspace of the tangent space at that point. At a point q, this subspace satisfies a system of equations of the type of (1.1.1). So, when we think about linear nonholonomic constraints we are equivalently thinking about a distribution.

Historically, the mathematical description of nonholonomic dynamics has been the source of some confusion, motivated by an erroneous derivation of the equations of motion for systems under nonholonomic constraints due to E. Lindelof [Lin95] later corrected by S.A. Chaplygin [Cha54]. The main reason behind this confusion, might be the fact that nonholonomic trajectories do not satisfy a variational principle such as the least action principle, in contrast with the unconstrained case. This was critically against the generalized philosophic belief that mechanical trajectories should follow a preferred path among all possible paths, as if nature was omniscient and knew how to distinguish the optimal possibility.

The term "nonholonomic" was coined by H. Hertz in 1894 [Her56], who was the first to realize that nonholonomic mechanics did not satisfied the same variational principles as unconstrained mechanics. Nowadays, we know due to Otto Holder that a slight variation of the least action principle is enough to obtain the equations of motion for nonholonomic mechanics. This principle is *Lagrange-d'Alembert principle* and states that the nonholonomic trajectories subjected to a distribution $\mathcal{D}$ are those that are critical values for the action $\mathcal{S}$ among any variation of paths satisfying

$$\delta\varphi = \left.\frac{d}{ds}\right|_{s=0} \varphi(t,s) \in \mathcal{D}.$$

(See [Blo15; Cor02; NF72; CS98; BM05; Blo+96b; BMZ05], also [Mar96] for a discussion on the validity of the Lagrange-d'Alembert principle and [GG08] or [BS08] for a general discussion on variational calculus with constraints). Then, a curve $q : I \to Q$ on a manifold Q is a nonholonomic trajectory if and only if it satisfies the *Lagrange-d'Alembert equations*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \lambda_a \mu_i^a,$$

where μ_i^a are the functions defining the distribution in the same way as in (1.1.1) and λ_a are Lagrange multipliers' to be computed with the additional equations

$$\mu_i^a(q)\dot{q}^i = 0,$$

which translate the fact that the trajectory must satisfy the constraints.

We will have the opportunity to meet some basic examples of nonholonomic systems with linear constraints such as the nonholonomic particle, the vertical rolling disk and the Chaplygin sleigh, which exhibit all the unusual characteristics of nonholonomic dynamics.

Besides failing to possess a variational derivation, nonholonomic dynamics also has another important difference with respect to unconstrained mechanics: the flow on phase space is not symplectic. This fact makes nonholonomic dynamics qualitatively very different from what we find in unconstrained mechanical systems: non-preservation of phase space volume in general (see [FGNM15] and the references there in and also [ZB03]), existence of attractors, etc. Nonetheless, when the constraints are linear on velocities, energy is still preserved along the flow.

1.2 Discrete variational mechanics

The second major topic on which we investigate is discrete mechanics and its applications to the development of *geometric integrators* for dynamics. These are numerical methods approximating the solutions of the equations of motion that preserve the geometric properties of mechanical dynamical systems (see [SSC94; HLW10; BC16] and, in particular, for discrete mechanics and variational integrators [MW01]). Hence, this is a new field born out of the marriage between differential geometry, mechanics and numerical analysis.

The ultimate goal of discrete variational theory in mechanics is the construction of the so-called geometric integrators. An integrator or numerical method for a system of differential equations of the form

$$\dot{x} = f(x),$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is smooth, is a useful tool in applied sciences because to find an explicit solution to a differential equation is difficult or even impossible. At the same time, in applications to engineering and other applied sciences, it is often enough to find an approximate solution of the differential equation. In order to do that, we formulate and study numerical methods to find approximate solutions.

The simplest of the numerical methods is the *Euler method*, introduced by Euler in 1768 and it is given by the equation

$$x_{n+1} = x_n + hf(x_n), \tag{1.2.1}$$

where $h > 0$ is called the *step size*. From a given initial value $x_0 \in \mathbb{R}^n$, we compute a sequence of points $x_1, x_2, ..., x_n, ...$ that *approximates* the true solution $x(t)$ in the sense that

$$x_1 \approx x(h),\ x_2 \approx x(2h), ...,\ x_n \approx x(nh), ...$$

up to some error order. Moreover, equation (1.2.1) defines a map $\Phi_h : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$\Phi_h(x_n) = x_{n+1}$$

called the *discrete flow*.

However, not all numerical methods satisfy the same geometrical properties as the mechanical systems they should approximate, such as conservation of energy, symplecticity, symmetries and existence of conserved quantities, etc. Geometric integrators are designed to reproduce these qualitative characteristics and, while doing this, they often have a better performance at a lower computational cost than standard algorithms. This property becomes essential in long-time simulations.

As far as we know, the theory of discrete variational mechanics has its roots in the 60's in the optimal control literature. Then it starts to appear in the context of mechanics in the 70's and the 80's, at which point the first concepts as the discrete principle and the discrete equations are already defined, and, finally, the theory evolves to a deeper level with the works of Veselov [Ves88] and Moser and Veselov [MV91] in the context of integrable systems. The numerical implementation of discrete variational mechanics is developed in several works by Marsden and collaborators (see [MW01] and references therein).

Discrete Lagrangian theory falls into the more general class of discrete variational integrators which are also geometric integrators. Given a configuration manifold Q, the starting point of the Lagrangian formalism is the choice of a Lagrangian function on the tangent space TQ. In order to develop a discrete Lagrangian formalism, the situation is similar except that we replace the tangent bundle TQ by its discretized version: the Cartesian product $Q \times Q$. Then we consider a function $L_d^h : Q \times Q \to \mathbb{R}$ (possibly depending on the time step $h > 0$) and call it the *discrete Lagrangian function*.

Next, we introduce a discrete version of the action functional, the *discrete action map*, defined to be the map,

$$S_d(q_d) = \sum_{k=0}^{N-1} L_d^h(q_k, q_{k+1}),$$

where q_d is a sequence of points in Q given by $\{q_0, q_1, ..., q_N\}$. In analogy with the continuous version, the discrete trajectory of the discrete Lagrangian system determined by the discrete Lagrangian function L_d is a critical value of the discrete action map S_d, among all sequences of points with fixed end-points. Then, one may prove that a sequence $\{q_k\}$ is a critical point of the functional S_d if and only if it is a solution of the *discrete Euler-Lagrange equations*

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0, \quad \text{for all } k = 1, ..., N-1,$$

which are the discrete version of Euler-Lagrange equations.

The discrete flow generated by discrete Euler-Lagrange equations possess all the expected geometric properties. We will review exactly which later in Chapter 3. One of the most remarkable facts about discrete Lagrangian theory is that if one wishes to construct an accurate numerical method using discrete Lagrangian mechanics, one usually regards the value of the discrete Lagrangian on a point (q_0, q_1) as being a sufficiently good approximation of the (continuous) action, i.e.

$$L_d^h(q_0, q_1) \approx \int_0^h L(q_{0,1}(t), \dot{q}_{0,1}(t)) \; dt,$$

where $q_{0,1}(t)$ is the *unique* solution of the Euler-Lagrange equations connecting q_0 and q_1. At least for two sufficiently near points in Q, this trajectory exists. Though this is a very intuitive fact, we only find a formal proof very recently in [MDM21] (see also [MDM16]). Anyway, to make the dissertation more self-contained, we include such a proof in Chapter 4. If one takes the right hand side of the equation above to be the discrete Lagrangian function, by proceeding as above, we deduce that the corresponding discrete flow is the *exact discrete flow*, which is a sequence of points over the continuous trajectory of the Lagrangian system determined by L (cf. [MW01; PC09] for the original exposition and proofs, [MDM16] for the case of reduced systems under symmetries and [DA18; FZG21] for forced systems).

1.3 Main contributions

In the present thesis, we follow five main research lines:

1. **The nonholonomic exponential map**: The exponential map is a central concept in Riemannian geometry playing a key role in some results on the global analysis of Riemannian manifolds together with

curvature and Jacobi fields. Moreover, as we will see in Chapter 3 and later in Chapter 7, the exponential map (associated with a vector field) lies in the interplay between discrete and continuous dynamics. In this thesis, we define the nonholonomic exponential map and we show that it allows to identify the space of nonholonomic constraints with its image. This fact, is explored later in Chapters 5, 7 and 8. In order to introduce the previous notion, we will use the geometric formulation of nonholonomic dynamics in terms of second order differential equation vector fields along the constraint distribution (see [LD96]).

2. **Radial trajectories of mechanical nonholonomic systems**: within this line we follow a new approach aiming to study further geometric properties of mechanical nonholonomic systems which will possibly open the door to qualitative results on the global properties of nonholonomic dynamics. For this matter, we will consider a Riemannian manifold (Q, g) and a nonintegrable distribution $\mathcal{D}$ determining a kinetic nonholonomic system, where the Lagrangian function is given simply by the kinetic energy.

   Along Chapter 5, we intended to put forward a geometric program of introducing some important concepts of standard Riemannian geometry into the nonholonomic setting. In this direction, one may find previous literature such as [Vra28], [Syn28], [Lew98], and [GNM20]. A fundamental concept in Riemannian geometry is, without any doubt, that of a geodesic. One of its main properties is that geodesics minimize the length among curves connecting nearby points. Conversely, any curve minimizing length is necessarily a geodesic. To prove this key fact in Riemannian geometry, it is necessary to introduce different concepts and results as, for instance, the notion of geodesic flow, the exponential map, the Gauss lemma, among others (see [Lee97; O'N83; Car92]).

   However, the introduction of a nonintegrable distribution $\mathcal{D}$ on Q makes the picture become much more complex. The Lagrange-d'Alembert principle for a curve $c : I \to Q$, mapping $t \in I \mapsto c(t) \in Q$, may be shown to be equivalent to the following equation:

$$\nabla^g_{\dot{c}(t)} \dot{c}(t) \in \mathcal{D}^{\perp}_{c(t)} \ , \quad \dot{c}(t) \in \mathcal{D}_{c(t)} \tag{1.3.1}$$

   where ∇^g is the Levi-Civita connection of g and $\mathcal{D}^{\perp}$ is the g-orthogonal complement to $\mathcal{D}$. Equivalently, we can describe the nonholonomic trajectories as the geodesics of an affine connection ∇^{nh} on Q with initial condition satisfying the nonholonomic constraints. Of course, analogous to what happens with unconstrained geodesics, a curve c is a solution of (1.3.1) if and only if it is the solution of Lagrange-d'Alembert equations for the nonholonomic Lagrangian system $(L_g, \mathcal{D})$. Only in very exceptional cases the connection ∇^{nh} is the Levi-Civita connection for a Riemannian metric, in particular, this would imply that the distribution was integrable (see [Lew98]). The fact that ∇^{nh} is not the Levi-Civita connection implies in particular that its geodesics, and among them nonholonomic trajectories, are not length minimizing for the Riemannian metric g.

In Chapter 5, we present the absolutely surprising result that we can characterize radial nonholonomic geodesics, i.e., nonholonomic solutions starting from a given point $q \in Q$, as true Riemannian geodesics for a family of Riemannian metrics g_q^{nh} defined in the image $\mathcal{M}_q^{nh}$ of the nonholonomic exponential map at q. In other words, we show that radial kinetic nonholonomic trajectories are length minimizing in a specified Riemannian manifold, that is, nonholonomic trajectories minimize the functional

$$L_q^{nh}(c) = \int_0^1 \|\dot{c}\|_{g_q^{nh}} \, dt$$

among all curves $c : [0, 1] \longrightarrow \mathcal{M}_q^{nh}$ with fixed endpoints and starting at q. As a consequence, they are Riemannian geodesics. Perhaps more importantly, this result opens the door to new developments in nonholonomic mechanics using Riemannian geometry techniques: Riemannian Jacobi fields, global minimizing properties of nonholonomic trajectories, construction of variational integrators for nonholonomic mechanics, Hamiltonization or Lagrangianization of nonholonomic systems,... (see Chapter 9 on conclusions and future work). In this direction, it would be interesting to study the kinetic Lagrangianization of kinetic nonholonomic systems, a theory which is closely related with the so-called Hamiltonization of nonholonomic systems. This last topic has been widely discussed in recent years by several authors (see [BGN12; Ehl+05; FJ04; GNM20; GNM18; Jov10; Koz02; VV88; BY20; BFM09; GN10; BM07; BBM10]).

We also extend the previous result on radial kinetic nonholonomic trajectories to radial trajectories of mechanical nonholonomic systems. For that matter, we discuss a contact bundle formulation of the nonholonomic Jacobi-Maupertuis principle (see also [Koi92]) and then we show that these last trajectories are reparametrizations of trajectories of certain kinetic nonholonomic systems. This fact implies the result.

3. **Nonholonomic Jacobi fields**: Given the importance of Jacobi fields in Riemannian geometry, there has been great interest in generalizing these results to different situations, for example, to general second order differential equations (SODE's) in [CM92] using the dynamical covariant derivative and the Jacobi endomorphism associated with the SODE [MCS93] (see also [HM20] and the references therein), to semi-Riemannian geometry [O'N83], to sub-Riemannian and Finsler geometry [ABR18; BR17], to the Lie algebroid setting [CGM15], to skew-symmetric algebroids [J1́3], etc.

   However, the case of systems subjected to nonholonomic constraints has not been properly considered in the previous literature. In this thesis, we introduce the notion of a nonholonomic Jacobi field and deduce the corresponding Jacobi equation. More concretely, following an analogous approach to the Riemannian case (see Table 1.1 for more details):

   (a) We have defined nonholonomic Jacobi fields in terms of infinitesimal nonholonomic geodesic variations. So, for a fixed point q on the configuration manifold, the tangent space of $\mathcal{M}_q^{nh}$ at each of its points is generated by the nonholonomic Jacobi fields along the radial nonholonomic trajectories starting from the point q. Moreover, with the previous definition, a nonholonomic Jacobi field along a nonholonomic trajectory c is not, in general, a section of the constraint distribution D along c. This is an important difference with previous approaches to the notion of a nonholonomic Jacobi field.

   (b) We have given new results to explicitly find nonholonomic Jacobi fields.

   (c) We have characterized nonholonomic Jacobi fields as trajectories of a lifted nonholonomic system.

   (d) Finally, we have derived the nonholonomic Jacobi equation in terms of the curvature and torsion of the corresponding nonholonomic connection.

On the other hand, to preserve as much as possible the Riemannian geometric flavour, we start our study with nonholonomic systems of kinetic type, but later we extend the results to the case of mechanical nonholonomic systems, where the Lagrangian function is the difference between the kinetic energy associated with a Riemannian metric and the potential energy.

4. **Discrete nonholonomic mechanics**: In the last few decades, several authors have proposed a discrete nonholonomic Lagrangian formalism, in order to obtain integrators for nonholonomic systems preserving their intrinsic geometric properties. This would be a natural extension of discrete Lagrangian formalism to the case where nonholonomic constraints are present. However, contrary to the unconstrained case, nonholonomic systems do not generate symplectic flows, neither should their discrete counterpart generate symplectic maps. So, the loss of geometric structure comparatively to the unconstrained case has been the main obstacle in the search for an adequate discrete description of nonholonomic systems. In this thesis, we address the following open problem proposed by R.I. MacLachlan and C. Scovel (see also [MP06] for an attempt to solve it):

   > *The problem for the more general class of non-holonomic constraints is still open, as is the question of the correct analogue of symplectic integration for non-holonomically constrained Lagrangian systems.* [MS96].

   To achieve this goal, we carefully define the *exact discrete flow* of nonholonomic systems. Then, we search for a discrete "variational" principle satisfied by the exact discrete flow and, finally, we propose a discrete nonholonomic formalism generating geometric integrators for nonholonomic systems. In order to accomplish that, we will rely on a map which is omnipresent across the various chapters and sections of the thesis: the *nonholonomic exponential map*.

   The importance of nonholonomic systems appears since they model a great variety of mechanical systems in engineering and robotics (see [Blo15] and references therein). However, at the moment (as we have

<table>
<tr><th>Riemannian geometry</th><th>Kinetic nonholonomic</th></tr>
<tr><td>A vector field $W : I \to TQ$ along a geodesic $c : I \to Q$ is said to be a Jacobi field for the Riemannian manifold (Q, g) if it is the infinitesimal variation vector field of a family of geodesics</td><td>A vector field $W : I \to TQ$ along a nonholonomic trajectory $c : I \to Q$ is said to be a nonholonomic Jacobi field for the system $(L_g, \mathcal{D})$ if it is the infinitesimal variation vector field of a family of nonholonomic trajectories (see Definition 6.1.1)</td></tr>
<tr><td>Every Killing vector field W for the Riemannian metric g is a Jacobi field along any geodesic</td><td>Every Killing vector field W for the Riemannian metric g which is an infinitesimal symmetry of $\mathcal{D}$ is a nonholonomic Jacobi field for any nonholonomic solution
(see Corollary 6.2.4)</td></tr>
<tr><td>The trajectories of the Lagrangian system $L_{g^c} : TTQ \to \mathbb{R}$ are just the Jacobi fields for the Riemannian manifold (Q, g)</td><td>The trajectories of the nonholonomic system $(L_{g^c}, \mathcal{D}^c)$ are just the Jacobi fields for the nonholonomic system determined by $(L_g, \mathcal{D})$
(see Theorem 6.2.11)</td></tr>
<tr><td>W is a Jacobi field if and only if $$\frac{D^2W}{dt^2} + R(W(t), \dot{c}(t))\dot{c}(t) = 0$$ or, equivalently, $$\nabla^g_{\dot{c}}\nabla^g_{\dot{c}}W + R(W, \dot{c})\dot{c} = 0$$</td><td>W is a nonholonomic Jacobi field if and only if $$\begin{aligned}&\nabla^{nh}_{\dot{c}}\nabla^{nh}_{\dot{c}}W + \nabla^{nh}_{\dot{c}}T^{nh}(W, \dot{c})\\ &+R^{nh}(W, \dot{c})\dot{c} = 0, \quad \dot{W}(t) \in \mathcal{D}^c.\end{aligned}$$ (see Theorem 6.2.21)</td></tr>
</table>

Table 1.1: Comparative notions between Jacobi fields for Riemannian geometry and kinetic nonholonomic systems.

mentioned before), there is no consensus in the scientific community on the best numerical methods to integrate a nonholonomic system among several existing possibilities that were inspired in the geometry of nonholonomic systems and suitable discretizations of Lagrange-d'Alembert principle (cf. [MV19]). We think that one of the reasons for these plethora of so different methods (see [CM01; MP06; FID08; BZ15; FBO12; Cel+19; Igl+08], among others) can be related with the difficulty to find an exact discrete version of the nonholonomic mechanics as it happens in the case of Lagrangian mechanics. This is precisely our main contribution to the subject.

5. **The geometry of discrete contact systems**: Contact Hamiltonian and Lagrangian systems have deserved a lot of attention in recent years [Bra17; Bra18] or [LLV19b]. One of the most relevant features of contact dynamics is the absence of conservative properties contrarily to the conservative character of the energy in symplectic dynamics; indeed, we have a dissipative behaviour. This fact suggests that contact geometry may be the appropriate framework to model many physical and mathematical problems with dissipation we find in thermodynamics, statistical physics, quantum mechanics (see [CCM18]), gravity or control theory, among many others. Consequently, it becomes an important necessity to develop numerical methods adapted to the contact setting for applications in the above mentioned subjects. The idea is to develop geometric integrators, that is, numerical methods for differential equations which preserve geometric properties like contact structure, symmetries, configuration space... This preservation of structural properties is often desirable to achieve correct qualitative behaviour and long time stability [HLW10; SSC94; BC16].

   As far as we know, the first attempt to develop geometric integrators for the contact case is achieved in the paper [VBS19] (see also [Bra+20]), where the authors present geometric numerical integrators for contact flows that stem from an heuristic discretization of Herglotz variational principle.

   Our contribution in this thesis is to go further in the discrete description of contact dynamics, by identifying the discrete geometric structures that replace the usual contact structure on the continuous side. Instead of deriving the discrete Herglotz equations by an heuristic argument, they are directly obtained from a clear discrete variational principle. In addition, to develop the discrete algorithm we use the natural discretization $Q \times Q \times \mathbb{R}$, which preserves all the contact geometry flavour.

Another relevant point is the discussion of the existence of an exact discrete Lagrangian function [MW01; PC09] for the contact case. We define the contact exponential map as an application of the results about the nonholonomic exponential map and we prove the existence of the exact discrete Lagrangian function. This construction is essential to develop a complete theory of variational error analysis for contact Lagrangian systems.

Finally, we consider a discrete version of the infinitesimal symmetries discussed in [LV20; Gas+20], jointly with the corresponding dissipated quantities.

The original results presented along the thesis are contained in the following scientific articles:

1. *Exact discrete Lagrangian in nonholonomic mechanics* [AMM20a], (submitted to *Numerische Mathematik*).

2. *Nonholonomic Jacobi fields*, [AMM20b] (submitted to *J. Phys. A: Math. Theoret.*).

3. *Kinetic nonholonomic radial trajectories are Riemannian geodesics!*, [AMM20c] (accepted for its publication in *Analysis and Mathematical Physics*)

4. *Contact bundle formulation of nonholonomic Maupertuis-Jacobi principle and length minimizing property of nonholonomic dynamics*, [AMM21].

5. *On the geometry of discrete contact mechanics*, [Ana+21] (published on *Journal of Nonlinear Science*).

1.4 Outline of the thesis

The thesis is divided in two parts. The first part is composed by Chapters 2 and 3, comprising what we could call the "background material". Chapter 2 is devoted to review necessary topics of differential geometry on which we will rely on during all the exposition of the thesis. The list of required machinery includes (semi-)Riemannian geometry, symplectic geometry, the geometry of the tangent bundle (including lifts, the vertical endomorphism, SODE vector fields), distributions and Lie group actions. Chapter 3 is devoted to review the fundamental results of Geometric Mechanics: Lagrangian and Hamiltonian mechanics, mechanics under the action of external forces, a brief review of discrete Lagrangian mechanics and, last but not the least, a review of nonholonomic mechanics and a final section where we selected a prior version of discrete nonholonomic mechanics representing what has been done in this field.

The reader may skip the first part if he/she is already familiar with these concepts without loosing any relevant information to understand the second part.

The second part of the thesis is dedicated to the exposition of the original results contained in the scientific articles mentioned before. It starts with Chapter 4, where we introduce the nonholonomic exponential map. To this end, we also discuss the exponential map for general second order differential equation (SODE) vector fields and present a proof that this map is a diffeomorphism, at least restricted to a local neighbourhood. In the following two chapters, we study nonholonomic systems determined by a mechanical Lagrangian function, finding results with an intense Riemannian flavour such as the characterization of radial kinetic nonholonomic trajectories as geodesics of a particular Riemannian manifold in Chapter 5 or the definition of a nonholonomic Jacobi field and the associated nonholonomic Jacobi equation in Chapter 6. In Chapter 7, we present a class of integrators for nonholonomic mechanics possessing a new feature distinguishing it from the previously proposed integrators: for the correct choice of geometric objects, the new integrator gives rise to the exact discrete flow. Finally, in Chapter 8 we introduce a geometric discretization of Herglotz principle from where we deduce the discrete Herglotz equations. We also examine what happens with discrete symmetries of the contact systems. To finish the chapter, we define the contact exponential map and use the nonholonomic exponential map to prove that the former is a local diffeomorphism. With this result we are able to define the exact discrete contact Lagrangian function.

In the conclusions chapter, we unveil some possible future directions from where to sharp our results. At the end, we include an appendix containing some technical results which we thought the reader might either be unfamiliar with or might want to recall, while we found that its placement in the main body broke the natural reading flow of the thesis.

Chapter 2

Review of differential geometry

This chapter is devoted to review the fundamental concepts, properties and results from differential geometry used along the thesis. Depending on the familiarity of the reader with the contents of the chapter, it can either be skipped, it can be quickly read as a mean to settle the notation or it can serve as an introduction to the mathematical concepts for the inexperienced reader.

The four main sections in this chapter are the following:

- (Semi-)Riemannian geometry.
- Symplectic geometry and Hamiltonian systems;
- Tangent bundle constructions and relevant results including lifts, canonical involution, the vertical endomorphsim, Second-order differential equation vector field and distributions;
- Basic notions of Lie group actions.

We remark that this introduction does not pretend to be exhaustive in any way. So that we assume previous knowledge of basic concepts about differentiable manifolds and we do not include proofs in most situations, redirecting the reader to the literature.

If nothing on the contrary is explicitly said, all objects along the thesis are assumed to be smooth. Einstein's summation convention is used.

2.1 Semi-Riemannian geometry

The concept of Riemannian metric and the machinery associated to it have revealed to be extremely useful in order to extend geometric concepts that one has in euclidean spaces, such as length, angle, volume form, gradient of a function or differentiation of vector fields, to more general differentiable manifolds. Perhaps the first important concept derived from a Riemannian metric is that of a geodesic. Geodesics will replace the role of straight lines as the curves which minimize distance between two fixed points in Riemannian manifolds more general than the euclidean space.

Naturally, this is where the interplay of mechanics and geometry begins since, according to Newton's second law of mechanics, particles that are not subjected to external forces will now follow a geodesic on a general differentiable manifold. Analogously, in Einstein's special relativity, if we are given a Lorentz metric, one finds that free particles follow geodesics.

This section will be a review of the more general concept of semi-Riemannian metric, which includes both Riemannian metric and Lorentz metric as particular cases (for more details see [O'N83; Lee97; Car92]).

2.1.1 Semi-Riemannian metric and basic constructions

Definition 2.1.1. A *semi-Riemannian metric* h on a manifold Q is a symmetric non-degenerate $(0, 2)$-tensor of constant signature. Accordingly, if h is a semi-Riemannian metric then the pair (Q, h) is called a semi-Riemannian manifold.

If the semi-Riemannian metric h is positive definite, then h is a Riemannian metric. On the contrary, if h is not definite and has signature of the type $(-1, 1, ..., 1)$ then g is called a Lorentz metric. See [O'N83] for more details.

If (q^i) are local coordinates on Q, then the local expression of the metric tensor h is

$$h = h_{ij} dq^i \otimes dq^j, \quad \text{with } h_{ij} = h\left(\frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j}\right).$$

With the help of the metric, we can also define a norm on each tangent space T_qQ. Given a vector $v_q \in T_qQ$ we define its *norm* as

$$\|v_q\| = \sqrt{|h(v_q, v_q)|}.$$

Moreover, if $c : [0, h] \to Q$ is a curve on Q, then its *length* $L(c)$ is defined to be

$$L(c) = \int_0^h \|\dot{c}(t)\| dt.$$

If (M, h_M) and (N, h_N) are two semi-Riemannian manifolds, a (local) *isometry* is a (local) diffeomorphism $\phi : M \to N$ such that

$$\phi^* h_N = h_M.$$

Given a semi-Riemannian metric we can define the *flat isomorphism* as the map $\flat_h : T_qQ \to T_q^*Q$ in each tangent space T_qQ given by

$$\langle \flat_h(X_q), Y_q \rangle = h(X_q, Y_q), \tag{2.1.1}$$

for every $X_q, Y_q \in T_qQ$. The fact that h is non-degenerate implies that the map $\flat_h$ is indeed an isomorphism on each tangent space. Its inverse isomorphism is denoted by $\sharp_h$ and is called the *sharp isomorphism*. Observe that the musical isomorphisms introduced here extend to maps between $\mathfrak{X}(Q)$ and $\Omega^1(Q)$. Given a function $f \in C^\infty(Q)$, the *gradient vector field* of f with respect to h is the vector field $\mathrm{grad}_h f \in \mathfrak{X}(Q)$ defined by

$$\mathrm{grad}_h f = \sharp_h \circ dV.$$

Let us recall now how a semi-Riemannian metric induces an intrinsic process to differentiate vector fields. Remember that a *linear connection* is a map $\nabla : \mathfrak{X}(Q) \times \mathfrak{X}(Q) \to \mathfrak{X}(Q)$ which is $C^\infty(Q)$-linear on the first factor, $\mathbb{R}$-linear in the second factor and if we denote the image of $X, Y \in \mathfrak{X}(Q)$ by $\nabla_X Y$, then ∇ satisfies the Leibniz rule

$$\nabla_X (fY) = X(f) \cdot Y + f \cdot \nabla_X Y,$$

for every $f \in C^\infty(Q)$. The vector field $\nabla_X Y$ is also called the *covariant derivative of* Y *with respect to* X.

If (q^i) are local coordinates on Q, then the connection is locally characterized by the *Christoffel symbols* which are real-valued functions on Q given by

$$\nabla_{\frac{\partial}{\partial q^i}} \frac{\partial}{\partial q^j} = \Gamma_{ij}^k \frac{\partial}{\partial q^k}.$$

Thus if X and Y are vector fields locally given by

$$X = X^i \frac{\partial}{\partial q^i} \quad \text{and} \quad Y = Y^i \frac{\partial}{\partial q^i},$$

then the covariant derivative may be written as

$$\nabla_X Y = \left(X^i \frac{\partial Y^k}{\partial q^i} + X^i Y^j \Gamma_{ij}^k \right) \frac{\partial}{\partial q^k}. \tag{2.1.2}$$

Notation. We will use the following convention regarding tensors: a tensor F of type (k,l) on the manifold Q is a $C^\infty(Q)$-multilinear map

$$F : \underbrace{\Omega^1(Q) \times ... \times \Omega^1(Q)}_{k\ times} \times \underbrace{\mathfrak{X}(Q) \times ... \times \mathfrak{X}(Q)}_{l\ times} \to C^\infty(Q).$$

The space of (k,l) tensors on Q will be denoted by $\mathcal{T}_l^k(Q)$. However, if $k=1$ we will identify the tensor F of type $(1,l)$ with the map

$$\hat{F} : \underbrace{\mathfrak{X}(Q) \times ... \times \mathfrak{X}(Q)}_{l\ times} \to \mathfrak{X}(Q)$$

uniquely defined by

$$\langle \alpha, \hat{F}(Y_1, ..., Y_l) \rangle = F(\alpha, Y_1, ..., Y_l),$$

for all $\alpha \in \Omega^1(Q)$ and $Y_1, ..., Y_l \in \mathfrak{X}(Q)$.

Given a vector field $X \in \mathfrak{X}(Q)$, the covariant derivative ∇_X may be generalized to a map from the space of (l,k) type tensors to itself in the following way. First, for 1-forms $\alpha \in \Omega^1(Q)$ we have that

$$\langle \nabla_X \alpha, Y \rangle = X(\langle \alpha, Y \rangle) - \langle \alpha, \nabla_X Y \rangle.$$

If F is a tensor of type (k,l), $Y_1, ..., Y_l$ are l vector fields and $\alpha_1, ..., \alpha_k$ are k 1-forms, then

$$\begin{aligned}(\nabla_X F)(\alpha_1, ..., \alpha_k, Y_1, ..., Y_l) =& X(F(\alpha_1, ..., \alpha_k, Y_1, ..., Y_l)) \\ & - \sum_{i=1}^{k} F(\alpha_1, ..., \nabla_X \alpha_i, ..., \alpha_k, Y_1, ..., Y_l) \\ & - \sum_{i=1}^{l} F(\alpha_1, ..., \alpha_k, Y_1, ..., \nabla_X Y_i, ..., Y_l)).\end{aligned}$$

Moreover, given F a tensor of type (k,l) we can define the tensor ∇F of type $(k, l+1)$ called the *total covariant derivative* of F given by

$$\nabla F(\alpha_1, ..., \alpha_k, X, Y_1, ..., Y_l) = (\nabla_X F)(\alpha_1, ..., \alpha_k, Y_1, ..., Y_l).$$

Now, we recall the definition of the Levi-Civita connection associated to a semi-Riemannian metric h.

Theorem 2.1.2 (Fundamental theorem of semi-Riemannian geometry)**.** *Let (Q, h) be a semi-Riemannian manifold. There is a unique connection ∇^h such that*

1. $\nabla^h_X Y - \nabla^h_Y X = [X, Y]$;
2. $X(h(Y, Z)) = h(\nabla^h_X Y, Z) + h(Y, \nabla^h_X Z)$,

for all $X, Y, Z \in \mathfrak{X}(Q)$. Then ∇^h is called the Levi-Civita connection associated to h and is characterized by the Koszul formula

$$
\begin{aligned}
2h(\nabla^h_X Y, Z) =& X(h(Y, Z)) + Y(h(Z, X)) - Z(h(X, Y)) \\
& - h(X, [Y, Z]) + h(Y, [Z, X]) + h(Z, [X, Y]).
\end{aligned}
$$

Proof. Cf. [O'N83]. □

If (q^i) are local coordinates on Q, then the local expression of the Christoffel symbols of the Levi-Civita connection associated with h are given by

$$
\Gamma^k_{ij} = \frac{1}{2} h^{km} \left(\frac{\partial h_{jm}}{\partial q^i} + \frac{\partial h_{im}}{\partial q^j} - \frac{\partial h_{ij}}{\partial q^m} \right),
$$

where (h^{km}) is the inverse matrix of h_{km}.

Now, we will generalize the fundamental formula of Riemannian geometry. We will denote by $\mathcal{L}_X$ the Lie derivative with respect to X.

Lemma 2.1.3. *Let h be a symmetric non-degenerate $(0, 2)$-tensor and ∇^h the Levi-Civita connection with respect to h. Then the Lie derivative of h satisfies*

$$
\mathcal{L}_X h(Y, Z) = 2h(\nabla^h_Y X, Z) - d(\flat_h(X))(Y, Z), \quad X, Y, Z \in \mathfrak{X}(Q). \tag{2.1.3}
$$

Proof. By definition of Lie derivative one has that

$$
\mathcal{L}_X h(Y, Z) = X(h(Y, Z)) - h([X, Y], Z) - h(Y, [X, Z]).
$$

Using the fact that the Levi-Civita connection ∇^h is symmetric and compatible with the metric h one gets

$$
\mathcal{L}_X h(Y, Z) = h(\nabla^h_Y X, Z) + h(Y, \nabla^h_Z X). \tag{2.1.4}
$$

Also from definition of differential of a one-form we know that

$$d(\flat_h(X))(Y,Z) = Y(h(X,Z)) - Z(h(X,Z)) - h(X,[Y,Z]).$$

It is not difficult to show using the Koszul's formula for ∇^h that

$$h(Y,\nabla^h_Z X) - h(\nabla^h_Y X, Z) = -d(\flat_h(X))(Y,Z)$$

and plugging in the last equation into (2.1.4), we get the desired formula for $\mathcal{L}_X h(Y,Z)$. □

In order to introduce the notion of geodesic, we will need first to show how the covariant derivative extends to vector fields along curves.

Let $c : I \to Q$ be a curve on the semi-Riemannian manifold (Q,h). A *vector field along* c is a smooth map $W : I \to TQ$ such that $W(t) \in T_{c(t)}Q$ for every $t \in I$. Note that the most trivial example of such a map is the derivative of the curve $\dot{c} : I \to TQ$ itself. Let $\mathfrak{X}(c)$ denote the space of vector fields along c.

Lemma 2.1.4. *Let ∇ be a linear connection on Q. For each curve $c : I \to Q$, there is a unique operator $\nabla_{\dot{c}} : \mathfrak{X}(c) \to \mathfrak{X}(c)$ called the covariant derivative along c that is $\mathbb{R}$-linear, satisfies the Leibniz rule*

$$\nabla_{\dot{c}}(f \cdot W) = \dot{f} \cdot W + f \cdot \nabla_{\dot{c}} W, \quad \textit{for any } f \in C^\infty(I),$$

and if $\widetilde{W}$ is a vector field on Q such that $\widetilde{W} \circ c = W$ then

$$\nabla_{\dot{c}} W(t) = \nabla_{\dot{c}} \widetilde{W}\Big|_{c(t)}. \tag{2.1.5}$$

Proof. See [Lee97] for proof. □

Remark 2.1.5. We remark that the right-hand side of equation (2.1.5) is well-defined over points in the image of c. Indeed, by equation (2.1.2) we see that for each point $q \in Q$, the covariant derivative $\nabla_X Y(q)$ just depends on the value of X at q and the values of Y along some curve $\gamma : (-\varepsilon, \varepsilon) \to Q$ with $\gamma(0) = q$ and $\dot{\gamma}(0) = X(q)$.

Moreover if $(q^i(t))$ are the local coordinate functions of the curve c and $W : I \to TQ$ is a curve over c locally given by

$$W(t) = W^i(t) \left.\frac{\partial}{\partial q^i}\right|_{c(t)},$$

then the covariant derivative of W along c has the local expression

$$\nabla_{\dot{c}} W(t) = \left(\dot{W}^k(t) + \dot{q}^i(t) W^j(t) \Gamma^k_{ij}(c(t)) \right) \frac{\partial}{\partial q^k}.$$

Definition 2.1.6. Given a linear connection ∇ on the manifold Q, a *geodesic* is a curve $c : I \to Q$ such that its tangent lift has vanishing covariant derivative along itself, that is,

$$\nabla_{\dot{c}}\, \dot{c} = 0. \tag{2.1.6}$$

In local coordinates, equations (2.1.6) is a second-order differential equation given by

$$\nabla_{\dot{c}}\, \dot{c}(t) = \left(\ddot{q}^k(t) + \dot{q}^i(t) \dot{q}^j(t) \Gamma^k_{ij}(c(t)) \right) \frac{\partial}{\partial q^k} = 0.$$

Thus, applying standard existence and uniqueness theorems of differential equations on a coordinate neighbourhood we deduce that given any $v \in T_qQ$, there is a unique geodesic denoted by $c_v : I \to Q$ such that $\dot{c}_v(0) = v$ and the domain interval I is maximal.

In fact, a curve $c : I \to Q$ is a geodesic if and only if its tangent lift $\dot{c} : I \to TQ$ is an integral curve of the geodesic vector field Γ^∇ (see Proposition 28, Chapter 3 in [O'N83]). If the flow of the geodesic vector field is denoted by $\phi_t^{\Gamma^\nabla} : TQ \to TQ$, then the geodesic c_v may be written as

$$c_v(t) = \tau_Q \circ \phi_t^{\Gamma^\nabla}(v)$$

for each $v \in TQ$.

Finally, let us recall the definitions of the *Torsion tensor* and the *Curvature tensor*. Given a linear connection ∇, the Torsion tensor is the $(1, 2)$ type tensor $T : \mathfrak{X}(Q) \times \mathfrak{X}(Q) \to \mathfrak{X}(Q)$ given by

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y].$$

Given local coordinates (q^i) on Q the local expression of T is

$$T = T^k_{ij} \frac{\partial}{\partial q^k}, \quad \text{where} \quad T^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji}.$$

Given a semi-Riemannian metric h and the associated Levi-Civita connection ∇^h, we denote by T^h the Torsion tensor with respect to ∇^h. By the definition of the Levi-Civita connection, T^h vanishes identically. On local coordinates, this implies that the Christoffel symbols are *symmetric* that is

$$\Gamma^k_{ij} = \Gamma^k_{ji}.$$

The *Curvature tensor* associated to a given linear connection ∇ is the $(1,3)$ type tensor $R : \mathfrak{X}(Q) \times \mathfrak{X}(Q) \times \mathfrak{X}(Q) \to \mathfrak{X}(Q)$ given by

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z,$$

where $R(X,Y)Z$ simply denotes $R(X,Y,Z)$ following the notation used in most of the literature. In local coordinates, the expression of R is

$$R = R^l_{ijk} \frac{\partial}{\partial q^l}, \quad \text{where} \quad R^l_{ijk} = \frac{\partial \Gamma^l_{jk}}{\partial q^i} - \frac{\partial \Gamma^l_{ik}}{\partial q^j} + \Gamma^m_{jk}\Gamma^l_{im} - \Gamma^m_{ik}\Gamma^l_{jm}.$$

As a final remark, now that we have introduced the Torsion tensor, we may rewrite the defining conditions of the Levi-Civita connection ∇_h associated to the semi-Riemannian metric h as the unique connection satisfying

$$T^h \equiv 0 \quad \text{and} \quad \nabla h \equiv 0.$$

2.1.2 The exponential map

Let us review the notion of exponential map in semi-Riemannian geometry. Being one of the main tools in Riemannian geometry, it is used to tackle a variety of problems. Its utility lies in the fact that it provides a natural identification of a neighbourhood of each point on the manifold with the tangent space at that point. This allows a simpler treatment of local properties.

Let $q_0 \in Q$ be a point on the semi-Riemannian manifold (Q,h). In this section, we will consider all geodesics relative to the Levi-Civita connection ∇^h. Then let

$$\mathcal{M}_{q_0} = \{v_{q_0} \in T_{q_0}Q \mid c_{v_{q_0}} \text{ is defined in } [0,1]\}.$$

This is an open subset containing the zero vector $0_{q_0} \in T_{q_0}Q$.

Then the *exponential map* at q_0 is the map

$$\exp^h_{q_0} : \mathcal{M}_{q_0} \to Q, \quad v_{q_0} \mapsto \tau_Q \circ \phi_1^{\Gamma^{\nabla^h}}(v_{q_0}) = c_{v_{q_0}}(1). \tag{2.1.7}$$

One consequence of the definition, together with the fact that geodesics rescale, i.e., $c_v(\lambda t) = c_{\lambda v}(t)$, is that if $v_{q_0} \in \mathcal{M}_{q_0}$, the geodesic $c_{v_{q_0}} : [0,1] \to Q$ is just

$$c_{v_{q_0}}(t) = \exp^h_{q_0}(t \cdot v_{q_0}).$$

The main result about the exponential map is that it is a diffeomorphism on a neighbourhood of the zero vector $0_{q_0} \in T_{q_0}Q$.

Theorem 2.1.7. *For any $q_0 \in Q$ there exists a neighbourhood $V_{q_0} \subset \mathcal{M}_{q_0}$ containing 0_{q_0} and a neighbourhood $\mathcal{U}_{q_0}$ of q_0 in Q such that*

$$\exp^h_{q_0} : V_{q_0} \to \mathcal{U}_{q_0}$$

is a diffeomorphism.

The previous theorem, which may very easily be proven using the fact that the tangent map of $\exp^h_{q_0}$ at the zero vector is the identity, has important geometric and analytical consequences. We may deduce that, within the neighbourhood $\mathcal{U}_{q_0} \subseteq Q$ of q_0 there exists a unique geodesic connecting any point to q_0! An important notion is that of *normal neighbourhoods.* An open subset $\mathcal{U}_{q_0}$ is called normal if it is the image by the exponential map $\exp^h_{q_0}$ of a star-shaped open subset V_{q_0}. In particular, we can always find normal neighbourhoods.

We may also consider the *extended exponential map.* Let $\mathcal{M}$ be the open subset of the tangent bundle TQ given by

$$\mathcal{M} = \bigcup_{q_0 \in Q} \mathcal{M}_{q_0}.$$

Then, we define

$$\exp^h : \mathcal{M} \to Q \times Q, \quad v_{q_0} \mapsto (\tau_Q(v_{q_0}), \tau_Q \circ \phi_1^{\Gamma^{\nabla^h}}(v_{q_0})),$$

or, equivalently, $\exp^h(v_{q_0}) = (c_{v_{q_0}}(0), c_{v_{q_0}}(1))$, where $c_{v_{q_0}}$ is the unique geodesic with initial velocity v_{q_0}.

Thus, the exponential map $\exp^h$ returns the end-points of each well-defined geodesic within the closed time interval $[0,1]$. Similarly to the restricted case before, we may prove that

Theorem 2.1.8. *There exists a neighbourhood V of the zero section in $\mathcal{M}$ and a tubular neighbourhood $\mathcal{U}$ around the diagonal set $\Delta_Q \subset Q \times Q$ given by*

$$\Delta_Q = \{(q, q) \in Q \times Q | q \in Q\},$$

such that

$$\exp^h : V \to \mathcal{U}$$

is a diffeomorphism.

2.1.3 Minimizing properties of geodesics

In this subsection, let (Q, g) be a Riemannian manifold.

The first ingredient we must present before discussing minimization properties of Riemannian geodesics is *Gauss lemma*, a fundamental result assuring that the orthogonality to radial directions is preserved.

We will write it using the vertical lift notation. Given vectors $v_q, u_q \in T_qQ$ the *vertical lift* of u_q at v_q is the tangent vector to TQ given by

$$(u_q)^{\mathbf{v}}_{v_q} = \left.\frac{d}{ds}\right|_{s=0} (v_q + su_q) \in T_{v_q}(TQ).$$

In Section 2.4.1 below, we will review more about the vertical lift. At the moment, it is enough to note that it induces an isomorphism between T_qQ and $T_{v_q}(T_qQ)$.

Lemma 2.1.9. *Let $q \in Q$ and let $v_q, w_q \in T_qQ$ be tangent vectors. Then,*

$$g_{\exp^g_q(v_q)}((T_{v_q}\exp^g_q)(v_q)^{\mathbf{v}}_{v_q}, (T_{v_q}\exp^g_q)(w_q)^{\mathbf{v}}_{v_q}) = g_q(v_q, w_q), \tag{2.1.8}$$

where $(v_q)^{\mathbf{v}}_{v_q}, (w_q)^{\mathbf{v}}_{v_q} \in T_{v_q}(T_qQ)$ are the vertical lifts to TQ at v_q of the vectors v_q and w_q, respectively.

Remark 2.1.10. Under the linear identification ${}^{\mathbf{v}}_{v_q} : T_qQ \to T_{v_q}(T_qQ)$ between T_qQ and $T_{v_q}(T_qQ)$, equation (2.1.8) is converted in the clearer expression

$$g(\exp^g_q(v_q))(T_{v_q}\exp^g_q(v_q), T_{v_q}\exp^g_q(w_q)) = g(q)(v_q, w_q). \tag{2.1.9}$$

Indeed, the Gauss Lemma is crucial to prove the following minimization property:

Proposition 2.1.11. *Let $\mathcal{U}_{q_0}$ be a normal neighbourhood of a point q_0 in the Riemannian manifold Q. If $q_1 \in \mathcal{U}_{q_0}$ then the radial geodesic curve $c : [0,1] \to \mathcal{U}_{q_0}$ from q_0 to q_1 is the unique shortest curve in $\mathcal{U}_{q_0}$ connecting both points.*

So, if $\sigma : [0, \lambda] \to \mathcal{U}_{q_0}$ is any other curve connecting q_0 to q_1 then

$$L(\sigma) \geqslant L(c)$$

and the equality holds if and only if one is a monotone reparametrization of the other.

The last Proposition states that, at least locally, geodesics are the unique curves minimizing the length between two points. Unfortunately, we can not hope that the same statement holds globally. This means either that a geodesic between two arbitrary points on the manifold may stop minimizing the length from some point on or that it is not the unique minimum length curve. There are several examples in non-euclidean manifolds. It is well-known that on the sphere, two antipodal points are connected by infinitely many geodesics with the same length.

Considering a given point q on the manifold, the set of points from which a geodesic starting at q stops minimizing the length is related with the singularities of the exponential map $\exp_q^g$.

2.1.4 Jacobi fields

In this section we will recall the notion of Jacobi field in Riemannian geometry which is linked to the way that different geodesics fall apart. In that sense, there is a connection between Jacobi fields and *variations*. We will introduce here some piece of notation related with calculus of variations that we will use later when discussing critical points of functionals, particularly in Lagrangian mechanics.

A *variation* of a curve $c : I \to Q$ is a map of the form

$$\begin{aligned} \Phi : (-\varepsilon, \varepsilon) \times I &\longrightarrow Q \\ (s,t) &\mapsto \Phi_s(t) \end{aligned} \tag{2.1.10}$$

such that $\Phi_0(t) = c(t)$ for all $t \in I$. The vector field $V : I \to TQ$ defined as

$$V(t) = \left.\frac{\partial}{\partial s}\right|_{s=0} \Phi_s(t) \tag{2.1.11}$$

is called the *infinitesimal variation vector field* of the variation Φ. Observe that for each $t \in I$ we have that $V(t) \in T_{c(t)}Q$.

Now the variation Φ is said to be a *variation by geodesics* if for each $s \in (-\varepsilon, \varepsilon)$ the curve $\Phi_s : I \to Q$ is a geodesic.

Definition 2.1.12. A *Jacobi field* along a geodesic $c : I \to Q$ is the infinitesimal variation vector field of a variation of c by geodesics.

Now, we have the following Theorem which asserts that Jacobi fields satisfy the so-called Jacobi equation:

Theorem 2.1.13. *Let $c : I \to Q$ be a geodesic and $W : I \to TQ$ a vector field along c. The vector field W is a Jacobi field if and only if W satisfies the equation*

$$\nabla_{\dot{c}} \nabla_{\dot{c}} W + R(W, \dot{c})\dot{c} = 0. \tag{2.1.12}$$

Usually, it is difficult to identify a Jacobi field when we do not know all the geodesics on the manifold. But there are some particular cases where we can deduce that a specific vector field is Jacobi. It is the case of the following Proposition.

Proposition 2.1.14. *Let $W \in \mathfrak{X}(Q)$ be a Killing vector field, i.e., $\mathcal{L}_W g = 0$. Then if $c : I \to Q$ is a geodesic, $W \circ c : I \to TQ$ is a Jacobi field along c.*

2.2 Symplectic geometry

As we have discussed previously, classical mechanics takes place in the symplectic playground. Hence, the knowledge of symplectic geometry is essential to unveil many underlying properties of mechanical systems (see [MS17; Lee13]).

2.2.1 Symplectic vector spaces

Let us start this section with a run through symplectic vector spaces. For a more comprehensive introduction we refer to [MS17]. Recall that a *symplectic vector space* is a finite dimensional real vector space V equipped with a non-degenerate skew-symmetric bilinear form $\omega : V \times V \to \mathbb{R}$ called the *symplectic form*. This means that

- For all $v, w \in V$, $\omega(v, w) = -\omega(w, v)$. (*skew-symmetry*)

- Given $v \in V$ if $\omega(v, w) = 0, \ \forall w \in V \ \Rightarrow v = 0$. (*non-degeneracy*)

The symplectic form induces a linear isomorphism $\flat_\omega : V \to V^*$, given by $\langle \flat_\omega(v), w \rangle = \omega(v, w)$ for any $v, w \in V$. In fact, the non-degeneracy condition is equivalent to this map being an isomorphism.

A linear isomorphism $\Psi : V \to V$ is said to be a *symplectomorphism* if it preserves the symplectic form in the sense that $\omega(\Psi(v), \Psi, (w)) = \omega(v, w)$ for $v, w \in V$.

The *symplectic complement* of a linear subspace $W \subseteq V$ is denoted by $W^\perp$ and defined as

$$W^\perp = \{v \in V | \omega(v, w) = 0 \ \forall w \in W\}.$$

Note that, unlike what happens with inner products, the skew-symmetry property does not force $W^\perp$ to be transversal to W. A subspace W might be

- *isotropic* if $W \subseteq W^\perp$;
- *coisotropic* if $W^\perp \subseteq W$;
- *symplectic* if $W \cap W^\perp = \{0\}$;
- *Lagrangian* if $W = W^\perp$.

Then, we have the following results.

Proposition 2.2.1. *W is isotropic if and only if ω vanishes on W and it is symplectic if and only if $\omega|_W$ is non-degenerate.*

Lemma 2.2.2. *For any subspace $W \subseteq V$, we have*

$$dim\ W + dim\ W^\perp = dim\ V.$$

Thus W is Lagrangian if and only if is isotropic and has half the dimension of V.

Example 2.2.3. The following example of symplectic form is fundamental because, by the next proposition, we will see that it is the canonical symplectic form. Let V be a real vector space with dimension $2n$ and

$\{A_1, ..., A_n, B_1, ..., B_n\}$ a basis of V with $\{\alpha^1, ..., \alpha^n, \beta^1, ..., \beta^n\}$ as the corresponding dual basis. The bilinear form

$$\omega = \sum_{i=1}^{n} \alpha^i \wedge \beta^i \tag{2.2.1}$$

is a symplectic form on V. △

The following result is Proposition 22.7 in [Lee13]:

Proposition 2.2.4. *Let ω be a symplectic form on an m-dimensional vector space V. Then V has even dimension $m = 2n$ and there exists a basis for V in which ω has the form* (2.2.1)*.*

Moreover, we have that:

Proposition 2.2.5. *A skew-symmetric bilinear form ω on the $2n$-dimensional vector space V is symplectic if and only if $\omega^n = \omega \wedge \cdots \wedge \omega \neq 0$.*

2.2.2 Symplectic manifolds

A *symplectic manifold* is a manifold M equipped with a closed 2-form $\omega \in \Omega^2(M)$ such that ω_p is a symplectic form for each $p \in M$ in the tangent space T_pM. Note that, by Proposition 2.2.4, the symplectic manifold (M, ω) must have even dimension. Also, following Proposition 2.2.5, the $2n$-form ω^n is non-vanishing. Hence, M is orientable.

Example 2.2.6. The standard model of symplectic manifold is the euclidean space $M = \mathbb{R}^{2n}$. If (x^i, y^i) are coordinates on M, then using Example 2.2.3, we know that

$$\omega_0 = \sum_{i=1}^{n} dx^i \wedge dy^i$$

is a symplectic form on each tangent space to M. Moreover, it is a trivial calculation to check that $d\omega_0 = 0$, thus, $(\mathbb{R}^{2n}, \omega_0)$ is a symplectic manifold. △

We will see now that all the previous constructions on symplectic vector spaces extend naturally to manifolds.

Similar to the vector space case, we may define the musical isomorphism $\flat_\omega : \mathfrak{X}(M) \to \Omega^1(M)$ as the isomorphism of $C^\infty(M)$-modules defined by

$\langle \flat_\omega(X), Y\rangle = \omega(X,Y)$, for all $X, Y \in \mathfrak{X}(M)$. Its inverse isomorphism is denoted by $\sharp_\omega : \Omega^1(M) \to \mathfrak{X}(M)$.

A diffeomorphism $\Psi : M_1 \to M_2$ between two symplectic manifolds (M_1, ω_1) and (M_2, ω_2) is said to be a *symplectomorphism* if it preserves the symplectic structures in the sense that

$$\omega_1 = \Psi^* \omega_2.$$

A submanifold $N \subseteq M$ is said to be isotropic, coisotropic, symplectic or Lagrangian if the tangent spaces T_pN have the corresponding property for all $p \in N$.

Example 2.2.7. The next example of symplectic manifold is of utmost importance in geometric mechanics, since it lays the foundations of Hamiltonian mechanics.

Given a manifold Q, its cotangent bundle T^*Q is equipped with a *canonical* symplectic structure. The *canonical 1-form* $\theta_Q \in \Omega^1(Q)$ on T^*Q is given by

$$\langle \theta_Q(\alpha_q), X_{\alpha_q}\rangle = \langle \alpha_q, T_{\alpha_q}\pi_Q(X_{\alpha_q})\rangle$$

where $\alpha_q \in T_q^*Q$, $X_{\alpha_q} \in T_{\alpha_q}T^*Q$ and $\pi_Q : T^*Q \to Q$ is the canonical projection. If (q^i, p_i) are local bundle coordinates on T^*Q then

$$\theta_Q(q,p) = p_i dq^i.$$

Then the exact 2-form $\omega_Q \in \Omega^2(Q)$ given by

$$\omega_Q = -d\theta_Q$$

is a symplectic form and it is called the *canonical symplectic structure* of T^*Q. In local coordinates, it is given by

$$\omega_Q(q,p) = dq^i \wedge dp_i.$$

△

The following result shows that all symplectic manifolds are locally equivalent.

Theorem 2.2.8 (Darboux Theorem)**.** *Let (M,ω) be a $2n$-dimensional symplectic manifold. For any $p \in M$, there is a neighbourhood U of p, where the* *smooth coordinates* $(x^1, ..., x^n, y^1, ..., y^n)$ *are defined and on which* ω *has the local expression*

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

Remark 2.2.9. The coordinates mentioned in Darboux Theorem are called *Darboux coordinates* or *canonical coordinates*. The standard coordinates in $\mathbb{R}^{2n}$ in Example 2.2.6 and the bundle coordinates on T^*Q in Example 2.2.7 are two examples of canonical coordinates for the symplectic forms ω_0 and ω_Q, respectively.

2.2.3 Hamiltonian vector fields and Poisson brackets

Given $X \in \mathfrak{X}(Q)$, recall that the *interior multiplication* by X is an operator from the space of (k, l)- tensors in Q to the space of $(k, l-1)$-tensors, that is

$$i_X : \mathcal{T}_l^k(Q) \to \mathcal{T}_{l-1}^k(Q)$$
$$(i_X F)(\alpha^1, ..., \alpha^k, X_1, ..., X_{l-1}) = F(\alpha^1, ..., \alpha^k, X, X_1, ..., X_{l-1}),$$

where $F \in \mathcal{T}_l^k(Q)$.

Naturally, if (M, ω) is a symplectic manifold and $X \in \mathfrak{X}(Q)$, then $i_X\omega = \flat_\omega(X)$.

Definition 2.2.10. Let (M, ω) be a symplectic manifold. A vector field $X \in \mathfrak{X}(M)$ is called *symplectic* if $\mathcal{L}_X\omega = 0$, it is called *Hamiltonian* if there exists a function $H \in C^\infty(M)$ such that $\flat_\omega(X) = dH$ and it is called *locally Hamiltonian* if any $p \in M$ has a neighbourhood U on which $\flat_{\omega|_U}(X|_U) = df$ for some function $f \in C^\infty(U)$.

These concepts are related to each other. A Hamiltonian vector field X is not only locally Hamiltonian but it is also symplectic since, by Cartan's identity, we have that

$$\mathcal{L}_X\omega = d(i_X\omega) + i_X d\omega = 0,$$

and due to the facts that ω is closed and $i_X\omega = dH$, for some function H on M and where the operator i denotes the inner multiplication.

Proposition 2.2.11. *A vector field* X *in the symplectic manifold* (M, ω) *is symplectic if and only if its flow* $\phi_t^X : M \to M$ *is a symplectomorphism for every* t.

Then an important consequence of the last Proposition is given by the next corollary.

Corollary 2.2.12. *The flow of every Hamiltonian vector field is a symplectomorphism.*

Given a function H on M, there exists a unique vector field $X_H \in \mathfrak{X}(M)$ satisfying the equation

$$i_{X_H}\omega = dH.$$

Moreover, this vector field is given by $X_H = \sharp_\omega(dH)$. By definition, X_H is Hamiltonian and it is called the Hamiltonian vector field associated to H.

Example 2.2.13. Let $(\mathbb{R}^{2n}, \omega_0)$ be the standard symplectic structure in even-dimensional euclidean space. Given a function H, the Hamiltonian vector field has the coordinate expression

$$X_H = \frac{\partial H}{\partial y^i}\frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i}\frac{\partial}{\partial y^i}.$$

△

One of the most remarkable properties of Hamiltonian systems is that the Hamiltonian function is a first integral of the Hamiltonian vector field.

Proposition 2.2.14. *Let (M, ω) be a symplectic manifold and H a function on M. Then, H is constant along the integral curves of X_H and, at each regular point of H, X_H is tangent to the level set of H.*

In a symplectic manifold (M, ω), using the fact that each function has a unique Hamiltonian vector field associated to it, we can define a Poisson bracket $\{\cdot,\cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$, which is given by

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

In fact, we can define Poisson brackets in any manifold. A bracket of functions in a manifold Q is a *Poisson bracket* if it is $\mathbb{R}$-bilinear, skew-symmetric, satisfies the Leibniz rule

$$\{f, gh\} = g\{f, h\} + \{f, g\}h, \quad \text{for } f, g, h \in C^\infty(Q)$$

and the *Jacobi identity*

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0, \quad \text{for } f, g, h \in C^\infty(Q).$$

Example 2.2.15. In $(\mathbb{R}^{2n}, \omega_0)$ the Poisson bracket associated to the symplectic structure is given by

$$\{f, g\} = \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial y^i} - \frac{\partial f}{\partial y^i}\frac{\partial g}{\partial x^i}.$$

$\triangle$

Now, it is easy to prove the following proposition.

Proposition 2.2.16. *Let (M, ω) be a symplectic manifold and $H \in C^\infty(M)$ a Hamiltonian function on M. A function $f \in C^\infty(M)$ is a conserved quantity along the integral curves of the Hamiltonian vector field X_H if and only if $\{f, H\} = 0$.*

Remark 2.2.17. By skew-symmetry of the Poisson bracket we trivially have that $\{H, H\} = 0$, so we find again that H is conserved along the motion, which we already knew by Proposition 2.2.14.

Remark 2.2.18. Note that, the local expression of the Hamiltonian vector field can be recovered using the Poisson brackets. That is, if (x^i, y^i) are Darboux coordinates on M then, the local expression of X_H is

$$X_H = \{x^i, H\}\frac{\partial}{\partial x^i} + \{y^i, H\}\frac{\partial}{\partial y^i}.$$

We can generalize the notion of Poisson brackets by observing that they are a manifestation of the $(2,0)$-type tensor $\Lambda : \Omega^1(M) \times \Omega^1(M) \to C^\infty(M)$ given by

$$\Lambda(\alpha, \beta) = \omega(\sharp_\omega(\alpha), \sharp_\omega(\beta)).$$

Then the Poisson bracket associated to the symplectic structure is simply given by

$$\{f, g\} = \Lambda(df, dg).$$

The map Λ is called the *Poisson structure* of M associated to the symplectic structure ω.

2.3 Contact geometry

In this section we will recall the main definitions and results on the theory of contact manifolds and Hamiltonian system. See [LLV19a; LL21] for a more detailed overview.

A *contact manifold* (M, η) is a $(2n+1)$-dimensional manifold with a *contact form* η [God69; LM87]. That is, η is a 1-form on M such that $\eta \wedge d\eta^n$ is a volume form. To give some motivation to the introduction of contact structures, some authors describe contact geometry to be an odd-dimensional analogue to symplectic geometry. This type of manifolds have a distinguished vector field: the so-called *Reeb vector field* $\mathcal{R}$, which is the unique vector field satisfying

$$i_{\mathcal{R}} d\eta = 0, \quad \eta(\mathcal{R}) = 1. \tag{2.3.1}$$

On a contact manifold (M, η), we define the following isomorphism of vector bundles:

$$\begin{aligned} \flat : \quad TM &\longrightarrow T^*M, \\ v &\longmapsto i_v d\eta + \eta(v)\eta. \end{aligned} \tag{2.3.2}$$

Notice that the image of the Reeb vector field under $\flat$ is exactly the contact form, i.e., $\flat(\mathcal{R}) = \eta$.

There is a Darboux theorem for contact manifolds. In a neighbourhood of each point in M one can find local coordinates (q^i, p_i, z) such that

$$\eta = dz - p_i dq^i. \tag{2.3.3}$$

In these coordinates, we have

$$\mathcal{R} = \frac{\partial}{\partial z}. \tag{2.3.4}$$

Example 2.3.1. An example of a contact manifold is $T^*Q \times \mathbb{R}$. Here, the contact form is given by

$$\eta_Q = dz - \theta_Q = dz - p_i dq^i, \tag{2.3.5}$$

where θ_Q is the pullback of the canonical 1-form of T^*Q, (q^i, p_i) are natural coordinates on T^*Q and z is the $\mathbb{R}$-coordinate. We call η_Q the *canonical contact structure* on $T^*Q \times \mathbb{R}$. △

We say that a (local) diffeomorphism between two contact manifolds $F : (M, \eta) \to (N, \tau)$ is a (local) *contactomorphism* if $F^*\tau = \eta$.

If $\alpha \in \Omega^1(Q)$, we may define the distribution $\ker \alpha \subseteq TQ$ given by

$$\ker \alpha = \{v \in TQ \mid \langle \alpha, v \rangle = 0\}.$$

Then, we say that F is a (local) *conformal contactomorphism* if $F^* \ker \tau = \ker \eta$ or, equivalently, $F^*\tau = \sigma\eta$, where $\sigma : M \to \mathbb{R} \setminus \{0\}$ is the *conformal factor*.

We say that a vector field X on M is an *infinitesimal (conformal) contactomorphism* if its flow F_t consists of (conformal) contactomorphisms.

From the general identity, where F_t is a flow and X is its infinitesimal generator

$$\frac{\partial}{\partial t} F_t^* \eta = F_t^* \mathcal{L}_X \eta, \tag{2.3.6}$$

we deduce that X is an infinitesimal contactomorphism if and only if

$$\mathcal{L}_X \eta = 0. \tag{2.3.7}$$

Furthermore, X is a conformal contactomorphism if and only if

$$\mathcal{L}_X \eta = a\eta, \tag{2.3.8}$$

for some smooth function $a : M \to \mathbb{R}$. The function a is related to the conformal factors σ_t of the conformal contactomorphisms F_t by

$$\sigma_t(x) = \exp\left(\int_0^t a(F_\tau(x)) d\tau\right). \tag{2.3.9}$$

Given a smooth function $f : M \to \mathbb{R}$, its *Hamiltonian vector field* X_f is given by

$$\flat(X_f) = df - (f + \mathcal{R}(f))\eta. \tag{2.3.10}$$

A vector field X is the Hamiltonian vector field of some function f if and only if it is an infinitesimal conformal contactomorphism. In that case $X = X_f$ for $f = -\eta(X)$. Moreover, $\mathcal{L}_X \eta = -\mathcal{R}(f)\eta$. Hence X is an infinitesimal contactomorphism if and only if $X = X_f$ for some function f such that $\mathcal{R}(f) = 0$.

We call the triple (M, η, H) a *contact Hamiltonian system*, where (M, η) is a contact manifold and $H : M \to \mathbb{R}$ is the *Hamiltonian function*.

In contrast to their symplectic counterpart, contact Hamiltonian vector fields do not preserve the Hamiltonian. In fact

$$X_H(H) = -\mathcal{R}(H)H. \tag{2.3.11}$$

Similarly to symplectic geometry, where a symplectic structure gives rise to a Poisson structure, in contact geometry the contact structure gives rise to a Jacobi structure, which is a generalization of Poisson structures. Indeed, the pair (Λ, E), where Λ is a bi-vector on M and E is a vector field, forms a *Jacobi structure* if it satisfies the equations

$$[\Lambda, \Lambda] = 2E \wedge \Lambda \quad \text{and} \quad [\Lambda, E] = 0,$$

where $[\cdot, \cdot]$ denotes the Schouten-Nijenhuis bracket (see, for instance, [MR13; Vai94]). Observe that when E vanishes the bi-vector Λ is a Poisson tensor.

Given a contact manifold (M, η), we may define a Jacobi structure formed by the bi-vector Λ on M given by

$$\Lambda(\alpha, \beta) = -d\eta(\flat^{-1}(\alpha), \flat^{-1}(\beta)), \qquad \alpha, \beta \in \Omega^1(M) \tag{2.3.12}$$

and the vector field $E = -\mathcal{R}$. In canonical coordinates,

$$\Lambda = \frac{\partial}{\partial p_i} \wedge \left(\frac{\partial}{\partial q^i} + p_i \frac{\partial}{\partial z} \right) \tag{2.3.13}$$

Define the $C^\infty(M)$-linear mapping

$$\sharp_\Lambda : \Omega^1(M) \to \mathfrak{X}(M)$$

by $\langle \beta, \sharp(\alpha) \rangle = \Lambda(\alpha, \beta)$ with $\alpha, \beta \in \Omega^1(M)$. Given a Hamiltonian function $H \in C^\infty(M)$, the corresponding contact Hamiltonian vector field X_H is characterized by

$$X_H = \sharp_\Lambda(dH) - H\mathcal{R}.$$

In canonical coordinates:

$$X_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} - \left(\frac{\partial H}{\partial q^i} + p_i \frac{\partial H}{\partial z} \right) \frac{\partial}{\partial p_i} + \left(p_i \frac{\partial H}{\partial p_i} - H \right) \frac{\partial}{\partial z}$$

From this Jacobi structure, we can define a *Jacobi bracket* as follows:

$$\{f, g\} = \Lambda(df, dg) + fE(g) - gE(f), \quad f, g \in C^\infty(M)$$

The mapping $\{\ ,\ \} : C^\infty(M) \times C^\infty(M) \longrightarrow C^\infty(M)$ is bilinear, skew-symmetric and satisfies the Jacobi's identity but, in general, it does not satisfy the Leibniz rule; this last property is replaced by a weaker condition:

$$\mathrm{Supp}\ \{f, g\} \subset \mathrm{Supp}\ f \cap \mathrm{Supp}\ g\ ,$$

where Supp$f$ denotes the support of the function f.

In this sense, this bracket also generalizes the well-known Poisson brackets. In local coordinates,

$$\{f, g\} = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial z}\left(p_i\frac{\partial g}{\partial p_i} - g\right) + \frac{\partial g}{\partial z}\left(p_i\frac{\partial f}{\partial p_i} - f\right)$$

2.4 Tangent bundle geometry

In this section we will present objects, operators and general constructions that we are able to define in tangent bundles. If Q is a smooth manifold, the tangent bundle TQ to Q possesses plenty of intrinsic structure: it is a smooth manifold with twice the dimension of Q; it is a vector bundle with projection $\tau_Q : TQ \to Q$, whose fibers are the tangent spaces T_qQ to each point q in the manifold Q; the sections of this bundle are the vector fields and there is a natural Lie bracket defined on them; tangent vectors of curves in Q live in the tangent bundle TQ.

Now, the double tangent bundle of Q, which is the tangent bundle of TQ will have an increased number of geometric and algebraic properties due to the fact that its base manifold is also a tangent bundle. We will introduce several concepts that will be used later on to formulate an intrinsic geometric version of Lagrangian mechanics (for more details see [LR89; YI73; LL66; HM20; CGM15] and see also [CCS87] for a similar discussion on the cotangent bundle).

2.4.1 The double tangent bundle

In this section we will define the vertical endomorphism, the Liouville vector field, and the canonical involution. Also, we will review some constructions on the theory of complete and vertical lifts in the tangent bundle (for more details, see [LR89] or [YI73]).

First let us recall the lift of vectors and vector fields on a manifold Q to its tangent bundle.

Complete and vertical lifts to the tangent bundle

Let $\tau_Q : TQ \to Q$ be the canonical projection of the tangent bundle.

Recall that the *complete* and *vertical lifts* of a function $f \in C^\infty(Q)$ are defined by

$$f^{\mathbf{c}}(v) = \langle df(q), v\rangle, \quad f^{\mathbf{v}}(v) = f \circ \tau_Q(v), \quad v \in T_qQ. \tag{2.4.1}$$

In other words, $f^{\mathbf{c}}$ is the fiberwise linear function on TQ induced by the 1-form df. Hence, the complete lift of f may be expressed on natural coordinates as

$$f^{\mathbf{c}}(q^i, v^i) = \frac{\partial f}{\partial q^i} v^i.$$

In what follows, $\mathcal{L}_X$ denotes the Lie derivative with respect to X. We will also use the *complete lift* of a vector field $X \in \mathfrak{X}(Q)$, which is the vector field $X^{\mathbf{c}} \in \mathfrak{X}(TQ)$ satisfying

$$X^{\mathbf{c}}(f \circ \tau_Q) = X(f) \circ \tau_Q, \text{ and } X^{\mathbf{c}}(\hat{\alpha}) = \widehat{\mathcal{L}_X \alpha}, \tag{2.4.2}$$

for any $f \in C^\infty(Q)$, $\alpha \in \Omega^1(Q)$ and where $\hat{\alpha} \in C^\infty(TQ)$ is the associated fiberwise linear function given by

$$\hat{\alpha}(v) = \langle \alpha(\tau_Q(v)), v\rangle, \quad v \in TQ.$$

We may also introduce the fiberwise quadratic function associated to a $(0, 2)$-tensor T on Q denoted by $T^{\mathfrak{q}} : TQ \to \mathbb{R}$ and defined by

$$T^{\mathfrak{q}}(v) = T_{\tau_Q(v)}(v, v), \quad v \in TQ. \tag{2.4.3}$$

Using equation (2.4.2) one may prove

Lemma 2.4.1. *If X is a vector field on Q and T is a $(0, 2)$-tensor on Q then*

$$X^{\mathbf{c}}(T^{\mathfrak{q}}) = (\mathcal{L}_X T)^{\mathfrak{q}}. \tag{2.4.4}$$

Proof. It is sufficient to prove the result for $T = \alpha \otimes \beta$, with α and β being 1-forms on Q. In this case,

$$T^{\mathfrak{q}} = \hat{\alpha} \cdot \hat{\beta}$$

and, using (2.4.2), it follows that

$$X^{\mathbf{c}}(T^{\mathfrak{q}}) = \widehat{\mathcal{L}_X \alpha} \cdot \hat{\beta} + \hat{\alpha} \cdot \widehat{\mathcal{L}_X \beta}.$$

This implies that

$$X^{\mathbf{c}}(T^{\mathfrak{q}}) = [\mathcal{L}_X\alpha \otimes \beta + \alpha \otimes \mathcal{L}_X\beta]^{\mathfrak{q}} = (\mathcal{L}_X T)^{\mathfrak{q}}.$$

□

We recall that the *vertical lift* of a vector $w \in T_qQ$ at $v \in T_qQ$ is given by

$$w_v^{\mathbf{v}} = \frac{d}{dt}_{|t=0}(v + tw), \quad \text{for } w \in T_qQ. \tag{2.4.5}$$

The concept of vertical lift may be generalized to vector fields in the following way: the *vertical lift of a vector field* is the vector field $X^{\mathbf{v}} \in \mathfrak{X}(TQ)$ satisfying

$$X^{\mathbf{v}}(f \circ \tau_Q) = 0, \text{ and } X^{\mathbf{v}}(\hat{\alpha}) = \langle \alpha, X \rangle \circ \tau_Q. \tag{2.4.6}$$

Similarly, we can define the *complete lift* of a k-form $\alpha \in \Omega^k(Q)$ to be the k-form $\alpha^{\mathbf{c}} \in \Omega^k(TQ)$ defined by

$$\alpha^{\mathbf{c}}(X_1^{\mathbf{c}}, \ldots, X_k^{\mathbf{c}}) = (\alpha(X_1, \ldots, X_k))^{\mathbf{c}}, \tag{2.4.7}$$

where $X_i \in \mathfrak{X}(Q)$. The expression above uniquely defines $\alpha^{\mathbf{c}}$ and, moreover,

$$d\alpha^{\mathbf{c}} = (d\alpha)^{\mathbf{c}} \tag{2.4.8}$$

and

$$i_{X^{\mathbf{c}}}\alpha^{\mathbf{c}} = (i_X\alpha)^{\mathbf{c}}. \tag{2.4.9}$$

On the other hand, the *vertical lift* of a k-form is simply the pullback by τ_Q, i.e.,

$$\alpha^{\mathbf{v}} = (\tau_Q)^*\alpha. \tag{2.4.10}$$

In local natural coordinates $(q^i, \dot{q}^i)$ on TQ, the expressions of the complete and vertical lifts of $X = X^i\frac{\partial}{\partial q^i}$ and $\alpha = \alpha_i dq^i$ are

$$\begin{gathered} X^{\mathbf{c}} = X^i\frac{\partial}{\partial q^i} + \frac{\partial X^i}{\partial q^j}\dot{q}^j\frac{\partial}{\partial \dot{q}^i}, \quad \alpha^{\mathbf{c}} = \frac{\partial \alpha_i}{\partial q^j}\dot{q}^j dq^i + \alpha_i d\dot{q}^i \\ X^{\mathbf{v}} = X^i\frac{\partial}{\partial \dot{q}^i}, \quad \alpha^{\mathbf{v}} = \alpha_i dq^i. \end{gathered} \tag{2.4.11}$$

To end this section, we will recall some useful identities satisfied by vertical and complete lifts. For any one form $\alpha \in \Omega^1(Q)$, one has that

$$\alpha^{\mathbf{v}}(Y^{\mathbf{c}}) = (\alpha(Y))^{\mathbf{v}}, \quad \alpha^{\mathbf{v}}(Y^{\mathbf{v}}) = 0. \tag{2.4.12}$$

Proposition 2.4.2. *The Lie bracket of complete and vertical lifts of vector fields satisfies the following relations*

$$\begin{aligned} &[X^{\mathbf{c}}, Y^{\mathbf{c}}] = [X,Y]^{\mathbf{c}}, \quad [X^{\mathbf{v}}, Y^{\mathbf{v}}] = 0, \\ &[X^{\mathbf{c}}, Y^{\mathbf{v}}] = [X^{\mathbf{v}}, Y^{\mathbf{c}}] = [X,Y]^{\mathbf{v}}. \end{aligned} \tag{2.4.13}$$

The canonical involution

Let Q be a smooth manifold of dimension n, $\tau_Q : TQ \to Q$ the canonical projection and TTQ the double tangent bundle to Q. Then, TTQ admits two vector bundle structures.

The first vector bundle structure is the canonical one with vector bundle projection $\tau_{TQ} : TTQ \to TQ$.

For the second vector bundle structure, the vector bundle projection is just the tangent map to τ_Q, that is, $T\tau_Q : TTQ \to TQ$ and the addition operation on the fibers is just the tangent map $T(+) : TTQ \times_{TQ} TTQ \to TTQ$ of the addition operation $(+) : TQ \times_Q TQ \to TQ$ on the fibers of τ_Q.

The canonical involution $\kappa_Q : TTQ \to TTQ$ is a vector bundle isomorphism (over the identity of TQ) between the two previous vector bundles. In fact, κ_Q is characterized by the following condition: let $\Phi : U \subseteq \mathbb{R}^2 \to Q$ be a smooth map, with U an open subset of $\mathbb{R}^2$

$$(t,s) \mapsto \Phi(t,s) \in Q.$$

Then,

$$\kappa_Q\left(\frac{d}{dt}\frac{d}{ds}\Phi(t,s)\right) = \frac{d}{ds}\frac{d}{dt}\Phi(t,s). \tag{2.4.14}$$

So, we have that κ_Q is an involution of TTQ, that is, $\kappa_Q^2 = id_{TTQ}$.

In fact, if $(q^i, \dot{q}^i)$ are canonical fibred coordinates on TQ and $(q^i, \dot{q}^i, v^i, \dot{v}^i)$ are the corresponding local fibred coordinates on TTQ then

$$\kappa_Q(q^i, \dot{q}^i, v^i, \dot{v}^i) = (q^i, v^i, \dot{q}^i, \dot{v}^i). \tag{2.4.15}$$

κ_Q may be characterized in a more intrinsic way, using the theory of complete and vertical lifts to TQ.

Indeed, if $X : Q \to TQ$ is a vector field on Q then

$$\kappa_Q \circ X^{\mathbf{c}} = TX, \quad \kappa_Q \circ X^{\mathbf{v}} = \widetilde{X}^{\mathbf{v}}, \tag{2.4.16}$$

where $TX : TQ \to TTQ$ is the tangent map to X (a section of the vector bundle $T\tau_Q$) and $\widetilde{X}^{\mathbf{v}} : TQ \to TTQ$ is the section of the vector bundle $T\tau_Q$ given by

$$\widetilde{X}^{\mathbf{v}}(u) = (T_q 0)(u) + X^{\mathbf{v}}(0(q)), \quad u \in T_q Q,$$

with $0 : Q \to TQ$ the zero section.

Note that, from (2.4.14), it follows that

$$\begin{aligned} T\kappa_Q \circ (X^{\mathbf{c}})^{\mathbf{c}} &= (X^{\mathbf{c}})^{\mathbf{c}} \circ \kappa_Q, \quad T\kappa_Q \circ (X^{\mathbf{v}})^{\mathbf{v}} = (X^{\mathbf{v}})^{\mathbf{v}} \circ \kappa_Q, \\ T\kappa_Q \circ (X^{\mathbf{c}})^{\mathbf{v}} &= (X^{\mathbf{v}})^{\mathbf{c}} \circ \kappa_Q, \quad T\kappa_Q \circ (X^{\mathbf{v}})^{\mathbf{c}} = (X^{\mathbf{c}})^{\mathbf{v}} \circ \kappa_Q, \end{aligned} \tag{2.4.17}$$

for $X \in \mathfrak{X}(Q)$.

As a consequence, we also deduce that

$$\begin{aligned} \kappa_Q^*((\alpha^{\mathbf{c}})^{\mathbf{c}}) &= (\alpha^{\mathbf{c}})^{\mathbf{c}} \circ \kappa_Q, \quad \kappa_Q^*((\alpha^{\mathbf{v}})^{\mathbf{v}}) = (\alpha^{\mathbf{v}})^{\mathbf{v}} \circ \kappa_Q, \\ \kappa_Q^*((\alpha^{\mathbf{c}})^{\mathbf{v}}) &= (\alpha^{\mathbf{v}})^{\mathbf{c}} \circ \kappa_Q, \quad \kappa_Q^*((\alpha^{\mathbf{v}})^{\mathbf{c}}) = (\alpha^{\mathbf{c}})^{\mathbf{v}} \circ \kappa_Q, \end{aligned} \tag{2.4.18}$$

for $\alpha \in \Omega^1(Q)$.

The vertical endomorphism and Liouville vector field

The vertical lift will allow us to construct the vertical endomorphism on TQ which plays an important role in the development of Lagrangian mechanics. The map $S : TTQ \to TTQ$ given by

$$S(X) = (T\tau_Q(X))^{\mathbf{v}}$$

is called the *vertical endomorphism* in Q. It may be considered as a $(1,1)$-tensor on TQ with local expression

$$S = \frac{\partial}{\partial \dot{q}^i} \otimes dq^i.$$

The vertical endomorphism satisfies the following properties related with complete and vertical lifts

$$SX^{\mathbf{c}} = X^{\mathbf{v}}, \quad SX^{\mathbf{v}} = 0, \quad \text{for } X \in \mathfrak{X}(Q). \tag{2.4.19}$$

An important property of the vertical endomorphism is given by the following theorem due to [LL66] (see also [LR89]):

Theorem 2.4.3. *The vertical endomorphism satisfies*

$$[SX, SY] = S([SX, Y]) + S([X, SY]). \qquad (2.4.20)$$

If $S^* : T^*TQ \to T^*TQ$ is the dual morphism of S defined by

$$\langle S^*\alpha_{v_q}, X_{v_q}\rangle = \langle \alpha_{v_q}, S\left(X_{v_q}\right)\rangle, \quad X_{v_q} \in T_{v_q}(TQ),$$

then we have that

$$S^*\alpha^{\mathbf{c}} = \alpha^{\mathbf{v}}, \quad S^*\alpha^{\mathbf{v}} = 0, \quad S^*(d\hat{\alpha}) = \alpha^{\mathbf{v}}, \quad \text{for } \alpha \in \Omega^1(Q). \qquad (2.4.21)$$

The *Liouville vector field* is the vector field $\Delta \in \mathfrak{X}(TQ)$ given by

$$\Delta(v_q) = \left.\frac{d}{dt}\right|_{t=0} (v_q + tv_q).$$

2.4.2 Second order differential equation vector fields

In this subsection, we will review the definition and main properties of SODE vector fields which are particularly important in mechanics.

A vector field $\Gamma \in \mathfrak{X}(TQ)$ is said to be a *second-order vector field* or simply a SODE vector field if it satisfies the property

$$(T\tau_Q \circ \Gamma)(v_q) = v_q, \quad \text{for all } v_q \in T_qQ.$$

This is equivalent to say that the vector field Γ on TQ is also a section of the vector bundle $T\tau_Q : TTQ \to TQ$, or also that

$$S(\Gamma) = \Delta.$$

In local coordinates, if $(q^i, \dot{q}^i)$ are local coordinates on TQ, then the local expression of a SODE Γ is of the form

$$\Gamma(q, \dot{q}) = \dot{q}^i \frac{\partial}{\partial q^i} + f^i(q, \dot{q}) \frac{\partial}{\partial \dot{q}^i},$$

where each f^i is a smooth function on TQ. Hence, $(q, \dot{q})$ is an integral curve of Γ if and only if

$$\ddot{q}^i = f^i(q, \dot{q})$$

and the curve q on the manifold Q is called a *trajectory* of Γ.

Example 2.4.4. In semi-Riemannian geometry, the geodesic vector field is a SODE with local expression

$$\Gamma^{\nabla}(q,\dot{q}) = \dot{q}^i \frac{\partial}{\partial q^i} - \Gamma^i_{kj}\dot{q}^k\dot{q}^j \frac{\partial}{\partial \dot{q}^i},$$

where $\{\Gamma^i_{kj}\}$ are the Christoffel symbols associated with the Levi-Civita connection ∇. $\triangle$

Proposition 2.4.5. *Let Γ be a SODE vector field. Then the Lie bracket of Γ with the Liouville vector field Δ satisfies*

$$S([\Delta,\Gamma]) = \Delta$$

and if $X \in \mathfrak{X}(Q)$, the Lie bracket with the vertical lift is simply

$$S([X^{\mathbf{v}},\Gamma]) = X^{\mathbf{v}}$$

Proof. If we compute the Lie bracket $[\Delta,\Gamma]$ in local coordinates, we may immediately check that its projection to TQ by $T\tau_Q : T(TQ) \to TQ$ is the vector

$$T\tau_Q([\Delta,\Gamma]) = \dot{q}^i \frac{\partial}{\partial q^i}.$$

Proceeding in the same way, the projection of $[X^{\mathbf{v}},\Gamma]$ to TQ is precisely X. Hence, the proposition follows. $\square$

SODE Exponential map

We will construct the exponential map of an arbitrary SODE in a similar fashion to that of semi-Riemannian geometry.

Let Γ be a SODE vector field on the tangent bundle TQ of a manifold Q. If $q_0 \in Q$ then we may consider the *exponential map at* q_0, which is defined as follows:

$$\exp^{\Gamma}_{h,q_0}(v_{q_0}) = \tau_Q(\phi^{\Gamma}_h(v_{q_0})), \quad v_{q_0} \in T_{q_0}Q$$

where $\{\phi^{\Gamma}_h\}$ is the flow of Γ for a sufficiently small non-negative number $h \geq 0$.

Denote also by

$$\exp^{\Gamma}_h(v_q) = (\tau_Q(v_q), \exp^{\Gamma}_{h,\tau_Q(v_q)}(v_q)) \subseteq Q \times Q, \quad q \in Q, v_q \in T_qQ. \quad (2.4.22)$$

Though one might intuitively convince himself that the exponential map is at least a local diffeomorphism and find numerous examples supporting it, we have not find in the literature no intrinsic proof whatsoever of this fact.

We postpone a complete proof until Chapter 4, where we will need this result in order to generalize further the concept of exponential map to the context of nonholonomic mechanics.

Jacobi fields for SODE's

Consider a vector field $X \in \mathfrak{X}(Q)$ on a manifold Q and denote by ϕ_t^X its flow. Fix an integral curve $\gamma_0 : I \to Q$ of X on a closed interval I, $0 \in I$. If $\gamma_0(0) = q$, then $\gamma_0(t) = \phi_t^X(q)$.

A *variation by integral curves of* X is a map $\Phi : (-\varepsilon, \varepsilon) \times I \to Q$ such that for each fixed $s \in (-\varepsilon, \varepsilon)$, $\Phi_s : I \to Q$ is an integral curve of X.

Given a SODE $\Gamma \in \mathfrak{X}(TQ)$, a variation by integral curves of Γ (see the definition (2.1.11) in Section 2.1.4) has the special form

$$\Phi(s,t) = \frac{\partial \varphi}{\partial t}(s,t),$$

where $\varphi(s,t)$ is a variation by trajectories in Q of Γ, i.e., for each fixed s, φ_s is a trajectory of Γ. If W denotes the infinitesimal variation vector field of φ (see (2.1.11) to recall the definition), then the infinitesimal variation vector field of the variation Φ is given by $t \in I \mapsto W^{\mathbf{c}}(t) \in T_{\frac{\partial \varphi}{\partial t}(0,t)}(TQ)$.

Definition 2.4.6. A vector field W along a trajectory of Γ denoted by γ_0 is said to be a *Jacobi field* if it is the infinitesimal variation vector field of a variation by trajectories.

We may prove the following proposition:

Proposition 2.4.7. *The vector field $\Gamma^{var} = (\kappa_Q)_*(\Gamma^{\mathbf{c}})$ called the variational vector field satisfies the following properties:*

1. *Γ^{var} is κ_Q-related to $\Gamma^{\mathbf{c}}$;*
2. *Γ^{var} is a SODE in TQ;*
3. *Γ^{var} is $T\tau_Q$-related to Γ;*
4. *The flow of Γ^{var} is $\phi_t^{var} = \kappa_Q \circ T\phi_t^{\Gamma} \circ \kappa_Q$.*

In fact, all properties may be immediately seen by inspecting the coordinate expression of Γ^{var}, which is

$$\Gamma^{var}(q,u,\dot q,\dot u)=\dot q^i\frac{\partial}{\partial q^i}+\dot u^i\frac{\partial}{\partial u^i}+f^i\frac{\partial}{\partial \dot q^i}+\left(u^j\frac{\partial f^i}{\partial q^j}+\dot u^j\frac{\partial f^i}{\partial \dot q^j}\right)\frac{\partial}{\partial \dot u^i},$$

provided $\Gamma=\dot q^i\frac{\partial}{\partial q^i}+f^i\frac{\partial}{\partial \dot q^i}$.

In semi-Riemannian geometry, it is well-known that Jacobi fields are solutions of the so-called Jacobi equation given by (2.1.12). In the general SODE case, we may find an expression which makes use of the *Ehresmann connection* associated to the SODE.

Theorem 2.4.8. *A vector field W along a trajectory of Γ denoted by γ_0 is a Jacobi field if and only if it satisfies the second-order differential equation*

$$\nabla\nabla W+\Phi(W)=0 \quad \textit{(Jacobi equation)}, \tag{2.4.23}$$

where Φ is the Jacobi endomorphism.

For the proof of any of the results in this section see [CGM15].

Recently, it was proven in [HM20] the relation between Jacobi fields, conjugate points and the SODE exponential map.

Definition 2.4.9. Let $q(t)$ be a trajectory of the SODE Γ, through q_0. If there exists a Jacobi field $W(t)$ not identically zero and such that $W(0)=W(h)=0$, then the point $q_1=q(h)$ is called a *conjugate point of q_0 along q*.

Then the authors prove the following result:

Proposition 2.4.10. *Let c be a trajectory of the SODE Γ joining two points q_0 and q_1 on Q with initial velocity $\dot c(0)=v_{q_0}$. Then q_1 is a conjugate point of q_0 along c if and only if the exponential map at v_{q_0} is singular.*

This result is analogous to what happens in Riemannian geometry (see [O'N83]), so it is evident that the strong interplay between these objects transcends the scope of Riemannian geometry.

2.4.3 Distributions

A central notion in the field of nonholonomic constraints on mechanical systems is that of a distribution. A *distribution* on a manifold Q is the assignment of a subspace $\mathcal{D}_q$ of T_qQ to each point $q\in Q$. We will just deal with *regular distributions*, which are those for which the vector space dimension of $\mathcal{D}_q$ is the same for all points $q \in Q$. Hence, a regular distribution is a vector subbundle of TQ and the dimension of each fiber $\mathcal{D}_q$ is called the *rank* of the distribution. We will denote its total space by $\mathcal{D}$, so that

$$\mathcal{D} = \bigcup_{q \in Q} \mathcal{D}_q.$$

Now, if $\mathcal{D}$ is a distribution on Q and there exists a submanifold $N \subseteq Q$ such that $T_n N = \mathcal{D}_n$ for all $n \in N$, then N is called an *integral manifold of* $\mathcal{D}$. A distribution $\mathcal{D}$ on the manifold Q is said to be *integrable* if any any point on Q is contained in an integral submanifold of $\mathcal{D}$.

The question of integrability has relevant consequences on mechanical systems. We will resume this discussion later with more detail, but for now let us just mention that if a mechanical system has its velocities constrained to lie in an integrable distribution, then the system may be seen as a *holonomic system*, which is a system with constrained configuration space (space of positions) $N \subseteq Q$. This is relevant because the dynamics of holonomic systems is completely understood as opposed to nonholonomic systems.

Fortunately, there is a practical criteria to decide whether a given distribution is integrable or not: involutivity. A distribution $\mathcal{D}$ is said to be *involutive* if for any pair X, Y of local sections of $\mathcal{D}$, their Lie bracket $[X, Y]$ is also a local section of $\mathcal{D}$. Our discussion culminates with Fröbenius theorem, which states the following:

Theorem 2.4.11. *A distribution is integrable if and only if it is involutive.*

2.5 Lie group actions

We end the introductory chapter with a brief introduction to Lie group actions, which appear naturally in mechanics and are often used to reduce the problem to a simpler dynamics which we can hopefully solve. We will restrict ourselves to the minimum amount of information needed to our purposes (for a complete treatment of the subject see, for instance, [Lee13; AM78; Hol+09; MR13]).

Given a Lie group and a manifold Q, consider a smooth map $\Phi : G \times Q \longrightarrow Q$ and denote by $g \cdot q \in Q$ the image of $\Phi(g, q)$. The map Φ is called a *left action* if:

1. $e \cdot q = q$ for any $q \in Q$, where $e \in G$ is the identity;
2. $(gh) \cdot q = g \cdot (h \cdot q)$, for any $g, h \in G$ and $q \in Q$.

A right action is defined analogously. Observe that the properties above imply that for every $g \in G$ the map $\Phi_g : Q \longrightarrow Q$ is a diffeomorphism.

Example 2.5.1. In the euclidean space $\mathbb{R}^n$, the group of invertible real matrices $GL(n)$ acts by matrix multiplication:

$$\Phi(A, q) = Aq.$$

△

Example 2.5.2. Given a complete vector field X on a manifold Q, its flow $\{\phi_t^X\}_{t\in\mathbb{R}}$ generates a Lie group action of $\mathbb{R}$ on Q by

$$\Phi(t, q) = \phi_t^X(q).$$

△

Let $\mathfrak{g}$ the Lie algebra of the group G. To each element ξ in the Lie algebra, we may associate a vector field on Q called the *infinitesimal generator associated to* ξ, which we denote by ξ_Q and defined by

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

where $\exp : \mathfrak{g} \to G$ is the Lie group exponential.

Proposition 2.5.3. *The map $\mathfrak{g} \to \mathfrak{X}(Q)$, mapping $\xi \mapsto \xi_Q$ is linear and satisfies*

$$[\xi, \eta] = -[\xi_Q, \eta_Q].$$

Thus, it is a Lie algebra anti-homomorphism.

In mechanics, we face many times tangent and cotangent lifted actions. Given a Lie group action Φ on Q the *tangent lift* Φ^T *of* Φ is the action on TQ given by

$$\Phi^T(g, v_q) = T_q\Phi_g(v_q).$$

Likewise, the *cotangent lift* Φ^{T^*} *of* Φ is the action on T^*Q given by

$$\Phi^{T^*}(g, \alpha_q) = \left(T^*_{g\cdot q}\Phi_{g^{-1}}\right)(\alpha_q). \tag{2.5.1}$$

The cotangent lift of any action Φ enjoys a remarkable property: for any $g \in G$, the map $\Phi_g^{T^*} : T^*Q \to T^*Q$ is both a symplectic and a Poisson map with respect to the corresponding canonical symplectic structure on T^*Q. In fact, $\Phi_g^{T^*}$ is an exact symplectic map, i.e., it is a diffeomorphism satisfying $(\Phi_g^{T^*})^*\theta_Q = \theta_Q$.

Chapter 3

Introduction to geometric mechanics

In this chapter we will cover a fairly modern treatment of mechanics in a differential geometric language. The main gain by proceeding this way is that all constructions are coordinate independent, globally defined and, many times, ready to be generalized to broader contexts. We will define the main concepts and results that lay the foundations of the main contributions submitted to this thesis.

Like in the last chapter, the reader who is familiarized with geometric mechanics might proceed to Chapter 4.

We will review the geometric formulation of Lagrangian mechanics, paying special attention to mechanical systems arising from a mechanical Lagrangian, i.e., of the form $L = K - V$, where K is the kinetic energy associated with a semi-Riemannian metric on the configuration space and $V \in C^\infty(Q)$ is the potential energy. We will discuss the interplay of trajectories and the semi-Riemannian structure. Next, we review Hamiltonian mechanics and mechanical systems under the action of external forces. Finally, we are left with probably the most relevant sections for our purposes which are the review of nonholonomic mechanics and discrete Lagrangian mechanics (for more details see [AM78; LR89; Blo15; MR13; Cra83; JM70; LLV19b; LD96; CMR01; Cor+03; LM95; VF72; VM94; Koi92; MW01; PC09; DA18; CM01; MP06; FID08; BZ15; FBO12; Cel+19; Igl+08; MV19; Cor02]).

3.1 Lagrangian mechanics

A *mechanical system* is a pair formed by a smooth manifold Q called the *configuration space* and a smooth function $L : TQ \to \mathbb{R}$ on its tangent bundle called the *Lagrangian* (see [LR89], [AM78]). If the system is not subjected to any constraint or external forces, a *trajectory* of the mechanical system is a solution of the *Euler-Lagrange equations*, whose expression on natural coordinates relative to a chart (q^i) for Q and the induced coordinates $(q^i, \dot{q}^i)$ on TQ is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0. \tag{3.1.1}$$

Example 3.1.1. Consider the example of a mechanical Lagrangian function in the euclidean space $Q = \mathbb{R}^n$, which is given by

$$L(q^i, \dot{q}^i) = \frac{1}{2}m\|\dot{q}\|^2 - V(q^i),$$

where $\|\dot{q}\|$ is the euclidean norm, $V : Q \to \mathbb{R}$ is called the *potential energy* and $m > 0$ is the mass of a particle. Then, Euler-Lagrange equations give the second-order differential equations

$$m\ddot{q}^i = -\frac{\partial V}{\partial q^i},$$

which are Newton's second law of mechanics for a system subjected to a conservative force in the euclidean space. △

Example 3.1.2. Let us consider an example of a simple Lagrangian system on a non-euclidean manifold. This is the example of a simple pendulum with configuration manifold $Q = \mathbb{S}^1$. Let $L : TQ \to \mathbb{R}$ be the Lagrangian function defined by

$$L(\theta, \dot{\theta}) = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta,$$

where m is the mass of the pendulum, l is the length of the rod and g is the acceleration of gravity. Then, Euler-Lagrange equations imply that

$$\ddot{\theta} = -\frac{g}{l}\sin\theta.$$

It is very common in physics to approximate the last equations using the small angle hypothesis under which $\sin\theta \approx \theta$. The equations obtained considering this substitution are the Euler-Lagrange equations relative to the Lagrangian function $L : TQ \to \mathbb{R}$ given by

$$L(\theta, \dot{\theta}) = \frac{1}{2} m l^2 \dot{\theta}^2 - mgl \frac{\theta^2}{2}.$$

The Lagrangian system determined by this Lagrangian is called the *harmonic oscillator*. △

As it is well-known, the *Hamilton's principle* asserts that the trajectories of mechanical systems are obtained by minimizing the action functional defined over curves with fixed end-points. Denote the set of twice differentiable curves with fixed end-points $q_0, q_1 \in Q$ by

$$C^2(q_0, q_1) = \{q : [0, T] \longrightarrow Q | \; q(\cdot) \text{ is } C^2, q(0) = q_0, q(T) = q_1\}.$$

We have that:

Definition 3.1.3 (Hamilton's principle)**.** A curve $q : I \to Q$ in $C^2(q_0, q_1)$ is a trajectory of the mechanical system determined by a Lagrangian function $L : TQ \to \mathbb{R}$ if it is a critical value of the *action functional* $\mathcal{S} : C^2(q_0, q_1) \longrightarrow \mathbb{R}$ defined by

$$\mathcal{S}(q(\cdot)) = \int_0^T L(q(t), \dot{q}(t)) \; dt,$$

among all curves in $C^2(q_0, q_1)$.

A necessary and sufficient condition for the curve q to be a trajectory is that it satisfies Euler-Lagrange equations in any coordinate chart (see [AM78],[LR89],[Blo15],[MR13]). We will not proof it here since it is a widely well-known result coming from the literature in calculus of variations.

We can give a coordinate-free set of equations equivalent to Euler-Lagrange equations (3.1.1) (cf. [Cra83]).

Proposition 3.1.4. *A curve $q(t)$ is a solution of Euler-Lagrange equations* (3.1.1) *if and only if*

$$X^{\mathbf{c}}(L)(q, \dot{q}) - \frac{d}{dt} \left(X^{\mathbf{v}}(L)(q, \dot{q}) \right) = 0, \quad \forall \; X \in \mathfrak{X}(Q). \tag{3.1.2}$$

Proof. Let $X \in \mathfrak{X}(Q)$ be a vector field locally given by

$$X = X^i \frac{\partial}{\partial q^i}.$$

Then

$$X^{\mathbf{c}}(L) = X^i \frac{\partial L}{\partial q^i} + \dot{q}^j \frac{\partial X^i}{\partial q^j} \frac{\partial L}{\partial \dot{q}^i} \quad \text{and} \quad X^{\mathbf{v}}(L) = X^i \frac{\partial L}{\partial \dot{q}^i}.$$

Thus equations (3.1.2), along a curve $q(t)$ have the local form

$$X^i \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) \right) = 0.$$

Hence, we deduce that (3.1.2) is satisfied for all vector fields if and only if the curve $q(t)$ is a solution of Euler-Lagrange equations. □

We will see now how the Lagrangian dynamics may be seen from the perspective of symplectic geometry. A Lagrangian function $L : TQ \to \mathbb{R}$ is *regular* if the Hessian matrix $\text{Hess}(L) := \left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \right)$ is non-singular.

Let us define the *Poincaré-Cartan 1-form* associated to a Lagrangian function L to be the 1-form $\theta_L \in \Omega^1(TQ)$ given by

$$\theta_L = S^*(dL),$$

where $S : TTQ \to TTQ$ is the vertical endomorphism of the tangent bundle. Using the canonical coordinates of the tangent bundle, θ_L may be written as

$$\theta_L(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}^i} dq^i.$$

Now, the *Poincaré-Cartan 2-form* associated to L is defined from the corresponding Poincaré-Cartan 1-form as the 2-form $\omega_L \in \Omega^2(TQ)$ with

$$\omega_L = -d\theta_L.$$

In the canonical tangent bundle coordinates we obtain

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

It is natural to wonder when is ω_L a symplectic form on the tangent bundle. This will happen when the corresponding Lagrangian function is regular.

Proposition 3.1.5. *Let L be a Lagrangian function. Then, L is regular if and only if its Poincaré-Cartan 2-form ω_L is a symplectic form.*

Proof. We give an intrinsic proof of the proposition after Lemma 3.3.2. At the moment just observe that ω_L is symplectic if and only if $\omega_L^n \neq 0$. But we have that

$$\omega_L^n = c \det\left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\right) dq^1 \wedge \cdots \wedge dq^n \wedge d\dot{q}^1 \wedge \cdots \wedge d\dot{q}^n,$$

where c is a non-zero constant function. Thus, ω_L is symplectic if and only if L is regular. $\square$

Suppose from now on that L is regular. Under the regularity assumption, by non-degeneracy of the symplectic form ω_L, there is a unique vector field Γ_L satisfying the geometric equation

$$i_{\Gamma_L}\omega_L = dE_L. \tag{3.1.3}$$

where $E_L = \Delta(L) - L$ is the *energy function* and Δ is the Liouville vector field on TQ. We have the following result describing the situation:

Proposition 3.1.6. *Let L be a regular Lagrangian function. Then there is a unique vector field Γ_L satisfying equation* (3.1.3) *called the Lagrangian vector field. Moreover, Γ_L is a SODE vector field on TQ and its trajectories satisfy the Euler-Lagrange equations* (3.1.1)*.*

Proof. First note that the vector field Γ_L is well-defined since ω_L is non-degenerate. To prove that Γ_L is a SODE, it is enough to prove that $S \circ \Gamma_L = \Delta$. Note that

$$i_\Delta \omega_L = -\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\dot{q}^j dq^i = -S^*(dE_L).$$

Suppose that the vector field Γ_L is locally given by the expression

$$\Gamma_L(q,\dot{q}) = f^i(q,\dot{q})\frac{\partial}{\partial q^i} + g^i(q,\dot{q})\frac{\partial}{\partial \dot{q}^i}.$$

Then,

$$i_{S\circ\Gamma_L}\omega_L = -\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} f^j dq^i = -S^*(i_{\Gamma_L}\omega_L).$$

Thus, by equation (3.1.3) defining the vector field Γ_L and by the non-degeneracy of ω_L we deduce that $S \circ \Gamma_L = \Delta$. Equivalently, the local expression of Γ_L is

$$\Gamma_L(q,\dot{q}) = \dot{q}^i\frac{\partial}{\partial q^i} + g^i(q,\dot{q})\frac{\partial}{\partial \dot{q}^i}.$$

Moreover, if $q(t)$ is a trajectory of Γ_L, then the second derivative of q^i satisfies $\ddot{q}^i(t) = g^i(q(t), \dot{q}(t))$ and from (3.1.3) it also satisfies the local equations

$$\frac{\partial^2 L}{\partial \dot{q}^i \partial q^j}\dot{q}^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\ddot{q}^j - \frac{\partial L}{\partial q^i} = 0,$$

which are equivalent to Euler-Lagrange equations (3.1.1). □

As an interesting result we have that

Lemma 3.1.7. *Let L be a regular Lagrangian function. Then Γ_L satisfies*

$$\mathcal{L}_{\Gamma_L}\theta_L = dL. \tag{3.1.4}$$

Proof. Since the Lagrangian vector field is a SODE, we have that

$$\langle \theta_L, \Gamma_L \rangle = \langle dL, \Delta \rangle.$$

Then, using Cartan's magic formula and the definition of Lagrangian energy we deduce

$$\mathcal{L}_{\Gamma_L}\theta_L = i_{\Gamma_L} d\theta_L + d(i_{\Gamma_L}\theta_L) = -dE_L + d(\Delta(L)) = dL.$$

□

We have conservation of energy and, in addition, the flow of Γ_L is a symplectomorphism.

Theorem 3.1.8. *If L is a regular Lagrangian function, the flow of Γ_L preserves the symplectic form ω_L and the energy E_L is invariant along its integral curves.*

Proof. This is a consequence of Proposition 2.2.14 and the fact that Γ_L is the Hamiltonian vector field of E_L with respect to the symplectic form ω_L. □

3.2 Mechanical Lagrangian systems

In this section we will restrict our attention to Lagrangian functions of the form $L = K - V$, where K is the kinetic energy associated with a semi-Riemannian metric h and $V \in C^\infty(Q)$ is the potential energy. We will take a deep insight into the underlying geometry behind this case.

Definition 3.2.1. The Lagrangian function $L_h : TQ \to \mathbb{R}$ given by

$$L_h(v_q) = \frac{1}{2} h(v_q, v_q), \quad \text{with} \quad v_q \in T_qQ,$$

where h is a semi-Riemannian metric is called a *kinetic Lagrangian* function or the *kinetic energy*. A *mechanical Lagrangian* function is a Lagrangian $L_{(h,V)} : TQ \to \mathbb{R}$ of the type

$$L_{(h,V)}(v_q) = L_h(v_q) - V(q),$$

where the function $V : Q \to \mathbb{R}$ is the *potential energy*.

We have already found examples of mechanical Lagrangian functions in Examples 3.1.1 and 3.1.2. Indeed, mechanical Lagrangian systems are the prototype of conservative physical systems we may find in our universe.

Given a pseudo-Riemannian metric h, define the *gradient vector field with respect to* h to be the vector field

$$\text{grad}_h V = \sharp_h(dV), \tag{3.2.1}$$

where $\sharp_h : \Omega^1(Q) \to \mathfrak{X}(Q)$ is the inverse isomorphism of $\flat_h$.

Theorem 3.2.2. *Let $L_{(h,V)}$ be a mechanical Lagrangian function on Q associated to a pseudo-Riemannian metric h and a potential energy V. The curve $c : I \to Q$ is a trajectory of $\Gamma_{L_{(h,V)}}$ if and only if it satisfies the equations*

$$\nabla^h_{\dot{c}} \dot{c} = -\text{grad}_h V \circ c, \tag{3.2.2}$$

where ∇^h is the Levi-Civita connection associated with the pseudo-Riemannian metric h.

Remark 3.2.3. In particular, when the potential energy V vanishes, the curve c is the solution of Euler-Lagrange equations for $L_{(h,V)}$ if and only if it is a geodesic with respect to the Levi-Civita connection ∇^h.

Proof. We will give here an intrinsic proof. For a proof in coordinates see, for example, [AM78].

Let $c : I \to Q$ be a trajectory of $\Gamma_{L_{(h,V)}}$. Given any $X \in \mathfrak{X}(Q)$, we will pair the geometric equation which defines $\Gamma_{L_{(h,V)}}$ with the complete lift $X^{\mathbf{c}}$ of the vector field X. In fact,

$$\langle \left(i_{\Gamma_{L_{(h,V)}}} \omega_{L_{(h,V)}} - dE_{L_{(h,V)}} \right)(u), X^{\mathbf{c}}(u) \rangle = 0.$$

Using the skew-symmetry of $\omega_{L_{(h,V)}}$ and the fact that $E_{L_{(h,V)}} = L_h + V \circ \tau_Q$, we get

$$-\langle i_{X^{\mathbf{c}}}\omega_{L_{(h,V)}}, \Gamma_{L_{(h,V)}}\rangle - X^{\mathbf{c}}(L_h) - X^{\mathbf{c}}(V \circ \tau_Q) = 0.$$

We will use now the equations given in (A.0.5) and in Lemma A.0.1 in Appendix A.1, namely the equations

$$i_{X^{\mathbf{c}}}\omega_{L_{(h,V)}} = d(\widehat{\flat_h(X)}) + 2\theta_{L_{(\nabla^h X)}}, \quad X^{\mathbf{c}}(L_h) = L_{\mathcal{L}_X h},$$

where $(\nabla^h X)$ is the $(0,2)$-tensor field defined by

$$(\nabla^h X)(Y, Z) = h(\nabla^h_X Y, Z),$$

(see equation (A.0.2) in Appendix A.1) and $\widehat{\flat_h(X)}$ is the fiberwise linear function on TQ induced by the 1-form $\flat_h(X)$.

Then, we deduce that

$$-\langle d(\widehat{\flat_h(X)}) + 2\theta_{L_{(\nabla^h X)}}, \Gamma_{L_{(h,V)}}\rangle - L_{\mathcal{L}_X h} - X^{\mathbf{c}}(V \circ \tau_Q) = 0.$$

However, by Lemma A.0.2, we have that $L_{\mathcal{L}_X h} = 2L_{(\nabla^h X)}$ and using that the vector field $\Gamma_{L_{(h,V)}}$ is a SODE, it follows that $S \circ \Gamma_{L_{(h,V)}} = \Delta$. So, we have that

$$\langle \theta_{L_{(\nabla^h X)}}, \Gamma_{L_{(h,V)}}\rangle = \Delta(L_{(\nabla^h X)}) = 2L_{(\nabla^h X)}.$$

Hence, we have deduced the following equation

$$\Gamma_{L_{(h,V)}}(\widehat{\flat_h(X)}) = 2L_{(\nabla^h X)} - X^{\mathbf{c}}(V \circ \tau_Q),$$

which evaluated at a point $u \in TQ$ is equivalent to

$$\Gamma_{L_{(h,V)}}(u)(\widehat{\flat_h(X)}) = h(\nabla^h_u X, u) - \langle dV, X\rangle \circ \tau_Q(u).$$

Then, evaluating the last equation over the curve $\dot{c}$, noting that $\Gamma_{L_{(h,V)}}(\dot{c})$ is just $\ddot{c}$ and using (3.2.1), we deduce

$$\ddot{c}\left(\widehat{\flat_h(X)}\right) = h(\nabla^h_{\dot{c}} X, \dot{c}) - h(\operatorname{grad}_h V \circ c, X \circ c).$$

Moreover, as the connection is compatible with the metric, the right-hand side is just

$$\ddot{c}\left(\widehat{\flat_h(X)}\right) = \frac{d}{dt}\left(\widehat{\flat_h(X)}(\dot{c})\right) = h(\nabla^h_{\dot{c}} X, \dot{c}) + h(X \circ c, \nabla^h_{\dot{c}} \dot{c})$$

and the previous expression reduces to

$$h(X\circ c, \nabla^h_{\dot c}\dot c + \mathrm{grad}_h V\circ c) = 0.$$

Since X is an arbitrary vector field and h is non-degenerate we proved that

$$\nabla^h_{\dot c}\dot c + \mathrm{grad}_h V\circ c = 0.$$

Conversely, let c be a solution of (3.2.2). Then c is a trajectory of the SODE vector field Γ defined by equation (3.2.2) given by

$$\Gamma = \Gamma^{\nabla^h} + (\mathrm{grad}_h V)^{\mathbf{v}},$$

where Γ^{∇^h} is the geodesic vector field associated with the Levi-Civita connection of the semi-Riemannian metric h.

By reversing all the arguments before, we deduce that

$$\Gamma_{L_{(h,V)}}(\dot c(t))(\widehat{\flat_h(X)}) = \Gamma(\dot c(t))(\widehat{\flat_h(X)}), \quad \forall X\in\mathfrak{X}(Q).$$

Therefore, since the $\flat_h$ is an isomorphism of $C^\infty(Q)$-modules, the action of $\Gamma_{L_{(h,V)}}$ and Γ over fiberwise linear functions matches. To establish that the two vector fields are the same, it remains to show that

$$\Gamma_{L_{(h,V)}}(\dot c(t))(f\circ\tau_Q) = \Gamma(\dot c(t))(f\circ\tau_Q), \quad \forall f\in C^\infty(Q).$$

But this follows immediately from the fact that Γ and $\Gamma_{L_{(h,V)}}$ are SODE vector fields since

$$\Gamma_{L_{(h,V)}}(\dot c(t))(f\circ\tau_Q) = \dot c(t)(f) = \Gamma(\dot c(t))(f\circ\tau_Q).$$

□

We finish this section with Jacobi theorem, stating that trajectories of mechanical systems may in fact be seen as geodesics with respect to a modified metric (cf., for instance, [AM78] for more details). In that sense, consider a Riemannian metric g and let the energy of the mechanical system determined by the Lagrangian $L_{(g,V)}$ be given by the function $E_{L_{(g,V)}}: TQ\to\mathbb{R}$ defined by

$$E_{L_{(g,V)}}(v) = \frac{1}{2}g(v,v) + V\circ\tau_Q(v), \quad v\in TQ.$$

Now, take $e \in \mathbb{R}$ to be a constant value of the energy such that the set

$$U_e = \{q \in Q \mid e > V(q)\}$$

is a non-empty subset of Q. In fact, U_e is an open subset of Q and if it is non-empty, it inherits the smooth manifold structure of Q. We can consider on it the *Jacobi metric* given by

$$g_e = (e - V)g. \tag{3.2.3}$$

We have the following result which we will call the *Maupertuis-Jacobi principle*:

Theorem 3.2.4. *The trajectories of the mechanical Lagrangian $L_{(g,V)}$ with energy e are the same as geodesics of the Jacobi metric up to a reparametrization, with unit energy.*

Proof. The proof consists of three steps. The first one is to move the problem into the cotangent bundle, where we have a single canonical symplectic structure. There we obtain the Hamiltonian function $H_{(g,V)}$ associated to $L_{(g,V)}$ and the Hamiltonian function H_{g_e} associated to the Lagrangian function L_{g_e}. The next step is to observe that

$$H_{(g,V)}(\alpha) = e \quad \text{if and only if} \quad H_{g_e}(\alpha) = 1$$

for $\alpha \in T^*U_e$. The third step is the most elaborated. We must prove that the fact that $H^{-1}_{(g,V)}(e) = H^{-1}_{g_e}(1) = \Sigma$ implies that the respective Hamiltonian vector fields are co-linear and, hence, their integral curves are the same up to reparametrization. Indeed, since Σ is a manifold with codimension 1 we have that the distribution along Σ

$$T\Sigma^{\perp} = \{v \in T_{\Sigma}(T^*Q) | i_v\omega(w) = 0, \ \forall w \in T\Sigma\}$$

must be one-dimensional at each point of Σ, by Lemma 2.2.2. Also, it is clear that

$$\langle dH_{(g,V)}(\alpha), w_\alpha \rangle = \langle dH_{g_e}(\alpha), w_\alpha \rangle = 0,$$

for all $\alpha \in \Sigma$ and $w_\alpha \in T\Sigma$. Hence the Hamiltonian vector fields $X_{H_{(g,V)}}$ and $X_{H_{g_e}}$ lie in $T\Sigma^{\perp}$. Therefore, they must be parallel. □

3.3 Hamiltonian mechanics

In this section we will briefly describe the standard Hamiltonian mechanics formalism. We will see that in very special occasions, the study of Euler-Lagrange equations is related to the study of a Hamiltonian vector field with respect to the canonical symplectic structure on the cotangent bundle T^*Q. In these cases, the Lagrangian flow automatically inherits all the distinguished properties that Hamiltonian vector fields possess.

Given a Hamiltonian function $H : T^*Q \to \mathbb{R}$ we define the corresponding Hamiltonian vector field X_H by

$$\imath_{X_H}\omega_Q = dH.$$

The integral curves of X_H are determined by *Hamilton's equations*, whose local expression is

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

Example 3.3.1. Consider the euclidean space $Q = \mathbb{R}^n$ and the Hamiltonian function given by

$$H(q^i, p_i) = \sum_{i=1}^{n} \frac{p_i^2}{2m} + V(q^i).$$

Then, Hamilton's equation in this case give

$$\frac{dq^i}{dt} = \frac{p_i}{m}, \quad \frac{dp_i}{dt} = -\frac{\partial V}{\partial q^i},$$

which are equivalent to Newton's second law. In fact, $p_i = m\dot{q}^i$ is usually known in physics as the *linear momentum* of the system. So, Hamilton's equations describe the evolution in time of linear momentum. △

Given a Lagrangian function L, we can define the *Legendre transformation* $\mathbb{F}L : TQ \to T^*Q$ by

$$\langle \mathbb{F}L(u_q), v_q\rangle = \left.\frac{d}{dt}\right|_{t=0} L(u_q + tv_q)$$

and if L is regular, its Legendre transformation is a local diffeomorphism. In local coordinates

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

We will see now how the Lagrangian and Hamiltonian descriptions are related.

Lemma 3.3.2. *Let L be a regular Lagrangian function. Then, the associated Poincaré-Cartan 1-form and 2-form denoted by θ_L and ω_L, respectively, satisfy*

$$\theta_L = \mathbb{F}L^*\theta_Q, \quad \omega_L = \mathbb{F}L^*\omega_Q.$$

Proof. We must prove the equality for the 1-forms. The relation between the 2-forms follows immediately from the first one.

It is enough check the equality on vector fields of the form $X^{\mathbf{v}}$ and $X^{\mathbf{c}}$ with $X \in \mathfrak{X}(Q)$, since they generate the tangent bundle TTQ.

Moreover, we only need to establish the first formula regarding the Poincaré-Cartan 1-from. The second one follows by applying the exterior derivative.

Let $X \in \mathfrak{X}(Q)$, using (2.4.19) we have that

$$\langle \theta_L, X^{\mathbf{v}} \rangle = \langle dL, S(X^{\mathbf{v}}) \rangle = 0.$$

Similarly,

$$\langle \mathbb{F}L^*\theta_Q, X^{\mathbf{v}} \rangle = \langle \theta_Q, (\mathbb{F}L)_*(X^{\mathbf{v}}) \rangle = 0.$$

Indeed, by definition of the canonical 1-form, given $\alpha_q \in T_q^*Q$

$$\langle \theta_Q(\alpha_q), (\mathbb{F}L)_*(X^{\mathbf{v}})(\alpha_q) \rangle = \langle \alpha_q, T_{\alpha_q}\pi_Q((\mathbb{F}L)_*(X^{\mathbf{v}})(\alpha_q)) \rangle.$$

Moreover given $v_q \in T_qQ$ we have that

$$(T_{\mathbb{F}L(v_q)}\pi_Q)\left((T_{v_q}\mathbb{F}L)(X^{\mathbf{v}}(v_q))\right) = T_{v_q}(\pi_Q \circ \mathbb{F}L)(X^{\mathbf{v}}).$$

But the Legendre transform is a fibre-preserving map, i.e., $\pi_Q \circ \mathbb{F}L = \tau_Q$. Thus the last expression vanishes since the vertical lift projects to the zero vector field.

On complete lifts the situation is very similar with the necessary adaptations. On one hand, using again (2.4.19), we obtain

$$\langle \theta_L(v_q), X^{\mathbf{c}}(v_q) \rangle = \langle dL(v_q), S(X^{\mathbf{c}})(v_q) \rangle = \langle dL(v_q), X^{\mathbf{v}}(v_q) \rangle = \langle \mathbb{F}L(v_q), X(q) \rangle.$$

On the other hand, we have that

$$\begin{aligned} \langle \mathbb{F}L^*\theta_Q(v_q), X^{\mathbf{c}}(v_q) \rangle &= \langle \theta_Q(\mathbb{F}L(v_q)), (T_{v_q}\mathbb{F}L)(X^{\mathbf{c}}(v_q)) \rangle \\ &= \langle \mathbb{F}L(v_q), (T_{v_q}\tau_Q)(X^{\mathbf{c}}(v_q)) \rangle \\ &= \langle \mathbb{F}L(v_q), X(q) \rangle, \end{aligned}$$

where the last line follows because the complete lift projects to X. □

As a corollary to the preceding lemma we deduce that the Poincaré-Cartan 2-form is symplectic if and only if the Legendre transform is a diffeomorphism, which is equivalent to the regularity of L. This is a result we had already proved.

Proposition 3.3.3. *Let L be a regular Lagrangian function and E_L its Lagrangian energy. Consider the Hamiltonian function $H = E_L \circ (\mathbb{F}L)^{-1}$. Then the vector fields Γ_L and X_H are $\mathbb{F}L$-related and, thus, its integral curves.*

Proof. Given a regular Lagrangian function L, let Γ_L be the Lagrangian vector field and let H be the Hamiltonian function $H = E_L \circ (\mathbb{F}L)^{-1}$. By definition, the vector field Γ_L satisfies the equation

$$i_{\Gamma_L}\omega_L = dE_L.$$

Applying the pullback by $(\mathbb{F}L)^{-1}$ to both sides of the equation we obtain

$$i_{(\mathbb{F}L)_*\Gamma_L}((\mathbb{F}L)^{-1})^*\omega_L = d(E_L \circ (\mathbb{F}L)^{-1}).$$

Using Lemma 3.3.2 and the definition of the Hamiltonian function H we conclude

$$i_{(\mathbb{F}L)_*\Gamma_L}\omega_Q = dH.$$

Therefore, by non-degeneracy of the canonical symplectic form ω_Q, the Hamiltonian vector field X_H must be $(\mathbb{F}L)_*\Gamma_L$. □

An extensive account of this subject is contained in [AM78; LR89], for instance.

Example 3.3.4. Let us now examine the example of a particle in a magnetic field. Let the configuration manifold be $Q = \mathbb{R}^3$ and suppose there is on Q a *magnetic field*, which is a vector field $Y_B \in \mathfrak{X}(Q)$ determined by the closed two-form $B \in \Omega^2(Q)$ in such a way that

$$i_{Y_B}(dx \wedge dy \wedge dz) = dB.$$

In addition, suppose that B is exact so there exists a one-form $A \in \Omega^1(Q)$ such that $B = dA$ called the *magnetic potential*.

Take the Hamiltonian function $H : T^*Q \to \mathbb{R}$ given by

$$H(q,p) = \frac{1}{2m}\left\| p - \frac{e}{c}A(q) \right\|^2, \tag{3.3.1}$$

where the norm is with respect to the euclidean metric, m is the mass of the particle, e is its electric charge and c is the speed of light. The Hamiltonian system associated to H is $\mathbb{F}L$-related to the Lagrangian system determined by the Lagrangian function $L : TQ \to \mathbb{R}$ given by

$$L(q, \dot{q}) = \frac{1}{2} m \|\dot{q}\|^2 + \frac{e}{c} \langle A(q), \dot{q} \rangle.$$

$\triangle$

The next example shows that the duality between Hamiltonian and Lagrangian formalisms exists even outside the scope of mechanics.

Example 3.3.5. In mathematical biology, *Lotka-Volterra* equations model the growth of two animal species. If $u(t)$ is the number of predators in time and $v(t)$ is the number of preys over time, then its evolution is given by

$$\dot{u} = u(v - \alpha), \quad \dot{v} = v(\beta - u),$$

where $\alpha, \beta > 0$ are constants. This may not be immediately recognized as a Hamiltonian system but it is indeed one. Let $Q = \mathbb{R}$ be the base manifold and let the cotangent bundle be $T^*Q \simeq \mathbb{R}^2$. Under the change of variables $u = e^q$ and $v = e^p$, Lotka-Volterra equations are equivalent to the Hamiltonian equations with respect to the canonical symplectic structure $\omega_Q = dq \wedge dp$ and the Hamiltonian function given by

$$H(q, p) = e^q - \beta q + e^p - \alpha p.$$

The corresponding Lagrangian formulation is obtained via the Legendre transformation $w = e^p - \alpha$ so that

$$L(q, w) = (w + \alpha)(\ln(w + \alpha) - 1) - e^q + \beta q.$$

$\triangle$

Not every Hamiltonian system has a corresponding Lagrangian dynamics as we will see in the next example.

Example 3.3.6. Consider the Hamiltonian function on $T^*\mathbb{R}^3$ given by

$$H(q, p) = \frac{1}{2} \left((p_1 + y p_3)^2 + (p_2)^2 \right).$$

This Hamiltonian function appears when one applies the Pontryagin Maximum Principle (see [Blo15]) in order to study the sub-Riemannian geodesics on $\mathbb{R}^3$ subjected to the constraint $\dot{z} = y\dot{x}$. It turns out that the trajectories of this Hamiltonian system, i.e., the projection to $\mathbb{R}^3$ of its integral curves, are the so-called *normal sub-Riemannian geodesics* minimizing the sub-Riemannian length

$$L(c) = \int_0^1 (\dot{x}^2 + \dot{y}^2)\, dt,$$

among all curves c in $\mathbb{R}^3$ satisfying the constraint $\dot{z} = y\dot{x}$. However, this Hamiltonian function is not regular since the Hessian matrix of H given by

$$\mathrm{Hess}(H) = \begin{pmatrix} 1 & 0 & y \\ 0 & 1 & 0 \\ y & 0 & y^2 \end{pmatrix}$$

has zero determinant everywhere and, thus, there is not a corresponding Lagrangian vector field which is $\mathbb{F}L$-related to the Hamiltonian vector field associated to H. $\triangle$

3.3.1 Mechanical Hamiltonian systems

Given a Riemannian metric g and a potential energy function V on the manifold Q, we may consider the Hamiltonian function $H_{(g,V)} : T^*Q \to \mathbb{R}$ given by

$$H_{(g,V)}(\alpha_q) = \frac{1}{2} g_q^\sharp(\alpha_q, \alpha_q) + V(q), \quad \alpha_q \in T_q^*Q,$$

where we are denoting by $g^\sharp$ the *co-metric* associated to the Riemannian metric g. Indeed, given a Riemannian metric g, there is an isomorphism of modules $\flat_g : \mathfrak{X}(Q) \to \Omega^1(Q)$ called the *flat isomorphism* given by

$$\langle \flat_g(X(q)), Y(q) \rangle = g_q(X(q), Y(q)), \quad X, Y \in \mathfrak{X}(Q).$$

Then the co-metric is the map $g^\sharp : \Omega^1(Q) \times \Omega^1(Q) \to C^\infty(Q)$ given by

$$g_q^\sharp(\flat_g(X(q)), \flat_g(Y(q))) = g_q(X(q), Y(q)), \quad X, Y \in \mathfrak{X}(Q).$$

It is also interesting to note that the Legendre transformation of the mechanical Lagrangian function $L_{(g,V)}$, denoted by $\mathbb{F}L_{(g,V)} : TQ \to T^*Q$, coincides with the flat isomorphism, i.e.,

$$\mathbb{F}L_{(g,V)} = \flat_g.$$

Moreover, we have that

$$(\mathbb{F}L_{(g,V)})^*\omega_Q = \omega_{L_{(g,V)}} \quad \text{and} \quad (\mathbb{F}L_{(g,V)})^* H_{(g,V)} = E_{L_{(g,V)}}, \tag{3.3.2}$$

where ω_Q is the canonical symplectic form on T^*Q. Therefore, we conclude that

$$(\flat_g)_* \Gamma_{L_{(g,V)}} = X_{H_{(g,V)}}.$$

3.3.2 Noether's theorem and symmetries

As we will see in this section, the symplectic formalism in which the equations of motion are written is specially useful to describe conserved quantities. We will see that when we have a symmetry then a conserved quantity arises. In the centre of the discussion will be a map called the momentum map introduced in its modern form by J. M. Souriau in 1970 (cf. [JM70]). The history of the discovery of this map is actually quite interesting and can be traced back until Sophus Lie's book *Theorie der Transformationsgruppen II*, who already wrote preliminary work on this concept and also coined the term.

As it will become clear in the examples, the terminology is not arbitrary and the momentum map can be seen as a generalization of the classical concepts of linear and angular momentum. For more details on this subject see the vast literature on mechanics: [AM78], [Blo15], [MR13].

Definition 3.3.7. Let (P, ω) be a symplectic manifold and $\Phi : G \times P \to P$ a symplectic action. The map $J : P \to \mathfrak{g}^*$ is called a *momentum map* if for every $\xi \in \mathfrak{g}$ the function $\hat{J}_\xi : P \to \mathbb{R}$ defined by $\hat{J}_\xi(p) = \langle J(p), \xi \rangle$, with $p \in P$, satisfies

$$i_{\xi_P}\omega = d\hat{J}_\xi,$$

where ξ_P is the infinitesimal generator vector field corresponding to ξ. In other words, ξ_P is the Hamiltonian vector field associated to the Hamiltonian function $\hat{J}_\xi$.

Now, we will see that if a Hamiltonian function H is invariant under a symplectic action, then the components of the momentum map are conserved quantities along integral curves of the Hamiltonian vector field X_H.

Theorem 3.3.8. *Let $\Phi : G \times P \to P$ be a symplectic action on the symplectic manifold (P, ω), with a momentum map $J : P \to \mathfrak{g}^*$. If the Hamiltonian* *function $H : P \to \mathbb{R}$ is invariant under the action of G, that is, $(H \circ \Phi_g)(p) = H(p)$, for any $p \in P$ and $g \in G$, then $\hat{J}_\xi$ is a conserved quantity along integral curves of the Hamiltonian vector field X_H.*

Proof. We will prove that $\{\hat{J}_\xi, H\} = 0$ which is enough to deduce that $\hat{J}_\xi$ is a conserved quantity, by Proposition 2.2.16. Indeed, we have that

$$\{\hat{J}_\xi, H\} = \omega(\xi_P, X_H) = -dH(\xi_P) = -\left.\frac{d}{dt}\right|_{t=0} (H \circ \Phi_{\exp(t\xi)}) = 0,$$

where we used the definitions of the canonical Poisson brackets, momentum map, Hamiltonian vector field and infinitesimal generator. The last step follows directly by invariance of the Hamiltonian function. □

Momentum maps not always exist. We will give two important examples of momentum maps.

Theorem 3.3.9. *Let $\Phi : G \times P \to P$ be a symplectic action on the symplectic manifold (P, ω). Suppose that $\omega = -d\theta$, for some $\theta \in \Omega^1(P)$ and that $\Phi_g^*\theta = \theta$, that is, Φ is an exact symplectomorphism. Then the map $J : P \to \mathfrak{g}^*$ defined by*

$$\langle J(p), \xi \rangle = \langle \theta(p), \xi_P(p) \rangle$$

is a momentum map.

Proof. Since the action is an exact symplectomorphism it follows that the Lie derivative of the one-form θ with respect to the infinitesimal generator ξ_P vanishes. Hence, using Cartan's magic formula and the fact that $\omega = -d\theta$ we deduce that

$$i_{\xi_P}\omega = d\left(i_{\xi_P}\theta\right).$$

Therefore the function defined by $\hat{J}_\xi = i_{\xi_P}\theta$ is the Hamiltonian function with Hamiltonian vector field ξ_P. Thus, by definition, J is a momentum map. □

Remark 3.3.10. In the special case where $P = T^*Q$ and θ_Q is the canonical one-form, any action of G on Q lifts to an action on T^*Q defined in (2.5.1), which preserves θ_Q and so, we can apply the previous theorem to the lifted action on T^*Q.

As a corollary of the previous theorem we obtain Noether's theorem, which summarizes the situation when $P = TQ$ and the action is Φ^T, that is, $\Phi_g^T = T\Phi_g$.

Corollary 3.3.11. *Let $\Phi : G \times Q \to Q$ be an action on Q and denote also by Φ the tangent lift of the action to TQ. Let $L : TQ \to \mathbb{R}$ be a regular Lagrangian function which is invariant under the tangent lifted action Φ^T. Then, Φ^T_g is an exact symplectomorphism for $\omega_L = -d\theta_L$ associated to the momentum map*

$$\hat{J}_\xi(v_q) = \langle \mathbb{F}L(v_q), \xi_Q(q)\rangle.$$

Moreover, the function $\hat{J}_\xi$ is a conserved quantity along integral curves of the Lagrangian vector field Γ_L.

Proof. The proof lies on three facts: the first is that G acts on TQ by exact symplectomorphisms (with respect to the Liouville one-form θ_L), secondly that the Hamiltonian function associated with the momentum map satisfies the equality

$$(i_{\xi_{TQ}}\theta_L)(v_q) = \langle \mathbb{F}L(v_q), \xi_Q(q)\rangle$$

and third, that the energy is invariant. Then, the combination of the previous two theorems leads to the conclusion.

Indeed, from the G-invariance of L, we may prove that

$$\Phi^{T^*}_g \circ \mathbb{F}L = \mathbb{F}L \circ \Phi^T_g.$$

Then using the previous equality together with the characterization of θ_L provided by Lemma 3.3.2 we deduce that

$$(\Phi^T_g)^*\theta_L = \theta_L.$$

As for the function $\hat{J}_\xi$, we know from the previous theorem that

$$\hat{J}_\xi(v_q) = \langle \theta_L(v_q), \xi_{TQ}(v_q)\rangle$$

and so using again Lemma 3.3.2 and the definition of canonical one-form we deduce that

$$\begin{aligned}\hat{J}_\xi(v_q) =& \langle \theta_Q(\mathbb{F}L(v_q)), T\mathbb{F}L(\xi_{TQ}(v_q))\rangle \\ =& \langle \mathbb{F}L(v_q), T\pi_Q \circ T\mathbb{F}L(\xi_{TQ}(v_q))\rangle\end{aligned}$$

while the fact that $\mathbb{F}L$ is fiber-preserving and ξ_{TQ} projects over ξ_Q implies that

$$\begin{aligned}\hat{J}_\xi(v_q) =& \langle \mathbb{F}L(v_q), T\tau_Q(\xi_{TQ}(v_q))\rangle \\ =& \langle \mathbb{F}L(v_q), \xi_Q(v_q)\rangle,\end{aligned}$$

as we have claimed. It remains to be shown that the energy is invariant with respect to the lifted action Φ_g^T, but as we have already said, the Legendre transform satisfies

$$\langle \mathbb{F}L(v), w\rangle = \langle \mathbb{F}L \circ \Phi_g^T(v), \Phi_g^T(w)\rangle, \quad \forall v, W \in T_qQ.$$

Hence, having into account that

$$E_L(v) = \Delta(L)(v) - L(v) = \langle \mathbb{F}L(v), v\rangle - L(v)$$

we conclude that E_L must be invariant.

$\square$

Example 3.3.12. Consider the manifold $Q = \mathbb{R}^n$ and the Lie group $G = \mathbb{R}^n$ acting by translations Q such that the lifted action on T^*Q is given by

$$g \cdot (q, p) = (q + g, p).$$

The infinitesimal generator of the lifted action is given by

$$\xi_{T^*Q}(q, p) = (\xi, 0).$$

It is not difficult to check by direct computation that the map $J : T^*Q \to \mathfrak{g}^*$ given by

$$J(q, p) = p$$

is a momentum map called *linear momentum*. Hence the infinitesimal generator is the Hamiltonian vector field associated to the Hamiltonian function

$$\hat{J}_\xi(q, p) = p \cdot \xi.$$

$\triangle$

Example 3.3.13. Consider now the manifold $Q = \mathbb{R}^3$ and the matrix Lie group $G = SO(3)$. If G acts on Q according with

$$\Phi_g(q) = g \cdot q,$$

where the dot denotes the action of the matrix g on the vector q. The lifted action to T^*Q is then given by

$$g \cdot (q, p) = (g \cdot q, g \cdot p).$$

The infinitesimal generator is given by

$$\xi_{T^*Q}(q,p) = (\xi \cdot q, \xi \cdot p) = (\hat{\xi} \times q, \hat{\xi} \times p),$$

where $\xi \in \mathfrak{so}(3)$ and the map $\hat{(\cdot)} : \mathfrak{so}(3) \to \mathbb{R}^3$ is the standard Lie algebra isomorphism. It is not hard to show that the map $J : T^*Q \to \mathfrak{so}(3)^*$ given by

$$J(q,p) = q \times p$$

is a momentum map called *angular momentum* with associated Hamiltonian function given by

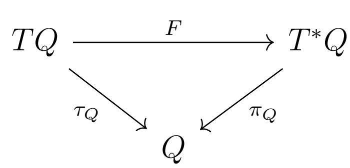
$$\hat{J}_\xi(q,p) = (q \times p) \cdot \hat{\xi}.$$

△

3.4 Mechanics under external forces

Lagrangian and Hamiltonian mechanics may be considered the building blocks of mechanics. Although they have intrinsic mathematical elegance and possess many nice geometric features they do not cover many relevant physical examples such as systems under the action of arbitrary external forces.

An *external force* can be interpreted as a fiber-preserving map denoted by $F : TQ \to T^*Q$ satisfying $\pi_Q \circ F = \tau_Q$. In canonical bundle coordinates (q^i, p_i) on T^*Q, the coordinate representation of the external force is of the form $F(q^i, \dot{q}^i) = (q^i, F_i(q^i, \dot{q}^i))$. So, F is fiber-preserving if and only if the following diagram is commutative:



It is well-know that to each such map we can associate a semibasic one-form on TQ defined by

$$\langle \mu_F(v_q), W \rangle = \langle F(v_q), T\tau_Q(W) \rangle, \quad v_q \in TQ \text{ and } W \in T_{v_q}TQ.$$

In coordinates $\mu_F = F_i(q^i, \dot{q}^i)\, dq^i$.

A *forced Lagrangian system* (L, F) is a mechanical system described by a Lagrangian function $L : TQ \to \mathbb{R}$ and subjected to an external force F. A trajectory of a forced mechanical system is a curve on Q satisfying the *Lagrange-d'Alembert principle.*

Definition 3.4.1 (Lagrange-d'Alembert principle)**.** The trajectory of the forced system given by the pair (L, F), where $L : TQ \to \mathbb{R}$ and $F : TQ \to T^*Q$ is a force map, between two fixed points $q_0, q_1 \in Q$ is a curve $q \in C^2(q_0, q_1)$ satisfying

$$\left.\frac{d}{ds}\right|_{s=0} \int_0^h L(q(t,s), \dot{q}(t,s))\, dt + \int_0^h \left\langle F(q(t), \dot{q}(t)), \frac{\partial q}{\partial s}(t,0) \right\rangle dt = 0, \quad (3.4.1)$$

for all smooth variations $q(s) \in C^2(q_0, q_1)$ of the curve q.

Using standard arguments from calculus of variations, it is not difficult to show that a curve $q : I \to Q$ satisfies Lagrange-d'Alembert principle if and only if in every coordinate neighbourhood it satisfies the *forced Euler-Lagrange* equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = F_i\, . \quad (3.4.2)$$

Example 3.4.2. If we are in the euclidean space $Q = \mathbb{R}^n$ and we have the Lagrangian function

$$L(q^i, \dot{q}^i) = \sum_{i=1}^{n} \frac{1}{2} m (\dot{q}^i)^2,$$

then the forced Euler-Lagrange equations give the second-order differential equations

$$m\ddot{q}^i = F_i,$$

which are Newton's second law in the presence of external forces. △

Example 3.4.3. A more interesting class of examples are the ones with *dissipative forces.* These forces are associated with friction or with motion in viscous media. In these cases, there is a force map $F : TQ \to T^*Q$ that is coordinate-wise proportional to the velocity of the trajectory, i.e.,

$$F_i = \lambda \dot{q}^i, \quad \text{for } 1 \leqslant i \leqslant n \text{ and } \lambda \in \mathbb{R}.$$

Moreover, if there is a function $R : TQ \to \mathbb{R}$ such that the force F is just the Legendre transform of $-R$, then R is called a *Rayleigh dissipation function*. In local coordinates, the forced Euler-Lagrange equations in this case give

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = -\frac{\partial R}{\partial \dot{q}^i} \,. \tag{3.4.3}$$

A simple example illustrating the situation is the damped pendulum. It is described by the Lagrangian function $L : T\mathbb{S}^1 \to \mathbb{R}$ given by

$$L(\theta, \dot{\theta}) = \frac{1}{2} m l^2 \dot{\theta}^2 + mgl \cos\theta,$$

where the first term corresponds to the kinetic energy, the second one to the potential energy, m is the mass of the pendulum, l is the length of the rod and g is the acceleration of gravity. Moreover, the function $R : T\mathbb{S}^1 \to \mathbb{R}$ given by

$$R(\theta, \dot{\theta}) = \frac{1}{2} \lambda l^2 \dot{\theta}^2, \quad \lambda \in \mathbb{R}$$

is a Rayleigh function. Then the forced Euler-Lagrange equations imply

$$\ddot{\theta} = -\frac{g}{l} \sin\theta - \frac{\lambda}{m} \dot{\theta},$$

which is a second-order differential equation known as the *damped harmonic oscillator*. △

As in the case of unconstrained systems, we may introduce a coordinate-free equation equivalent to forced Euler-Lagrange equations.

Proposition 3.4.4. *A curve $q(t)$ on Q satisfies the forced Euler-Lagrange equations* (3.4.2) *if and only if*

$$X^{\mathbf{c}}(L)(q, \dot{q}) - \frac{d}{dt}\left(X^{\mathbf{v}}(L)(q, \dot{q})\right) = \langle F(q, \dot{q}), X \circ q \rangle, \quad \forall\, X \in \mathfrak{X}(Q). \tag{3.4.4}$$

Proof. Let $X \in \mathfrak{X}(Q)$ be a vector field locally given by

$$X = X^i \frac{\partial}{\partial q^i}.$$

Then

$$X^{\mathbf{c}}(L) = X^i \frac{\partial L}{\partial q^i} + \dot{q}^j \frac{\partial X^i}{\partial q^j} \frac{\partial L}{\partial \dot{q}^i} \quad \text{and} \quad X^{\mathbf{v}}(L) = X^i \frac{\partial L}{\partial \dot{q}^i}.$$

Thus equations (3.4.4), along a curve $q(t)$ have the local form

$$X^i\left(\frac{\partial L}{\partial q^i}-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)\right)=F_iX^i.$$

Hence, we deduce that (3.4.4) is satisfied for all vector fields if and only if the curve $q(t)$ is a solution of forced Euler-Lagrange equations. □

Analogously to the case in which no forces appear, if L is regular, by non-degeneracy of the symplectic form ω_L, there is a unique vector field $\Gamma_{(L,F)}$ satisfying the geometric equation

$$i_{\Gamma_{(L,F)}}\omega_L = dE_L - \mu_F. \tag{3.4.5}$$

Proposition 3.4.5. *Let L be a regular Lagrangian function. Given a force map $F : TQ \to T^*Q$, there is a unique vector field $\Gamma_{(L,F)}$ satisfying equation* (3.4.5) *called the forced Lagrangian vector field. Also, $\Gamma_{(L,F)}$ is a SODE vector field on TQ and its integral curves satisfy the forced Euler-Lagrange equations* (3.4.2).

Proof. If L is regular, then the vector field $\Gamma_{(L,F)}$ in (3.4.5) exists and it is unique since ω_L is a symplectic form.

Moreover it is a SODE, since the one-form μ_F is semi-basic we have that $S^*(\mu_F) = 0$. Hence, we have that

$$i_\Delta\omega_L = i_{S\circ\Gamma_{(L,F)}}\omega_L,$$

using the same arguments as in the proof of Proposition 3.1.6.

Hence, the local expression of $\Gamma_{(L,F)}$ is of the form

$$\Gamma_{(L,F)}(q,\dot{q}) = \dot{q}^i\frac{\partial}{\partial q^i}+g^i(q,\dot{q})\frac{\partial}{\partial \dot{q}^i}.$$

If $q(t)$ is a trajectory of $\Gamma_{(L,F)}$, from (3.4.5) it satisfies the local equations

$$\frac{\partial^2 L}{\partial \dot{q}^i\partial q^j}\dot{q}^j+\frac{\partial^2 L}{\partial \dot{q}^i\partial \dot{q}^j}\ddot{q}^j-\frac{\partial L}{\partial q^i}=F_i,$$

which are equivalent to forced Euler-Lagrange equations (3.4.2). □

However, in the presence of forces neither the energy or the symplectic form are necessarily preserved by the flow of $\Gamma_{(L,F)}$. Indeed, the forced Lagrangian vector field may not be a Hamiltonian vector field.

Lemma 3.4.6. *We have that the energy along the flow of* $\Gamma_{(L,F)}$ *satisfies*

$$\langle dE_L(v_q), \Gamma_{(L,F)}(v_q)\rangle = \langle F(v_q), v_q\rangle, \quad \forall\ v_q \in T_qQ.$$

The symplectic form satisfies

$$\mathcal{L}_{\Gamma_{(L,F)}}\omega_L = -d\mu_F.$$

Proof. As for the equation concerning energy, note that by equation (3.4.5) and using the skew-symmetry of the symplectic form, we have that

$$\langle dE_L(v_q), \Gamma_{(L,F)}(v_q)\rangle = \langle \mu_F(v_q), \Gamma_{(L,F)}(v_q)\rangle.$$

Hence, by definition of the 1-form μ_F associated to the force map F and since $\Gamma_{(L,F)}$ is a SODE vector field we deduce that

$$\langle dE_L(v_q), \Gamma_{(L,F)}(v_q)\rangle = \langle F(v_q), v_q\rangle.$$

To compute the Lie derivative of the symplectic form, we use Cartan's identity

$$\mathcal{L}_{\Gamma_{(L,F)}}\omega_L = i_{\Gamma_{(L,F)}}(d\omega_L) + d(i_{\Gamma_{(L,F)}}\omega_L)$$

and then, noting that the first term on the right-hand side of the above equation vanishes since ω_L is exact, we deduce by (3.4.5) that

$$\mathcal{L}_{\Gamma_{(L,F)}}\omega_L = -d\mu_F.$$

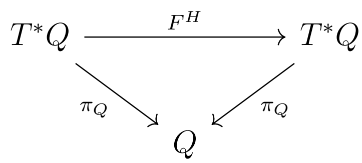
□

Remark 3.4.7. The last lemma implies that if $c : I \to Q$ is a solution of forced Euler-Lagrange equations, the energy changes along the integral curves according to the equation

$$\frac{d}{dt}E_L(c(t), \dot{c}(t)) = \langle F(c, \dot{c}), \dot{c}\rangle,$$

or, equivalently, the change in energy equals the *work done by the force*, that is,

$$E(c(h), \dot{c}(h)) - E(c(0), \dot{c}(0)) = \int_0^h \langle F(c, \dot{c}), \dot{c}\rangle\ dt.$$



Now, we move onto the Hamiltonian description of systems subjected to external forces. Given a Hamiltonian function $H : T^*Q \to \mathbb{R}$ we may define the *Hamiltonian external force* to be a fiber-preserving map $F^H : T^*Q \to T^*Q$, i.e., the following diagram is commutative:

We will say the the forced Hamiltonian system is determined by the pair (H, F^H). The Hamiltonian force map is associated to a semi-basic one-form β_{F^H} defined by

$$\langle \beta_{F^H}(\alpha_q), W \rangle = \langle F^H(\alpha_q), (T_{\alpha_q}\pi_Q)(W) \rangle, \quad \text{for } W \in T_{\alpha_q}(T^*Q).$$

It is possible to modify the Hamiltonian equations to obtain the forced Hamiltonian equations. The trajectories of forced Hamiltonian systems are the integral curves of the vector field $X_{(H,F^H)}$ determined by

$$i_{X_{(H,F^H)}}\omega_Q = dH - \beta_{F^H}. \tag{3.4.6}$$

Locally, we obtain forced Hamilton's equations as follows:

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}(q,p), \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}(q,p) + F_i^H(q,p). \tag{3.4.7}$$

In fact, the forced Hamiltonian vector field may be written as $X_H + Y_F^v$ where the vector field $Y_F^v \in \mathfrak{X}(T^*Q)$ is defined by

$$Y_F^v(\alpha_q) = \left.\frac{d}{dt}\right|_{t=0} (\alpha_q + tF^H(\alpha_q)) \, .$$

In coordinates,

$$Y_F^v = F_i^H\left(q^j, \frac{\partial H}{\partial p_j}(q,p)\right)\frac{\partial}{\partial p_i} = F_i^H(q,p)\frac{\partial}{\partial p_i}.$$

If L is a regular Lagrangian function and $F : TQ \to T^*Q$ is an external force map, it is possible to construct a Hamiltonian external force map $F^H : T^*Q \to T^*Q$ defined by $F^H = F \circ (\mathbb{F}L)^{-1}$. Then the forced Hamiltonian vector field $X_{(H,F^H)}$ associated to $H = E_L \circ (\mathbb{F}L)^{-1}$ is $\mathbb{F}L$-related to the forced Lagrangian vector field $\Gamma_{(L,F)}$.

3.5 A brief introduction to Contact mechanics

In this section, we will review the Lagrangian version of contact mechanics. This very recent formalism of mechanics (see [LLV19b] and the references therein), is especially useful to describe dissipative systems. As we will see, this formalism allows to incorporate some specific kinds of force maps in such a way that the dynamics will observe some geometric properties related with the contact structure.

Let Q be an n-dimensional *configuration manifold* and consider the *extended phase space* $TQ \times \mathbb{R}$ and a *contact Lagrangian function* $L : TQ \times \mathbb{R} \to \mathbb{R}$. In this section, we will assume that the Lagrangian is regular, that is, the Hessian matrix with respect to the velocities $Hess(L)$ is regular, where

$$Hess(L) = (W_{ij}) = \left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \right), \tag{3.5.1}$$

as before and $(q^i, \dot{q}^i, z)$ are natural bundle coordinates for $TQ \times \mathbb{R}$. Equivalently, L is regular if and only if the one-form

$$\eta_L = dz - \theta_L \tag{3.5.2}$$

is a contact form. Here,

$$\theta_L = S^*(dL) = \frac{\partial L}{\partial \dot{q}^i} dq^i, \tag{3.5.3}$$

where S is the extension of the canonical vertical endomorphism to $TQ \times \mathbb{R}$, that is, in local bundle coordinates on $TQ \times \mathbb{R}$

$$S = dq^i \otimes \frac{\partial}{\partial \dot{q}^i}. \tag{3.5.4}$$

The energy of the system is still defined by

$$E_L = \Delta(L) - L = \dot{q}^i \frac{\partial L}{\partial \dot{q}^i} - L, \tag{3.5.5}$$

where Δ is the extension of the Liouville vector field on TQ to $TQ \times \mathbb{R}$ in the natural way.

The Reeb vector field of η_L, which we will denote by $\mathcal{R}_L$ is given by

$$\mathcal{R}_L = \frac{\partial}{\partial z} - (W^{ij}) \frac{\partial^2 L}{\partial \dot{q}^i \partial z} \frac{\partial}{\partial \dot{q}^j}, \tag{3.5.6}$$

where (W^{ij}) is the inverse of the Hessian matrix of L with respect to the velocities (W_{ij}) (Equation (3.5.1)).

The Hamiltonian vector field of the energy E_L, denoted by $\xi_L = X_{E_L}$, hence

$$\flat_L(\xi_L) = dE_L - (\mathcal{R}_L(E_L) + E_L)\eta_L, \tag{3.5.7}$$

where $\flat_L(v) = i_v d\eta_L + \eta_L(v)\eta_L$ is the isomorphism defined in Equation (2.3.2) for this particular contact structure.

The vector field ξ_L will be called the *contact Lagrangian vector field*. It is a second order differential equation (SODE) and its trajectories are just the solutions of the Herglotz equations (also called generalized Euler-Lagrange equations) for L (see [LLV19b]):

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \frac{\partial L}{\partial \dot{q}^i}\frac{\partial L}{\partial z}. \tag{3.5.8}$$

There exists a *Legendre transformation* for contact Lagrangian systems. Given the vector bundle $TQ \times \mathbb{R} \to Q \times \mathbb{R}$, one can consider the fiber derivative $\mathbb{F}L$ of $L : TQ \times \mathbb{R} \to \mathbb{R}$, which has the following coordinate expression in natural coordinates:

$$\begin{aligned} \mathbb{F}L : TQ \times \mathbb{R} &\to T^*Q \times \mathbb{R} \\ (q^i, \dot{q}^i, z) &\mapsto (q^i, \frac{\partial L}{\partial \dot{q}^i}, z). \end{aligned} \tag{3.5.9}$$

If we consider the contact structure η_Q on $T^*Q \times \mathbb{R}$ given by (2.3.5) and the one-form η_L on $TQ \times \mathbb{R}$ then $\mathbb{F}L$ is a local contactomorphism.

In the case that $\mathbb{F}L$ is a global contactomorphism, then we say that L is *hyperregular*. In this situation, we can define a Hamiltonian $H : T^*Q \times \mathbb{R} \to \mathbb{R}$ such that $E_L = H \circ \mathbb{F}L$ and the Lagrangian and Hamiltonian dynamics are $\mathbb{F}L$-related, that is, $\mathbb{F}L_* \xi_L = X_H$.

3.5.1 Herglotz variational principle

Equations (3.5.8) can be derived from a modified variational principle [Her30]. In contrast to the symplectic case, the action is not a definite integral. The contact action is the endpoint value of a solution to a non-autonomous ODE (see [LLV19b] for more details).

Let Ω be the (infinite dimensional) manifold of twice-differentiable curves $c : [0,1] \to Q$ on Q. Given $q_0, q_1 \in Q$, we denote by $\Omega(q_0, q_1) \subseteq \Omega$ the submanifold whose elements are the smooth curves $c \in \Omega$ such that $c(0) = q_0$, $c(1) = q_1$. The tangent space of Ω at a curve c is given by vector fields over c, i.e.,

$$T_c\Omega = \{\delta c : [0,1] \to TQ \mid \tau_Q \circ \delta c = c\ \},$$

while the tangent vectors in $T_c\Omega(q_0, q_1)$ are the vector fields over c vanishing at the endpoints. Thus,

$$T_c\Omega(q_0, q_1) = \{\delta c \in T_c\Omega \mid \delta c(0) = 0,\ \delta c(1) = 0\ \}.$$

Now, we define the *contact action functional* as the map which assigns to each curve c and initial condition z_0, the integral of the Lagrangian over a curve $(c(t), \dot{c}(t), z(t))$ on $TQ \times \mathbb{R}$ where z is the solution of the following ODE:

$$\begin{cases} \frac{dz}{dt} = L(c, \dot{c}, z), \\ z(0) = z_0. \end{cases} \tag{3.5.10}$$

Thus we define the contact action functional as the map

$$\begin{aligned} &\mathcal{A} : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}, \\ &\mathcal{A}(c, z_0) = \int_0^1 L(c(t), \dot{c}(t), z(t))\ dt. \end{aligned} \tag{3.5.11}$$

When restricted to $\Omega(q_0, q_1) \times \{z_0\}$, the critical points of $\mathcal{A}_{z_0} = \mathcal{A}(\cdot, z_0)$ are the solutions to Herglotz equation. More precisely,

Theorem 3.5.1 (Herglotz variational principle)**.** *Let $L : TQ \times \mathbb{R} \to \mathbb{R}$ be a Lagrangian function and let $c \in \Omega(q_0, q_1)$ and $z_0 \in \mathbb{R}$. Then, $(c, \dot{c}, z)$ satisfies the Herglotz equations* (3.5.8)*, with $z(0) = z_0$, if and only if c is a critical point of $\mathcal{A}_{z_0}|_{\Omega(q_0,q_1)}$.*

We will prove the preceding theorem, since the result is not so well-known as the common Hamiltonian principle. Moreover, the careful inspection of this proof will help us later in Chapter 8 proving the discrete version of this principle.

Proof. Given $z_0 \in \mathbb{R}$, let $c \in \Omega(q_0, q_1)$ be a curve and consider some tangent vector $\delta c \in T_c\Omega(q_0, q_1)$.

Consider a variation $c_s \in \Omega$ of the curve c and let

$$\delta c = \left.\frac{dc_s}{ds}\right|_{s=0}, \quad c_0 = c.$$

Note that if $z(s,t)$ is the solution of the differential equation (3.5.10), associated to the curve c_s, then define

$$\delta z(t) = \left.\frac{dz}{ds}(s,t)\right|_{s=0}.$$

Note that $z(s,t)$ is a variation of $z(t)$, defined to be the solution of the differential equation (3.5.10), associated to the curve c and satisfies $z(0,0) = z_0$.

We compute the differential of the action to be

$$\langle d\mathcal{A}_{z_0}, \delta c\rangle = \int_0^1 \frac{\partial L}{\partial q^i}(\chi(t))\delta c^i(t) + \frac{\partial L}{\partial \dot{q}^i}(\chi(t))\delta \dot{c}^i(t) + \frac{\partial L}{\partial z}(\chi(t))\delta z(t)\ dt,$$

where $\chi(t)$ simply denotes the curve $(c(t), \dot{c}(t), z(t))$.

But, the function δz must satisfy the variational equation of (3.5.10), i.e.,

$$\frac{d}{dt}\delta z = \frac{\partial L}{\partial q^i}(\chi(t))\delta c^i(t) + \frac{\partial L}{\partial \dot{q}^i}(\chi(t))\delta \dot{c}^i(t) + \frac{\partial L}{\partial z}(\chi(t))\delta z(t).$$

Thus, we deduce that

$$\langle d\mathcal{A}_{z_0}, \delta c\rangle = \delta z(1) - \delta z(0).$$

So the problem reduces to solving the variational equation above. Using that $\delta z(0) = 0$, we can explicitly solve the Cauchy problem and obtain

$$\delta z(t) = \frac{1}{\sigma(t)} \int_0^t \sigma(\tau) \left(\frac{\partial L}{\partial q^i}(\chi(\tau))\delta c^i(\tau) + \frac{\partial L}{\partial \dot{q}^i}(\chi(\tau))\delta \dot{c}^i(\tau) \right) d\tau, \qquad (3.5.12)$$

where

$$\sigma(t) = \exp\left(- \int_0^t \frac{\partial L}{\partial z}(\chi(\tau))d\tau \right). \qquad (3.5.13)$$

Integrating by parts, we get the following expression

$$\delta z(1) = \frac{1}{\sigma(1)} \int_0^1 \delta c^i(t)\sigma(t) \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} + \frac{\partial L}{\partial \dot{q}^i}\frac{\partial L}{\partial z} \right) dt.$$

It is now clear that c is a critical value of $\mathcal{A}_{z_0}$ if and only if $c(t)$ and its associated curve $z(t)$ satisfy the Herglotz equations. □

3.5.2 Symmetries and dissipated quantities on contact Lagrangian systems

As explained in [Gas+20; LV20], given a symmetry on a contact system, one does not obtain a conserved quantity, but a quantity f that dissipates at the same rate as the Hamiltonian.

Given a contact Hamiltonian system (M, η, H), we say that a quantity $f : M \to \mathbb{R}$ is *dissipated* if

$$\mathcal{L}_{X_H} f = -\mathcal{R}(H) f, \tag{3.5.14}$$

or, equivalently,

$$\phi_t{}^*(f) = \sigma_t, \tag{3.5.15}$$

where ϕ is the flow of X_H and σ_t, its conformal factor.

Notice that the quotient of two dissipated quantities (if it is well defined) is a conserved quantity.

We end this section by stating a Noether's theorem in this setting, which provides a link between symmetries of the Lagrangian and conserved quantities.

Let $L : TQ \times \mathbb{R} \to \mathbb{R}$ be a regular Lagrangian. Let G be a Lie group acting on Q

$$\Phi : G \times Q \to Q. \tag{3.5.16}$$

We defined the lifted action as

$$\tilde{\Phi} : G \times TQ \times \mathbb{R} \to TQ \times \mathbb{R}, \tag{3.5.17}$$

given by $\tilde{\Phi}(g, v_q, z) = (T_q\Phi(v_q), z)$, where $v_q \in T_qQ$. We denote by $\xi_{TQ\times\mathbb{R}}$ to the vector field on $TQ \times \mathbb{R}$ which is the infinitesimal generator by the lifted action of an element ξ of the Lie algebra $\mathfrak{g}$ of G.

We define the momentum map J_L:

$$\begin{aligned} &J_L : TQ \times \mathbb{R} \to \mathfrak{g}^*, \\ &\langle J_L(v_q, z), \xi\rangle = -\left[\eta_L(\xi_{TQ\times\mathbb{R}})\right](v_q, z). \end{aligned} \tag{3.5.18}$$

and we define $\hat{J}(\xi) : TQ \times \mathbb{R} \to \mathbb{R}$ by $\hat{J}(\xi)(v_q, z) = \langle J_L(v_q, z), \xi\rangle$.

Then we have the following [LV20, Section 4.1]

Theorem 3.5.2. *If the lifted action $\tilde{\Phi}$ preserves the Lagrangian L, then $\tilde{\Phi}$ acts by contactomorphisms on $(TQ \times \mathbb{R}, \eta_L, E_L)$ and $\hat{J}(\xi)$ is a dissipated quantity for every $\xi \in \mathfrak{g}$.*

3.6 Nonholonomic mechanical systems

This section is a brief introduction to the subject of mechanics subjected to nonholonomic constraints. More generally, one finds many examples of systems which are constrained to move either in some submanifold of the ambient configuration manifold or which are constrained to move along a set of allowed directions, i.e., they move on a submanifold of the velocity space. The first kind of constraints are called *holonomic constraints* while the second kind is known as *nonholonomic constraints*. Some examples of holonomic constraints are a pendulum which while being inserted in a plane is constrained to move on a circle or the motion of a planet around a star which in principle lives in the euclidean three space but it is constrained to move on a plane.

The trajectories and properties of holonomically constrained systems have been well-understood for a long time. With nonholonomically constrained systems this was not the case. It was not until the end of the nineteenth century that the equations of motion for nonholonomic systems were established and in the last few decades mathematicians have been investigating further properties.

We will review the precise mathematical definition of nonholonomic constraint, state the equations of motion and give some examples (for further details see [Cor02; Blo15; LD96; CMR01; Cor+03; LM95; VF72; VM94; Koi92; BS93; SCS95; SSC96; Blo11; BMZ05; Blo+96b; FB08]).

3.6.1 Lagrange-d'Alembert principle

First we introduce the general notion of nonholonomic constraint. Then, along the remaining of the section we will treat the case of linear nonholonomic constraint.

Definition 3.6.1. A *nonholonomic constraint* on a system with configuration manifold Q is a submanifold M of the tangent space TQ, which is not the tangent space to any submanifold of Q.

We remark that this is the most general definition. If M is a distribution we say it is a *linear nonholonomic constraint*, which we will carefully examine in what follows, otherwise we are in the presence of a *nonlinear nonholonomic constraint*. Observe that if M was the tangent space of some submanifold N of Q, i.e., $M = TN$, the constraint would be implied by a constraint on the configuration manifold and we would be in the presence of a holonomic constraint. Also, if the constraint M is an integrable distribution, then by Fröbenius theorem the manifold Q is foliated by immersed submanifolds of Q whose tangent space at each point coincides with the subspace determined by the distribution at that point. Hence, each trajectory of a constrained system of this type is confined to evolve in a submanifold $N \subseteq Q$. This is why this constraint is called *semi-holonomic*.

A *(linear) nonholonomic mechanical system* is a pair $(L, \mathcal{D})$ where $\mathcal{D}$ is a distribution on Q and $L : TQ \to \mathbb{R}$ is a Lagrangian function. The trajectories of a nonholonomic system are subjected to the constraint in the sense that its velocity vectors belong to the distribution $\mathcal{D}$, i.e.,

$$\dot{q}(t) \in \mathcal{D}_{q(t)},$$

where $q : I \to Q$ is a curve.

Locally, linear nonholonomic constraints are given by a set of $n-k$ equations that are linear on the velocities

$$\mu_i^a(q)\dot{q}^i = 0,$$

where $1 \leqslant a \leqslant n-k$, k is the rank of the distribution $\mathcal{D}$ and n is the dimension od the manifold Q. Geometrically, these equations define the vector subbundle $\mathcal{D}^o \subseteq T^*Q$, called the *annihilator* of $\mathcal{D}$, spanned at each point by the one forms $\{\mu^a\}$ locally given by $\mu^a = \mu_i^a(q)dq^i$.

The trajectories of nonholonomic mechanical systems satisfy the following "pseudo-variational" principle (see [LM95; FB08] for a discussion about variational principles on systems subjected to nonholonomic constraints):

Definition 3.6.2. A curve $q : I \to Q$ with $q(0) = q_0$ and $q(h) = q_1$ satisfying $\dot{q} \in \mathcal{D}$ is a trajectory of the nonholonomic mechanical system $(L, \mathcal{D})$ if it is a critical point of the action integral $\mathcal{S} : C^2(q_0, q_1) \to \mathbb{R}$ given by

$$\mathcal{S}(q(\cdot)) = \int_0^h L(q(t), \dot{q}(t))\ dt,$$

among all variations satisfying $\delta q(0) = \delta q(h) = 0$ and $\delta q(t) \in \mathcal{D}_{q(t)}$.

Remark 3.6.3. Let us make a small parenthesis just to set the notation for infinitesimal variations. We will denote by $\delta q(t)$ the infinitesimal variation vector field along the curve q (recall the definition in (2.1.11)).

Analogously to the previous sections, we find a system of equations that give necessary and sufficient conditions for a curve to satisfy the Lagrange-d'Alembert principle.

Theorem 3.6.4. *A curve is a trajectory of the nonholonomic mechanical system $(L, \mathcal{D})$ if and only if it satisfies the Lagrange-d'Alembert equations, whose local expression is*

$$\begin{aligned} &\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \lambda_a \mu_i^a(q) \\ &\mu_i^a(q)\dot{q}^i = 0, \end{aligned} \tag{3.6.1}$$

for some Lagrange multipliers λ_a, which may be determined with the help of the constraint equations.

Proof. The proof is very similar to the standard proof of Euler-Lagrange equations that one finds in calculus of variations. Choosing a variation of the curve $q : I \to Q$ and denoting by $\delta q(t)$ the infinitesimal variation vector field, the differentiation of the action integral gives

$$d\mathcal{S}(q) = \int_0^h \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)\right) \delta q(t)\ dt,$$

where we apply differentiation by parts. Since $\delta q \in \mathcal{D}$, the integration vanishes for all such variations if the term between brackets is in the annihilator of $\mathcal{D}$, i.e., there exist functions λ_a called the Lagrange multipliers' such that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \lambda_a \mu_i^a(q). \tag{3.6.2}$$

The second equation in (3.6.1) comes from the fact that q satisfies the nonholonomic constraint. □

There is a great variety of examples of nonholonomic systems in the literature (see [Blo15; Cor02; Blo+96b; FGNM19; VV88; FJ04; FBZ14; FGN10; BM01; BM03]). We will introduce three of them where we will be able to compute the Lagrange multipliers' explicitly and then we will see that equations (3.6.1) belong to the family of *differential algebraic equations*, that is, they are formed by a system of differential equations and a coupled algebraic condition, the nonholonomic constraint.

Example 3.6.5. We will introduce here an example of a simple nonholonomic system to which we will get back all along the text: the *nonholonomic particle*. Consider a mechanical system in the configuration manifold $Q = \mathbb{R}^3$ defined by the Lagrangian

$$L(x, y, z, \dot{x}, \dot{y}, \dot{z}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

and subjected to the nonholonomic constraint $\dot{z} - y\dot{x} = 0$. The one-form $\mu = dz - y\, dx$ spans the vector subbundle $\mathcal{D}^o$, which is the annihilator of the distribution

$$\mathcal{D} = \operatorname{span}\left\{ \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}, \frac{\partial}{\partial y} \right\}.$$

Then the equations of motion of this system are given by Lagrange-d'Alembert equations (3.6.1), which in this case hold

$$\begin{cases} \ddot{x} = -\lambda y \\ \ddot{y} = 0 \\ \ddot{z} = \lambda \\ \dot{z} - y\dot{x} = 0 \end{cases} \qquad \Rightarrow \qquad \begin{cases} \ddot{x} = -y\frac{\dot{x}\dot{y}}{1+y^2} \\ \ddot{y} = 0 \\ \ddot{z} = \frac{\dot{x}\dot{y}}{1+y^2} \\ \dot{z} - y\dot{x} = 0, \end{cases} \tag{3.6.3}$$

where the value of λ is computed with the help of the constraints. △

Example 3.6.6. The *vertical rolling disk* is another classical example of a nonholonomic system for which we can easily find the explicit solution of the equations of motion.

The configuration space is $Q = \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$ and the Lagrangian function $L : TQ \to \mathbb{R}$ is given by

$$L(x, y, \varphi, \theta, \dot{x}, \dot{y}, \dot{\varphi}, \dot{\theta}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2,$$

where m is the mass of the disk and I, J are moments of inertia about an axis perpendicular to the plane of the disk and contained in the plane of the disk, respectively.

The disk is subjected to a constraint determined by the equations

$$\dot{x} = R\cos\varphi\ \dot{\theta}, \quad \dot{y} = R\sin\varphi\ \dot{\theta},$$

where R is the radius of the disk. The equations of motion are given by

$$\begin{cases} m\ddot{x} = \lambda_1 \\ m\ddot{y} = \lambda_2 \\ I\ddot{\theta} = -\lambda_1 R\cos\varphi - \lambda_2 R\sin\varphi \\ J\ddot{\varphi} = 0 \\ \dot{x} = R\cos\varphi\ \dot{\theta} \\ \dot{y} = R\sin\varphi\ \dot{\theta}, \end{cases} \tag{3.6.4}$$

where λ_1 and λ_2 are Lagrange multipliers' that can be determined with the help of the constraint equations. △

Example 3.6.7 (Chaplygin sleigh)**.** Consider the configuration manifold $Q = \mathbb{R}^2 \times \mathbb{S}^1$ where we define a Lagrangian function $L : TQ \to \mathbb{R}$ by

$$L(x, y, \theta, \dot{x}, \dot{y}, \dot{\theta}) = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 - 2a\dot{x}\dot{\theta}\sin\theta + 2a\dot{y}\dot{\theta}\cos\theta) + (I + ma^2)\frac{\dot{\theta}^2}{2},$$

where m is the mass of the sleigh, I is the moment of inertia about the centre of mass and a is the distance between the point (x, y) and the centre of mass (Check Figure 3.1). The limit case when $a = 0$, leads to the same kinetic energy one considers in the knife edge example which we will find later on.

The sleigh is constrained to move in the direction of its orientation at each point, i.e., it does not move sideways. Consequently, the constraint is given by the equation

$$\dot{y}\cos\theta - \dot{x}\sin\theta = 0.$$

Computing Lagrange-d'Alembert equations we obtain the equations of motion

$$\begin{cases} m\ddot{x} - am\cos\theta\dot{\theta}^2 - am\sin\theta\ddot{\theta} = -\lambda\sin\theta \\ m\ddot{y} - am\sin\theta\dot{\theta}^2 + am\cos\theta\ddot{\theta} = \lambda\cos\theta \\ (I + ma^2)\ddot{\theta} + ma\dot{\theta}(\dot{x}\cos\theta + \dot{y}\sin\theta) = 0 \\ \dot{y}\cos\theta - \dot{x}\sin\theta = 0. \end{cases} \tag{3.6.5}$$

Additionally, the Lagrange multiplier is completely determined by the equations and it may be computed to be

$$\lambda = \frac{\dot{\theta}(\dot{x}\cos\theta + \dot{y}\sin\theta)}{2}.$$

△

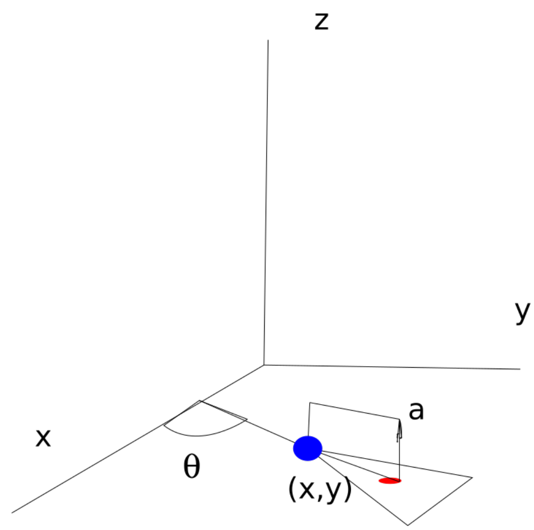


Figure 3.1: Representation of the Chaplygin sleigh.

3.6.2 The geometric formalism of nonholonomic systems

Now, we want to establish in which conditions are we able to determine the Lagrange multipliers'. In order to do so, we will reproduce a more general formalism, a geometric description of nonholonomic mechanics (see [LD96],[CMR01], [Cor+03], [LM95],[ZBM98] or [VF72] for a first geometrical approach to nonholonomic mechanics and also [SCS95; SSC96]). Consider the geometric equations

$$\begin{cases} \left(i_{\Gamma_{(L,\mathcal{D})}}\omega_L - dE_L\right)\Big|_{\mathcal{D}} \in (\mathcal{D}^o)^{\mathbf{v}} \\ \Gamma_{(L,\mathcal{D})} \in \mathfrak{X}(\mathcal{D}), \end{cases} \tag{3.6.6}$$

where $(\mathcal{D}^o)^{\mathbf{v}}$ is the vector subbundle of $T^*_{\mathcal{D}}(TQ) \to \mathcal{D}$ whose fiber at $v_q \in \mathcal{D}_q$ is spanned by

$$\{\alpha^{\mathbf{v}}(v_q) \mid \alpha \in \Gamma(\mathcal{D}^o)\}.$$

We will use the notation $F^o = (\mathcal{D}^o)^{\mathbf{v}}$.

Now, the nonholonomic system $(L, \mathcal{D})$ is called *regular* if L is a regular Lagrangian and the following condition is satisfied (again see [LD96]):

$$T_v\mathcal{D} \cap (\sharp_{\omega_L})_v(F^o_v) = \{0\} \text{ for all } v \in \mathcal{D} \text{ (\textit{compatibility condition})}.$$

Here, $\sharp_{\omega_L} : T^*(TQ) \to T(TQ)$ is the *sharp isomorphism* and it is the inverse map of the *flat isomorphism* defined by $\flat_{\omega_L}(X) = i_X \omega_L$.

For a regular nonholonomic system we have the following result regarding the geometric equations:

Theorem 3.6.8. *If the nonholonomic system* $(L, \mathcal{D})$ *is regular, then equations* (3.6.6) *have a SODE denoted by* $\Gamma_{(L,\mathcal{D})}$ *as a unique solution on* $\mathcal{D}$ *and its integral curves satisfy equations* (3.6.1). *In fact, for each* $v_q \in \mathcal{D}_q$*, there exists a unique* $F_{nh}(v_q) \in \mathcal{D}_q^o \subseteq T_q^*Q$ *such that*

$$i_{\Gamma_{(L,\mathcal{D})}} \omega_L(v_q) - dE_L(v_q) = F_{nh}^{\boldsymbol{v}}(v_q).$$

The following theorem is a useful sufficient condition to prove that the nonholonomic system is regular (see [LD96]).

Theorem 3.6.9. *If the Lagrangian L has either a positive definite or a negative definite Hessian matrix Hess*(L)*, then the nonholonomic system is also regular.*

Example 3.6.10. For mechanical Lagrangian systems of the form $L = K - V$ where the kinetic energy is associated with a Riemannian metric as in Examples 3.6.5, 3.6.6 and 3.6.7, there always is a unique well-defined nonholonomic Lagrangian vector field $\Gamma_{(L,\mathcal{D})}$ on $\mathcal{D}$. As it will become clear below, this is intimately related with the fact that when the nonholonomic system is regular, the Lagrange multipliers' are determined by the nonholonomic constraints. Indeed, if Hess(L) is positive (or negative) definite then the Lagrange multipliers' will automatically be determined. △

Before proving the preceding theorems we will introduce some useful notation. To each of the one-forms μ^a associate the fiberwise linear function $\widehat{\mu}^a : TQ \to \mathbb{R}$ defined by $\widehat{\mu}^a(v_q) = \langle \mu^a(q), v_q \rangle$, for $v_q \in T_qQ$. In local coordinates, equation (3.6.6) may be written like

$$i_{\Gamma_{(L,\mathcal{D})}} \omega_L - dE_L = \lambda_a S^*(d\widehat{\mu}^a) = \lambda_a \mu_i^a dq^i,$$

for some Lagrange multipliers λ_a. Indeed, if $\mathcal{D}$ is characterized as the zero set of $\widehat{\mu}^a$, then $(T\mathcal{D})^o$ is spanned by the one-forms $d\widehat{\mu}^a$.

If L is a regular Lagrangian, ω_L is symplectic and recalling the definition of the Lagrangian vector field Γ_L in (3.1.3), a solution $\Gamma_{(L,\mathcal{D})}$ of (3.6.6) must be of the form

$$\Gamma_{(L,\mathcal{D})} = \Gamma_L + \lambda_a Z^a,$$

where $Z^a = \sharp_{\omega_L}(\mu_i^a dq^i)$. The Lagrange multipliers may be computed by imposing the tangency condition in (3.6.6), which is equivalent to

$$0 = \Gamma_{(L,\mathcal{D})}(\widehat{\mu}^b) = \Gamma_L(\widehat{\mu}^b) + \lambda_a Z^a(\widehat{\mu}^b), \quad \text{for } b = 1, ..., n-k.$$

This equation has a unique solution for the Lagrange' multipliers if and only if the matrix $(C^{ab}) = (Z^a(\widehat{\mu}^b))$ is invertible at all points of $\mathcal{D}$. In local coordinates, one finds that

$$C^{ab} = \mu_i^a W^{ij} \mu_j^b, \tag{3.6.7}$$

where W^{ij} is the inverse matrix of $W_{ij} = \text{Hess}(L)$. In fact, in [LD96], the authors prove that the regularity of the matrix (C^{ab}) is equivalent to the compatibility condition.

Proposition 3.6.11. *The nonholonomic system $(L, \mathcal{D})$ is regular if and only if the matrix (C^{ab}) is non-singular on $\mathcal{D}$.*

Proof. Suppose first that the nonholonomic system $(L, \mathcal{D})$ is regular. Take a linear combination of any column of the matrix (C^{ab}), e.g.,

$$w^b = \lambda_a C^{ab} = \lambda_a Z^a(\widehat{\mu}^b).$$

If $w^b = 0$ and since $d\widehat{\mu}^a$ span $(T\mathcal{D})^o$, we must have that $\lambda_a Z^a(v) \in T_v\mathcal{D}$. But, by definition of the vector field Z^a, the vector $\lambda_a Z^a(v) \in (\sharp_{\omega_L})_v(F_v^o)$ and since $T_v\mathcal{D} \cap (\sharp_{\omega_L})_v(F_v^o) = \{0\}$ for all $v \in \mathcal{D}$, we conclude that

$$\lambda_a Z^a(v) = 0.$$

Since Z^a are linearly independent, we have that $\lambda_a = 0$, and (C^{ab}) is non-singular.

Conversely, if (C^{ab}) is non-singular, take a vector $X_v \in T_v\mathcal{D} \cap (\sharp_{\omega_L})_v(F_v^o)$. Thus, by definition we have that

$$X_v = \lambda_a Z^a(v) \quad \text{and} \quad X_v(\widehat{\mu}^a) = 0,$$

which implies that $\lambda_a Z^a(v)(\widehat{\mu}^b) = 0$. But since (C^{ab}) is non-singular, any linear combination of its columns vanishes if and only if $\lambda_a = 0$. Thus, $X_v = 0$. □

Proof of Theorem 3.6.8. Following the discussion above, if $(L, \mathcal{D})$ is a regular system then there is a unique solution of equations (3.6.6) of the form $\Gamma_{(L,\mathcal{D})} = \Gamma_L + \lambda_a Z^a$, where the Lagrange multipliers' λ_a are given by

$$\lambda_a = -\mathcal{C}_{ab}\Gamma_L(\widehat{\mu}^b),$$

where $(\mathcal{C}_{ab})$ is the inverse matrix of (C^{ab}).

Since Γ_L is a SODE and the vector fields Z^a are vertical, meaning that

$$S(Z^a) = 0,$$

the nonholonomic Lagrangian vector field $\Gamma_{(L,\mathcal{D})}$ is also a SODE.

Finally, computing the local expression of $\Gamma_{(L,\mathcal{D})}$ we deduce that if $q(t)$ is one of its trajectories, then it must satisfy the local equations

$$\begin{aligned}
&\frac{\partial^2 L}{\partial \dot{q}^i \partial q^j}\dot{q}^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\ddot{q}^j - \frac{\partial L}{\partial q^i} = \mathcal{C}_{ab}\Gamma_L(\widehat{\mu}^b)\mu_i^a \\
&\mu_i^a \dot{q}^i = 0.
\end{aligned}$$

These are exactly equations (3.6.1) provided we compute the corresponding Lagrange multipliers'. Indeed, from (3.6.1) we obtain that

$$\begin{aligned}
&\frac{\partial^2 L}{\partial \dot{q}^i \partial q^j}\dot{q}^j + W_{ij}\ddot{q}^j - \frac{\partial L}{\partial q^i} = \lambda_a \mu_i^a \\
&\mu_i^a \ddot{q}^i = -\frac{\partial \mu_i^a}{\partial q^j}\dot{q}^j \dot{q}^i.
\end{aligned}$$

where $(W_{ij}) = (\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j})$ is the Hessian matrix. Then, multiplying the first equation by the inverse matrix (W^{ij}) of (W_{ij}) and contracting it with μ_j^b we get

$$C^{ab}\lambda_a = \mu_j^b \ddot{q}^j - f^i(q, \dot{q})\mu_j^b,$$

where $f^i = W^{ij}(\frac{\partial L}{\partial q^i} - \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j}\dot{q}^j)$. We have chosen to denote this expression by f^i on purpose since the Lagrangian vector field Γ_L is precisely given by

$$\Gamma_L = \dot{q}^i \frac{\partial}{\partial q^i} + f^i \frac{\partial}{\partial \dot{q}^i}.$$

Then, using the fact that (C^{ab}) is invertible we conclude that

$$\begin{aligned}
\lambda_a &= C_{ab}\left(-\frac{\partial \mu_j^b}{\partial q^i}\dot{q}^i \dot{q}^j - f^i(q, \dot{q})\mu_j^b\right) \\
&= C_{ab}\Gamma_L(\widehat{\mu}^b),
\end{aligned}$$

which is just what we wanted to prove. □

Proof of Theorem 3.6.9. If the matrix (W^{ij}) in equation (3.6.7) is positive or negative definite, it generates for each fixed $v_q \in T_qQ$ an inner product with respect to which we may find an orthonormal basis of vectors in $\mathcal{D}$. In this way, the matrix (C^{ab}) is equivalent to the identity or minus the identity meaning that $(C^{ab}) = A^T(\pm I)A$, for some invertible matrix A. Thus (C^{ab}) is non-singular. □

Recall from symplectic geometry that $F^\perp = \sharp_{\omega_L}(F^o)$ for any distribution F, where $\perp$ denotes the symplectic orthogonal relative to ω_L. Hence, the compatibility condition also implies the Whitney sum decomposition

$$T(TQ)|_{\mathcal{D}} = T\mathcal{D} \oplus F^\perp,$$

to which we may associate two complementary projectors $P : T(TQ)|_{\mathcal{D}} \to T\mathcal{D}$ and $P' : T(TQ)|_{\mathcal{D}} \to F^\perp$ with coordinate expressions

$$P(X) = X - \mathcal{C}_{ab}X(\widehat{\mu}^b)Z^a, \quad P'(X) = \mathcal{C}_{ab}X(\widehat{\mu}^b)Z^a.$$

Proposition 3.6.12. *[LD96] The nonholonomic dynamics is given by*

$$\Gamma_{(L,\mathcal{D})} = P(\Gamma_L|_{\mathcal{D}}).$$

Proof. This is a consequence of the fact that

$$\Gamma_{(L,\mathcal{D})} = \Gamma_L - \mathcal{C}_{ab}\Gamma_L(\widehat{\mu}^b)Z^a.$$

□

We will finish the review of nonholonomic mechanics discussing some of its geometric properties. The first one is conservation of the energy. It is known that if the constraints are not linear on velocities, then the conservation of energy is not guaranteed. However, in this dissertation, we will only be concerned with linear constraints, determined by a distribution. Therefore, linear nonholonomic systems always verify the conservation of energy.

Proposition 3.6.13. *The Lagrangian energy E_L is conserved along the flow of the nonholonomic vector field $\Gamma_{(L,\mathcal{D})}$.*

Proof. Let us prove that the energy E_L is conserved along the flow of the nonholonomic vector field $\Gamma_{(L,\mathcal{D})}$ by computing the Lie derivative at points on $\mathcal{D}$. Noting that $\Gamma_{(L,\mathcal{D})} = \Gamma_L + \lambda_a Z^a$ and that by Theorem 3.1.8 the energy is conserved along the flow of Γ_L, we deduce that

$$\begin{aligned}\mathcal{L}_{\Gamma_{(L,\mathcal{D})}} E_L &= \lambda_a dE_L(Z^a) \\ &= \lambda_a dE_L(\sharp_{\omega_L}(S^* d\widehat{\mu}^a)),\end{aligned}$$

where in the last step we used the definition of the vector field Z^a. So by definition of the musical isomorphism $\sharp_L$ we have that

$$\begin{aligned}\mathcal{L}_{\Gamma_{(L,\mathcal{D})}} E_L &= \lambda_a S^* d\widehat{\mu}^a(\Gamma_L) \\ &= \lambda_a d\widehat{\mu}^a(\Delta) = \lambda_a \widehat{\mu}^a,\end{aligned}$$

where we have used the SODE property $S(\Gamma_L) = \Delta$. The fact that $d\widehat{\mu}^a(\Delta) = \widehat{\mu}^a$ may immediately be seen using natural bundle coordinates on TQ. Therefore, we have proved that

$$\mathcal{L}_{\Gamma_{(L,\mathcal{D})}} E_L(v_q) = \lambda_a(v_q)\widehat{\mu}^a(v_q), \quad v_q \in \mathcal{D}_q.$$

Since $\widehat{\mu}^a$ vanishes on $\mathcal{D}$, the result follows. □

Observe that, from the physical point of view, the conservation of the energy in nonholonomic systems is intimately related with the fact that the *nonholonomic force* annihilates the distribution and, hence, does no work along the flow. Indeed, we may define the nonholonomic force to be the map $F_{nh} : \mathcal{D} \to \mathcal{D}^o \subseteq T^*Q$ (see Theorem 3.6.8) determined in coordinates by the right hand-side of equations (3.6.1), i.e.,

$$F_{nh,i}(v_q) = \lambda_a(v_q)\mu_i^a(q), \quad v_q \in \mathcal{D}_q. \tag{3.6.8}$$

Hence, we have that

$$\dot{E}_L(q(t), \dot{q}(t)) = \langle F_{nh,i}(q(t), \dot{q}(t)), \dot{q}^i(t)\rangle$$

which must vanish along the trajectories of the nonholonomic vector field. Note that, if the nonholonomic force is extended to a map on the entire tangent bundle $F : TQ \to T^*Q$, then nonholonomic trajectories may be seen as trajectories of a forced Lagrangian system with initial conditions on $\mathcal{D}$.

The equations of motion can also be written on the cotangent bundle under the so-called nonholonomic Hamiltonian equations [VM94]. Note that, these will not be true Hamiltonian equations in the standard sense of symplectic geometry. In fact, note that the flow of $\Gamma_{(L,\mathcal{D})}$ is not symplectic. Indeed, computing the Lie derivative of ω_L in the direction of the nonholonomic vector field and using Cartan's magic formula, we deduce that

$$\mathcal{L}_{\Gamma_{(L,\mathcal{D})}}\omega_L = d(\lambda_a S^*(d\widehat{\mu}^a)),$$

where the right-hand side is in general, non-vanishing. So, the symplectic form is not conserved in general.

Nonetheless, one might transport Lagrange-d'Alembert equation to the cotangent setting using the Legendre transform $\mathbb{F}L$. Suppose that the nonholonomic system $(L, \mathcal{D})$ is regular and define, as usual, the Hamiltonian function $H : T^*Q \to \mathbb{R}$ given by

$$H(q,p) = E_L \circ (\mathbb{F}L)^{-1}(q,p)$$

and consider the constrained phase space

$$\mathcal{M} = \{(q,p) \in T^*Q \mid (\mathbb{F}H)(q,p) \in \mathcal{D}_q\} = \left\{(q,p) \in T^*Q \mid \mu_i^a(q)\frac{\partial H}{\partial p_i}(q,p) = 0\right\}.$$

Transporting the nonholonomic vector field $\Gamma_{(L,\mathcal{D})}$ to $\mathcal{M}$ using the Legendre transform we obtain the nonholonomic Hamiltonian vector field $X_{H,\mathcal{M}} = (\mathbb{F}L)_*(\Gamma_{(L,\mathcal{D})})$, whose integral curves satisfy the nonholonomic Hamiltonian equations

$$\begin{aligned}\dot{q} &= \frac{\partial H}{\partial p}\\ \dot{p} &= -\frac{\partial H}{\partial q} + \lambda_a \mu^a,\end{aligned}$$

where the Lagrange multipliers' may be computed using the equations defining the constraint phase space $\mathcal{M}$. It is also shown in [VM94] that the vector field $X_{H,\mathcal{M}}$ is the Hamiltonian vector field with respect to an almost Poisson structure on $\mathcal{M}$.

Finally, we will just mention that under some restrictive conditions we may state a nonholonomic version of Noether's Theorem relating invariant Lagrangians with respect to some lifted action by a Lie group, with the existence of conserved quantities. The first modern treatment of symmetry and reduction on nonholonomic systems dates back to the early 90's with the paper [Koi92]. For a modern treatment the literature is quite vast. See [FJ04; BGN12; BY20; Ehl+05; BFM09; GN10; BM07; VV88; Koz02; Jov10; BBM10; BBM15; GN19; BS15; Mes05; ML05; SCS99; FMB09; FBZ14; Ohs+11; BMZ09; Shi+17; SZB20; Can+98].

Note that in the unconstrained case, when the Lagrangian function is invariant with respect to the tangent lifted action $\Phi^T : G \times TQ \to TQ$ (as we have seen before, this implies that the energy E_L is also invariant), if $J : TQ \to \mathfrak{g}^*$ is the momentum map, then (see Corollary 3.3.11)

$$\Gamma_L(\hat{J}_\xi) = -dE_L(\xi_{TQ}) = 0,$$

where the last term vanishes due to the G-invariance of the energy. However, if we carried a similar computation in the nonholonomic case, replacing Γ_L by $\Gamma_{(L,\mathcal{D})}$, we would obtain extra terms which in general do not vanish. In fact, the associated Hamiltonian function $\hat{J}_\xi$ is conserved only if ξ_Q lies in the distribution $\mathcal{D}$. In this case, we say that $\xi \in \mathfrak{g}$ is a *horizontal* symmetry.

Summarizing the situation, we have that if $\xi \in \mathfrak{g}$ is a horizontal symmetry then the function

$$\hat{J}^{nh}_\xi(v_q) = \langle \mathbb{F}L(v_q), \xi_Q(q) \rangle, \quad v_q \in \mathcal{D}_q$$

is conserved along the nonholonomic flow.

3.7 Discrete-time Lagrangian mechanics

We will now describe a theory of discrete mechanics on the discretized velocity space $Q \times Q$. Discrete mechanics differs from continuous mechanics on the description of motion, which will no longer be a curve on the configuration manifold Q, it will be rather a sequence of points on Q. This is obviously connected with numerical calculation since any numerical method approximating the solution of a continuous mechanical system is first and foremost a sequence of points on its configuration manifold. Therefore when we compute a discrete trajectory we will also be obtaining a numerical method.

We describe a variational discrete theory based on a discretized Hamilton's principle. From here we see that much of the theory evolves in parallel with the continuous Lagrangian theory. See [MW01] for the main bibliographic account on the subject.

3.7.1 Discrete Euler-Lagrange equations

Given a configuration manifold Q, the starting point of the Lagrangian formalism is the choice of a Lagrangian function on the tangent space TQ. In order to develop a discrete Lagrangian formalism, the situation is similar except that we replace the tangent bundle TQ by the Cartesian product $Q \times Q$. Then we consider a function $L_d : Q \times Q \to \mathbb{R}$ and call it the *discrete Lagrangian function*.

For reasons we will understand later, the discrete Lagrangian function over a pair (q_0, q_1) is intended to approximate the action over the unique solution of mechanics connecting those points. Recall that the action is given by

$$\mathcal{S}[c] = \int_0^h L(c(t), \dot{c}(t))\ dt,$$

so

$$L_d(q_0, q_1) \approx \mathcal{S}[c_{01}],$$

where $c_{01} : I \to Q$ is the unique trajectory connecting q_0 and q_1. We remark that the existence of a unique trajectory is not a trivial question and we postpone its discussion until Chapter 4. For the time being, we will just admit that for two sufficiently close points and for a small enough positive number h this trajectory exists and it is unique.

We will introduce two operators we will use intensively. Given a function $F : Q \times Q \to \mathbb{R}$, we will denote by D_1F the differential of the function F relative to the first variable and by D_2F the differential of the function F relative to the second variable. So if we denote by $F_y : Q \to \mathbb{R}$ the function $F_y(x) = F(x, y)$ then

$$D_1F(x, y) = dF_y(x)$$

and analogously for the other variable.

Remark 3.7.1. An usual way to obtain a discrete Lagrangian function from a continuous one is by considering a map $R_h : \mathcal{U} \subseteq Q \times Q \to U \subseteq TQ$ (with $\mathcal{U}$ and U open subsets) that may depend on a real parameter $h > 0$. Such a map is called an *inverse retraction map*. Inverse retraction maps may be used to construct discrete Lagrangian functions from continuous-time ones.

Suppose we have two nearby points in Q, say q_0 and q_1, such that $(q_0, q_1) \in \mathcal{U}$. Also, consider the Lagrangian function $L : TQ \to \mathbb{R}$. Then we may define the discrete Lagrangian function $L_d^h : Q \times Q \to \mathbb{R}$ given by

$$L_d^h(q_0, q_1) = h \cdot (L \circ R_h)(q_0, q_1).$$

Example 3.7.2. Let $Q = \mathbb{R}^n$ and consider the map

$$\begin{aligned} R_h : Q \times Q &\to TQ \\ (q_0, q_1) &\mapsto \left(q_0, \frac{q_1 - q_0}{h} \right). \end{aligned}$$

Take a mechanical Lagrangian function $L : TQ \to \mathbb{R}$, i.e., a kinetic minus potential Lagrangian $L = K - V$, given by

$$L(q, \dot{q}) = \frac{m}{2}\dot{q}^2 - V(q)$$

and define $L_d^h : Q \times Q \to \mathbb{R}$ by $L_d^h(q_0, q_1) = h \cdot (L \circ R_h)(q_0, q_1)$, which gives

$$L_d^h(q_0^i, q_1^i) = \sum_{i=1}^{n} \frac{m(q_1^i - q_0^i)^2}{2h} - h \cdot V(q_0^i).$$

△

Now we will define a discrete analogue to Hamilton's principle, by making use of the calculus of variations on an appropriate discrete setting. The first step is to find a new functional that replaces the action and the infinite dimensional space where it is defined.

The *discrete path space* of length N is given by the space of sequences

$$C_d^N(Q) = \{\{q_k\}_{k=0}^N \, | q_k \in Q\}.$$

Furthermore, given $q_0, q_N \in Q$, we will work with the subset $C_d^N(q_0, q_N) \subseteq C_d^N(Q)$ formed by sequences with fixed end-points q_0 and q_N.

The *discrete action map* is defined to be the map $S_d : C_d^N(Q) \to \mathbb{R}$,

$$S_d(q_d) = \sum_{k=0}^{N-1} L_d^h(q_k, q_{k+1}). \tag{3.7.1}$$

Definition 3.7.3 (Discrete Hamilton's principle)**.** The discrete trajectory of the discrete Lagrangian system determined by the discrete Lagrangian function L_d^h is a critical value of the discrete action map (3.7.1) among all sequences of points with fixed end-points.

Proposition 3.7.4. *A sequence $\{q_k\}$ is a critical point of the functional S_d if and only if it is a solution of the discrete Euler-Lagrange equations*

$$D_2 L_d^h(q_{k-1}, q_k) + D_1 L_d^h(q_k, q_{k+1}) = 0, \quad \textit{for all } k = 1, ..., N-1. \tag{3.7.2}$$

Proof. Given q_0 and q_N in Q, take an arbitrary *variation of sequences* in $C_d^N(q_0, q_N)$, i.e., a sequence of curves of the form

$$q_d(s) = \{q_0(s), q_1(s), ..., q_N(s)\},$$

where $q_0(s) = q_0$ and $q_N(s) = q_N$ are fixed and let $q_i(0) = q_i \in Q$. The differential of the action is computed using an arbitrary variation in the following way:

$$\left.\frac{d}{ds}\right|_{s=0} S_d(q_d(s)) = \langle dS_d(q_d), \delta q_d \rangle,$$

where $\delta q_d = \left.\frac{d}{ds}\right|_{s=0} q_d(s)$. Then we immediately see that

$$\langle dS_d(q_d), \delta q_d \rangle = \sum_{k=0}^{N-1} \left[\langle D_1 L_d^h(q_k, q_{k+1}), \delta q_k \rangle + \langle D_2 L_d^h(q_k, q_{k+1}), \delta q_{k+1} \rangle \right].$$

Now using a discrete integration by parts (which is a just a rearrangement of the summation index), we deduce

$$\langle dS_d(q_d), \delta q_d \rangle = \sum_{k=1}^{N-1} \langle D_1 L_d^h(q_k, q_{k+1}) + D_2 L_d^h(q_{k-1}, q_k), \delta q_k \rangle.$$

The reader might have observed that we omitted two terms in the last line. We did so since they vanish when the differential acts on a variation with fixed end-points, namely

$$\langle D_1 L_d^h(q_0, q_1), \delta q_0 \rangle = \langle D_2 L_d^h(q_{N-1}, q_N), \delta q_N \rangle = 0.$$

Since the variation is arbitrary the proposition follows immediately. □

Example 3.7.5. Let $Q = \mathbb{R}^n$ and consider the discrete Lagrangian function introduced before, i.e.,

$$L_d^h(q_0^i, q_1^i) = \sum_{i=1}^{n} \frac{m(q_1^i - q_0^i)^2}{2h} - h \cdot V(q_0^i).$$

Then, discrete Euler-Lagrange equations imply that the sequence (q_0, q_1, q_2) must satisfy

$$q_2^i = 2q_1^i - q_0^i - \frac{h^2}{m}\frac{\partial V}{\partial q^i}(q_1).$$

which is a second-order discrete evolution map. △

Given a discrete Lagrangian $L_d^h : Q \times Q \to \mathbb{R}$ we can define two *discrete Legendre transformations* $\mathbb{F}^{\pm} L_d^h : Q \times Q \to T^*Q$ given by

$$\begin{aligned} \mathbb{F}^+ L_d^h(q_{k-1}, q_k) &= (q_k, D_2 L_d^h(q_{k-1}, q_k)) \, , \\ \mathbb{F}^- L_d^h(q_{k-1}, q_k) &= (q_{k-1}, -D_1 L_d^h(q_{k-1}, q_k)) \, . \end{aligned}$$

We say that L_d^h if *regular* if $\mathbb{F}^+ L_d^h$ (or, equivalently, $\mathbb{F}^- L_d^h$) is a local diffeomorphism. This is equivalent to the regularity of the matrix $D_{12} L_d^h$.

Moreover, if L_d^h is regular, since the discrete Euler-Lagrange equations (3.7.2) may be rewritten as

$$\mathbb{F}^+ L_d^h(q_{k-1}, q_k) = \mathbb{F}^- L_d^h(q_k, q_{k+1}),$$

then we can obtain a well-defined *discrete Lagrangian map*

$$\begin{array}{rccc} F_{L_d^h} : & Q \times Q & \longrightarrow & Q \times Q \\ & (q_{k-1}, q_k) & \longmapsto & (q_k, q_{k+1}(q_{k-1}, q_k)) \, , \end{array}$$

which is the discrete dynamical flow of our system. Here, $q_{k+1}(q_{k-1}, q_k)$ is the unique solution of the DEL equations (3.7.2) for the given pair (q_{k-1}, q_k).

3.7.2 Discrete symplectic structure

This method to construct integrators for Lagrangian systems enjoys plenty of nice geometric features such as a symplectic discrete flow and discrete momentum conservation.

Proposition 3.7.6. *If L_d^h is regular and ω_Q is the canonical symplectic form on T^*Q, then the 2-forms $(\mathbb{F}^{\pm} L_d^h)^* \omega_Q$ are equal and define a symplectic form on $Q \times Q$ denoted by $\Omega_{L_d^h}$.*

Proof. First note that there are two 1-forms (which are the discrete analogue of the *Poincaré-Cartan 1-form*) given by

$$\theta^{\pm}_{L_d^h} = (\mathbb{F}^{\pm} L_d^h)^* \theta_Q,$$

where θ_Q is the canonical 1-form on T^*Q. Locally, if (q_0^i, q_1^i) are coordinates on $Q \times Q$ they are represented by the expressions

$$\theta^+_{L_d^h} = \frac{\partial L_d^h}{\partial q_1^i} dq_1^i \quad \text{and} \quad \theta^-_{L_d^h} = -\frac{\partial L_d^h}{\partial q_0^i} dq_0^i.$$

Thus, we have that

$$dL_d^h = \theta^+_{L_d^h} - \theta^-_{L_d^h} \tag{3.7.3}$$

and so

$$-d\theta^+_{L_d^h} = -d\theta^-_{L_d^h}.$$

But since $\omega_Q = -d\theta_Q$, we conclude that

$$(\mathbb{F}^+L_d^h)^*\omega_Q = (\mathbb{F}^-L_d^h)^*\omega_Q =: \Omega_{L_d^h}.$$

Moreover, $\Omega_{L_d^h}$ is symplectic which is just a consequence of the fact that L_d^h is regular. □

We can check that the discrete flow preserves the symplectic form $\Omega_{L_d^h}$. Indeed, consider again the differential of the discrete action this time over a solution (q_0, q_1, q_2) of the discrete Euler-Lagrange equations and apply it to an arbitrary variation of the solution. In other words, we have that

$$(q_1, q_2) = F_{L_d^h}(q_0, q_1) \tag{3.7.4}$$

and so we deduce

$$\langle dS_d(q_d), \delta q_d\rangle = \langle D_1L_d^h(q_0, q_1), \delta q_0\rangle + \langle D_2L_d^h(q_1, q_2), \delta q_2\rangle.$$

Moreover note that

$$\langle -D_1L_d^h(q_0, q_1), \delta q_0\rangle = \langle \theta^-_{L_d^h}(q_0, q_1), (\delta q_0, \delta q_1)\rangle$$
$$\langle D_2L_d^h(q_0, q_1), \delta q_0\rangle = \langle \theta^+_{L_d^h}(q_0, q_1), (\delta q_0, \delta q_1)\rangle$$

and that from (3.7.4) we also have that

$$(\delta q_1, \delta q_2) = (F_{L_d^h})_*(\delta q_0, \delta q_1).$$

Thus, the differential of the discrete action may be rewritten as

$$\langle dS_d(q_d), \delta q_d\rangle = -\langle \theta^-_{L_d^h}(q_0, q_1), (\delta q_0, \delta q_1)\rangle + \langle (F^*_{L_d^h}\theta^+_{L_d^h})(q_0, q_1), (\delta q_0, \delta q_1)\rangle.$$

Finally, applying again the differential to the last expression, since $ddS_d = 0$ and

$$-d\theta^-_{L_d^h} = -d\theta^+_{L_d^h} = \Omega_{L_d^h}$$

we obtain

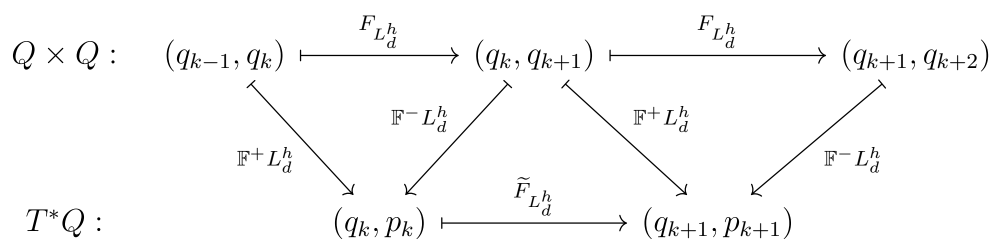
$$F^*_{L_d^h}\Omega_{L_d^h} = \Omega_{L_d^h}.$$

3.7.3 Hamiltonian viewpoint

Alternatively, using the discrete Legendre transformations, we can also define the evolution of the discrete system on the cotangent bundle by introducing the *discrete Hamiltonian map* $\widetilde{F}_{L_d^h} : T^*Q \longrightarrow T^*Q$ through any of the formulas

$$\widetilde{F}_{L_d^h} = \mathbb{F}^+ L_d^h \circ (\mathbb{F}^- L_d^h)^{-1} = \mathbb{F}^+ L_d^h \circ F_{L_d^h} \circ (\mathbb{F}^+ L_d^h)^{-1} = \mathbb{F}^- L_d^h \circ F_{L_d^h} \circ (\mathbb{F}^- L_d^h)^{-1} ,$$

which are equivalent due to the commutativity of the following diagram:



The discrete Hamiltonian map $\widetilde{F}_{L_d^h} : (T^*Q, \omega_Q) \longrightarrow (T^*Q, \omega_Q)$ is symplectic with respect to the canonical symplectic form ω_Q on the cotangent bundle.

3.7.4 Discrete Noether's theorem

Suppose that there is a symmetry of the discrete Lagrangian, i.e., there is a Lie group G acting on the manifold Q such that the discrete Lagrangian function is G-invariant with respect to the *diagonal action* on $Q \times Q$. If $\phi_g : Q \to Q$ is the Lie group action then the diagonal action is given by

$$\begin{aligned} \Phi_g : Q \times Q &\to Q \times Q \\ (q_0, q_1) &\mapsto (\phi_g(q_0), \phi_g(q_1)) \end{aligned}$$

The fact that L_d^h is G-invariant means that

$$L_d^h \circ \Phi_g = L_d^h .$$

Let $\mathfrak{g}$ be the Lie algebra of G. It is not difficult to show that if $\xi_Q \in \mathfrak{X}(Q)$ is the infinitesimal generator of the Lie group action ϕ associated to $\xi \in \mathfrak{g}$ then the corresponding infinitesimal generator of the diagonal action is characterized by

$$\xi_{Q\times Q}(q_0, q_1) = (\xi_Q(q_0), \xi_Q(q_1)).$$

As it is the case with the Poincaré-Cartan 1-forms, there are, in principle, two possible discrete analogues of the Lagrangian momentum map, which are given by the maps $J^{\pm}_{L^h_d} : Q \times Q \to \mathfrak{g}^*$ such that

$$\langle J^{\pm}_{L^h_d}(q_0, q_1), \xi \rangle = \langle \theta^{\pm}_{L^h_d}(q_0, q_1), \xi_{Q\times Q}(q_0, q_1) \rangle$$

However, the fact that L^h_d is G-invariant implies that

$$\langle dL_d(q_0, q_1), \xi_{Q\times Q}(q_0, q_1) \rangle = 0, \tag{3.7.5}$$

hence, using (3.7.3) we deduce that $J^+_{L^h_d}$ and $J^-_{L^h_d}$ must be equal. From now on, we will just use the notation $J_{L^h_d}$ to refer to any of the two maps and we will call it *discrete Lagrangian momentum map*.

Theorem 3.7.7 (Discrete Noether's theorem)**.** *Let G be a Lie group acting on Q and suppose that the discrete Lagrangian function L^h_d is G-invariant with respect to the diagonal action. Then the discrete momentum map $J_{L^h_d} : Q \times Q \to \mathfrak{g}^*$ is a conserved quantity of the discrete flow $F_{L^h_d}$, that is,*

$$J_{L^h_d} \circ F_{L^h_d} = J_{L^h_d}.$$

Proof. Let $\xi \in \mathfrak{g}$ be an arbitrary vector in the Lie algebra. Using one of the two definitions of $J_{L^h_d}$, we have that

$$\begin{aligned} \langle J_{L^h_d} \circ F_{L^h_d}(q_0, q_1), \xi \rangle &= \langle \theta^-_{L^h_d}(F_{L^h_d}(q_0, q_1)), \xi_{Q\times Q}(F_{L^h_d}(q_0, q_1)) \rangle \\ &= -\langle D_1 L^h_d(F_{L^h_d}(q_0, q_1)), \xi_Q(q_1) \rangle \\ &= \langle D_2 L^h_d(q_0, q_1), \xi_Q(q_1) \rangle \end{aligned}$$

where in the last step we used discrete Euler-Lagrange equations and in the second we have used the fact that $F_{L^h_d}(q_0, q_1) = (q_1, q_2)$. Now, note that by the infinitesimal invariance of L^h_d given by (3.7.5) we have that

$$\langle D_2 L^h_d(q_0, q_1), \xi_Q(q_1) \rangle = -\langle D_1 L^h_d(q_0, q_1), \xi_Q(q_0) \rangle$$

so that

$$\begin{aligned} \langle J_{L^h_d} \circ F_{L^h_d}(q_0, q_1), \xi \rangle &= \langle \theta^-_{L^h_d}(q_0, q_1), \xi_{Q\times Q}(q_0, q_1) \rangle \\ &= \langle J_{L^h_d}(q_0, q_1), \xi \rangle, \end{aligned}$$

which finishes the proof. $\square$

3.7.5 Implementation of variational numerical methods

Now, we will computationally observe the symplecticity of the discrete Lagrangian flow by plotting the phase space in an explicit example.

Example 3.7.8. Consider the simple pendulum in $Q = \mathbb{S}^1$. Using the same technique described in Remark 3.7.1, we define a discrete Lagrangian function $L_d^h : Q \times Q \to \mathbb{R}$ given by

$$L_d^h(\theta_0, \theta_1) = \frac{1}{2h} ml^2(\theta_1 - \theta_0)^2 + hmgl\cos\theta_0.$$

Applying the discrete Euler-Lagrange equations to this system we obtain the integrator

$$\theta_{k+2} = 2\theta_{k+1} - \theta_k - \frac{h^2 g}{l}\sin\theta_{k+1}, \quad \text{for any } k \geqslant 0$$

with discrete momenta given by

$$p_{k+1} = \frac{ml^2}{h}(\theta_{k+1} - \theta_k), \quad \text{for any } k \geqslant 0.$$

In Figure 3.2, we plotted the trajectory in the phase space T^*Q with respect to four different initial conditions. We see that the trajectories exhibit the same qualitative behaviour than the exact solution of the continuous-time equations we are numerically integrating.

△

Example 3.7.9. Another interesting example showing the potential of the discrete Lagrangian formalism is the study of a satellite motion around the Earth. The configuration manifold is $Q = \mathbb{R}^2 \setminus \{0\}$ and the Lagrangian function $L : TQ \to \mathbb{R}$ is given in polar coordinates by

$$L(r, \theta, \dot{r}, \dot{\theta}) = \frac{1}{2} m(\dot{r}^2 + r^2\dot{\theta}^2) + \gamma\frac{Mm}{r},$$

where m is the mass of the satellite, γ is Newton's gravitational constant and M is the Earth's mass. Using an inverse retraction map similar to the one used in the previous example, we construct the discrete Lagrangian function given by

$$L_d^h(r_0, \theta_0, r_1, \theta_1) = \frac{1}{2h} m\left[(r_1 - r_0)^2 + r_0^2(\theta_1 - \theta_0)^2\right] + h\gamma\frac{Mm}{r_0}$$

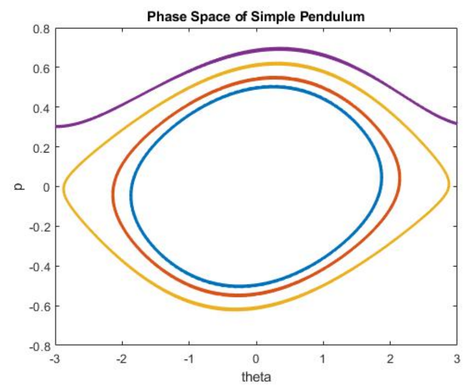


Figure 3.2: Plot of four different trajectories (θ_k, p_k) in the phase space T^*Q.

and by computing discrete Euler-Lagrange equations we obtain an integrator for the satellite motion. In Figure 3.3, we analyse the behaviour of the Hamiltonian function along the discrete trajectory (q_k, p_k) where $q_k = (r_k, \theta_k)$ and $p_k = ((p_r)_k, (p_\theta)_k)$, as well as the discrete analogue to the angular momentum which is $(p_\theta)_k$. We observe that both values remain approximately constant along the discrete flow generated by discrete Euler-Lagrange equations.

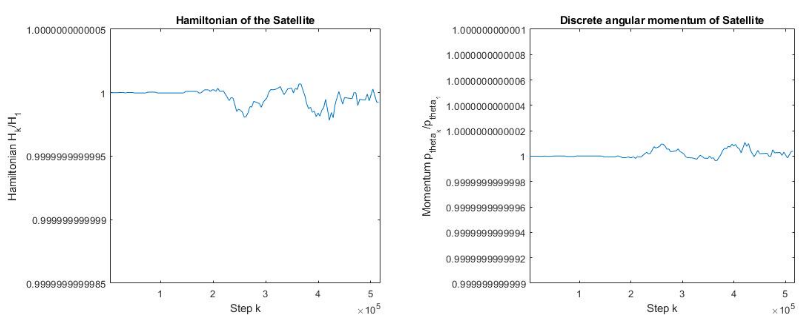


Figure 3.3: In the left-hand figure we plot the value of the Hamiltonian function along the trajectory (q_k, p_k) and divide it by its initial value $H(q_1, p_1)$. On the right we have the discrete analogue to the angular momentum $(p_\theta)_k$. We plot its value along the trajectory and divide it by its initial value.

$\triangle$

3.7.6 Exact discrete Lagrangian and discrete-continuous correspondence

When one wishes to construct a numerical method using discrete Lagrangian mechanics, one usually regards the value of the discrete Lagrangian on a point (q_0, q_1) as being an approximation of the (continuous) action, integrated over a solution connecting the two fixed points q_0, q_1 in a fixed time-step $h \in \mathbb{R}$, i.e.

$$L_d^h(q_0, q_1) \approx \int_0^h L(q_{0,1}(t), \dot{q}_{0,1}(t))\ dt,$$

where $q_{0,1}(t)$ is the unique solution of the Euler-Lagrange equations connecting q_0 and q_1. As we have mentioned at the beginning of this section, we will admit that at least for two sufficiently close points in Q and small enough $h > 0$ this trajectory exists.

We will describe the exact correspondence between continuous and discrete trajectories obtained by choosing the action integral as the discrete Lagrangian function.

Now suppose that L is a regular Lagrangian function $L : TQ \to \mathbb{R}$. We know from Section 3.1 that the dynamical vector field is a SODE Γ_L on TQ characterized by the geometric equation

$$i_{\Gamma_L}\omega_L = dE_L.$$

We will denote by

$$\exp_h^{\Gamma_L} : \mathcal{U}_h \subseteq TQ \to Q \times Q$$

the exponential map associated with Γ_L for a sufficiently small positive number h (see (2.4.22) for the definition). In Chapter 4, we will show that this map is a diffeomorphism when restricted to an appropriate neighbourhood and so we may consider the *exact inverse retraction* associated with Γ_L to be the inverse map R_h^{e-} of $\exp_h^{\Gamma_L}$.

Definition 3.7.10. The *exact discrete Lagrangian function* $L_d^{e,h} : Q \times Q \to \mathbb{R}$ is given by

$$L_d^{e,h}(q_0, q_1) = \int_0^h \left(L \circ \phi_t^{\Gamma_L} \circ R_h^{e-}\right)(q_0, q_1)\ dt,$$

where $\{\phi_h^{\Gamma_L}\}$ is the flow of Γ_L.

Now, denote by $q : Q \times Q \times [0, h] \to Q$ the function defined by

$$q(q_0, q_1, t) = q_{0,1}(t),$$

where $q_{0,1} : [0, h] \to Q$ is the solution of the Lagrangian system satisfying $q_{0,1}(0) = q_0$ and $q_{0,1}(h) = q_1$. Then it is clear that

$$q_{0,1}(t) = \left(\tau_Q \circ \phi_t^{\Gamma_L} \circ R_h^{e-}\right)(q_0, q_1).$$

So, with this notation, the map $L_d^{e,h}$ may be written as follows

$$L_d^{e,h}(q_0, q_1) = \int_0^h L(q_{0,1}(t), \dot{q}_{0,1}(t)) \, dt.$$

In [MW01], the authors prove the following theorem which gives us the correspondence between discrete and continuous Lagrangian mechanics:

Theorem 3.7.11. *Take a series of times $\{t_k = kh, k = 0, ..., N\}$ for a sufficiently small time-step $h \in \mathbb{R}$, a regular Lagrangian $L : TQ \to \mathbb{R}$ and its corresponding exact discrete Lagrangian function $L_d^{e,h} : Q \times Q \to \mathbb{R}$. Let $q(t)$ be a solution of Euler-Lagrange equations for L satisfying the boundary conditions $q(0) = q_0$ and $q(t_N) = q_N$. Define a sequence $\{q_k\}_{k=0}^N$ in Q by*

$$q_k = q(t_k), \quad \text{for } k = 0, ..., N.$$

Then $\{q_k\}_{k=0}^N$ is a solution of the discrete Euler-Lagrange equations for $L_d^{e,h}$.

Conversely, if we let $\{q_k\}_{k=0}^N$ be a solution of the discrete Euler-Lagrange equations for $L_d^{e,h}$, then the curve $q : [0, t_N] \to Q$ defined by

$$q(t) = q_{k,k+1}(t), \quad \text{for } t \in [t_k, t_{k+1}],$$

where $q_{k,k+1}(t)$ is the unique solution of the Euler-Lagrange equations connecting q_k and q_{k+1}, is a solution of Euler-Lagrange equations for L on the whole interval $[0, t_N]$.

In fact, the key ingredient to prove the previous theorem lies in a simple lemma also stated in [MW01]:

Lemma 3.7.12. *Given a regular Lagrangian function $L : TQ \to \mathbb{R}$, the discrete Legendre transformations of the exact discrete Lagrangian function satisfy*

$$\mathbb{F}^+ L_d^{e,h}(q_0, q_1) = \mathbb{F}L \circ R_h^{e+}(q_0, q_1), \quad \mathbb{F}^- L_d^{e,h}(q_0, q_1) = \mathbb{F}L \circ R_h^{e-}(q_0, q_1),$$

where R_h^{e-} is the exact inverse retraction associated with the Lagrangian vector field Γ_L and $R_h^{e+} = \phi_h^{\Gamma_L} \circ R_h^{e-}$.

Proof. Using the notation introduced previously, we will compute the discrete Legendre transformations of the exact discrete Lagrangian. For simplicity and in accordance with the line followed along these notes, we will use coordinate computations, though the proof can be done using exclusively intrinsic objects.

So, choosing coordinates (q_0^i, q_1^i) in $Q \times Q$ and since we have that

$$\mathbb{F}^- L_d^{e,h}(q_0, q_1) = -\frac{\partial L_d^{e,h}}{\partial q_0^i} dq_0^i,$$

we must compute the partial derivatives of $L_d^{e,h}$. This represents no trouble for us and we can proceed with differentiation under the integral sign, since we are supposing that all maps are smooth. Thus,

$$\frac{\partial L_d^{e,h}}{\partial q_0^i} = \int_0^h \left[\frac{\partial L}{\partial q^j} \frac{\partial q_{0,1}^j}{\partial q_0^i}(t) + \frac{\partial L}{\partial \dot{q}^j} \frac{\partial \dot{q}_{0,1}^j}{\partial q_0^i}(t) \right] dt.$$

Using integration by parts we get

$$\frac{\partial L_d^{e,h}}{\partial q_0^i} = \int_0^h \left[\frac{\partial L}{\partial q^j} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^j} \right] \frac{\partial q_{0,1}^j}{\partial q_0^i}(t)\ dt + \left[\frac{\partial L}{\partial \dot{q}^j} \frac{\partial q_{0,1}^j}{\partial q_0^i}(t) \right]_0^h.$$

The term under the integral sign vanishes, since it is being evaluated over the solution $q_{0,1}$ of Euler-Lagrange equations. Also, by definition of the function $q_{0,1}$ given previously, we have that

$$\frac{\partial q_{0,1}^j}{\partial q_0^i}(0) = \delta_i^j, \quad \frac{\partial q_{0,1}^j}{\partial q_0^i}(h) = 0.$$

Hence, we have that

$$\frac{\partial L_d^{e,h}}{\partial q_0^i} = -\frac{\partial L}{\partial \dot{q}^i}(q_{0,1}(0), \dot{q}_{0,1}(0)),$$

and thus

$$\begin{aligned} \mathbb{F}^- L_d^{e,h}(q_0, q_1) &= \frac{\partial L}{\partial \dot{q}^i}(q_{0,1}(0), \dot{q}_{0,1}(0)) dq_0^i \\ &= \mathbb{F}L \circ R_h^{e-}(q_0, q_1). \end{aligned}$$

The equality for the other discrete Legendre transformation is proved using the same argument. □

Next, let us see a simple example illustrating the statement of Theorem 3.7.11.

Example 3.7.13. Consider the harmonic oscillator with configuration manifold $Q = \mathbb{S}^1$ and Lagrangian function given by

$$L(\theta, \dot{\theta}) = \frac{\dot{\theta}^2}{2} - \frac{\theta^2}{2},$$

where we have set all physical constants equal to one for simplicity. Then Euler-Lagrange equations reduce to the second-order differential equation

$$\ddot{\theta} = -\theta,$$

which has as a general solution the curve $\theta : \mathbb{R} \to Q$

$$\theta(t) = \theta_0 \cos(t) + \dot{\theta}_0 \sin(t),$$

satisfying the initial conditions

$$\theta(0) = \theta_0, \quad \dot{\theta}(0) = \dot{\theta}_0,$$

for some $(\theta_0, \dot{\theta}_0) \in TQ$. Fixing an adequate positive number h, the exponential map $\exp_h^L : TQ \to Q \times Q$ is given by

$$\exp_h^L(\theta_0, \dot{\theta}_0) = \left(\theta_0, \theta_0 \cos(h) + \dot{\theta}_0 \sin(h)\right),$$

which is invertible provided $h \neq k\pi$ for any $k \in \mathbb{N}$ and its inverse map $R_h^{e-} : Q \times Q \to TQ$ is given by

$$R_h^{e-}(\theta_0, \theta_1) = \left(\theta_0, \frac{\theta_1 - \theta_0 \cos(h)}{\sin(h)}\right).$$

Notice that the curve $\theta_{0,1} : \mathbb{R} \to Q$ given by

$$\theta_{0,1}(t) = \theta_0 \cos(t) + \frac{\theta_1 - \theta_0 \cos(h)}{\sin(h)} \sin(t),$$

is the unique solution of Euler-Lagrange equations satisfying the boundary conditions

$$\theta_{0,1}(0) = \theta_0, \quad \theta_{0,1}(h) = \theta_1,$$

with $(\theta_0, \theta_1) \in Q \times Q$. Now, the exact discrete Lagrangian function $L_d^{e,h}$: $Q \times Q \to \mathbb{R}$ can be computed and it is given by

$$L_d^{e,h}(\theta_0, \theta_1) = \frac{\theta_0^2 \cos(h) + \theta_1^2 \cos(h) - 2\theta_0\theta_1}{2\sin(h)}.$$

After applying discrete Euler-Lagrange equations, we deduce that

$$\theta_2 = 2\cos(h)\theta_1 - \theta_0,$$

but this is exactly the expression of $\theta_{0,1}(2h)$ after trigonometric simplifications. Hence, we have experimentally obtained the result in Theorem 3.7.11. $\triangle$

Finally, we will just mention another reason why the methods arising from discrete Lagrangian mechanics are so useful. This last feature is summarized by the *variational error order theorem*, which states that the error of the algorithm may be estimated from the error committed when approximating the exact discrete Lagrangian function with a discrete Lagrangian L_d^h.

This result is originally contained in [MW01] and it was correctly proven later in [PC09].

Theorem 3.7.14. *If $\widetilde{F}_{L_d^h} : T^*Q \to T^*Q$ is the discrete Hamiltonian map of an order r discretization $L_d^h \colon Q \times Q \to \mathbb{R}$ of the exact discrete Lagrangian $L_d^{e,h} \colon Q \times Q \to \mathbb{R}$, then*

$$\widetilde{F}_{L_d^h} = \widetilde{F}_{L_d^{e,h}} + \mathcal{O}(h^{r+1}).$$

In other words, $\widetilde{F}_{L_d^h}$ gives an integrator of order r for the exact discrete Hamiltonian map $\widetilde{F}_{L_d^{e,h}} : T^*Q \to T^*Q$.

3.7.7 Forced discrete mechanics

When an external force is considered, we may formulate a forced Lagrangian discrete formalism by incorporating the necessary changes to the discrete theory we have just presented. In fact, one of the most important features of variational integrators is the capability to adapt to more complex situations, for instance, systems involving forces or constraints (see [MW01]).

Given an external force map $F : TQ \to T^*Q$, we introduce its discrete counterpart as two maps $F_d^+ : Q \times Q \longrightarrow T^*Q$ and $F_d^- : Q \times Q \longrightarrow T^*Q$

called the *discrete force maps*. These discrete forces satisfy $\pi_Q \circ F_d^+ = \mathrm{pr}_2$ and $\pi_Q \circ F_d^- = \mathrm{pr}_1$, where π_Q is the canonical projection of the cotangent bundle, and $\mathrm{pr}_{1,2} : Q \times Q \longrightarrow Q$ are the canonical projections onto the first and second factors, respectively.

Now, the discrete equations of motion are derived from the *discrete Lagrange-d'Alembert principle*:

$$\delta S_d(q_d) \cdot \delta q_d + \sum_{k=1}^{N-1} \left[F_d^+(q_{k-1}, q_k) + F_d^-(q_k, q_{k+1})\right] \cdot \delta q_k = 0 \tag{3.7.6}$$

for all variations δq_k, with $\delta q_0 = \delta q_N = 0$.

The *forced Euler-Lagrange equations* are given by

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) + F_d^+(q_{k-1}, q_k) + F_d^-(q_k, q_{k+1}) = 0 \ , \tag{3.7.7}$$

which implicitly define a discrete forced Lagrangian map $F_{L_d^f} : Q \times Q \to Q \times Q$.

As in the unforced case, we can define the corresponding discrete Legendre transformations $\mathbb{F}^{f\pm} L_d : Q \times Q \to T^*Q$ given by

$$\begin{aligned} \mathbb{F}^{f+} L_d(q_{k-1}, q_k) &= (q_k, D_2 L_d(q_{k-1}, q_k) + F_d^+(q_{k-1}, q_k)) \, , \\ \mathbb{F}^{f-} L_d(q_{k-1}, q_k) &= (q_{k-1}, -D_1 L_d(q_{k-1}, q_k) - F_d^-(q_{k-1}, q_k)) \, . \end{aligned}$$

If the discrete forced system is regular, that is, the discrete Legendre transformations $\mathbb{F}^{f\pm} L_d$ are local diffeomorphisms then we have an explicit discrete forced Lagrangian map $F_{L_d^f}$ which is a local diffeomorphism. In addition, we may consider the discrete forced Hamiltonian map $\widetilde{F}_{L_d^f} : T^*Q \to T^*Q$

$$\widetilde{F}_{L_d^f} = \mathbb{F}^{f\pm} L_d \circ F_{L_d^f} \circ \left(\mathbb{F}^{f\pm} L_d\right)^{-1} \ .$$

Now suppose that (L, F) is a forced continuous Lagrangian system with regular Lagrangian function $L : TQ \to \mathbb{R}$ and an external force $F : TQ \to T^*Q$. Then, as we know (see Section 3.4), the dynamical vector field is a SODE $\Gamma_{(L,F)}$ on TQ which is characterized by condition (3.4.5).

We will denote by

$$\exp_h^{\Gamma_{(L,F)}} : \mathcal{U}_h \subseteq TQ \to Q \times Q$$

the exponential map associated with $\Gamma_{(L,F)}$ for a sufficiently small positive number h. We will again assume that this map is a diffeomorphism when restricted to a sufficiently small open subset $\mathcal{U}_h$ and so we may consider the corresponding exact inverse retraction, which is its inverse map $R^{e-}_{h,F}$. On Chapter 4 we will prove that our assumption is correct and we can always consider an open set where the exponential map of any SODE is a diffeomorphism.

Using the flow $\phi_h^{\Gamma_{(L,F)}}$ of $\Gamma_{(L,F)}$ and the associated exact retraction we may introduce the forced exact discrete Lagrangian function $L^{e,h}_{d,F} : Q \times Q \to \mathbb{R}$ given by

$$L^{e,h}_{d,F}(q_0, q_1) = \int_0^h \left(L \circ \phi_t^{\Gamma_{(L,F)}} \circ R^{e-}_{h,F} \right) (q_0, q_1) \; dt,$$

and the double exact discrete force $F^{e,h}_d : Q \times Q \to T^*(Q \times Q)$ defined by

$$\langle F^{e,h}_d(q_0, q_1, h), (X_{q_0}, X_{q_1}) \rangle = \int_0^h \langle \left(F \circ \phi_t^{\Gamma_{(L,F)}} \circ R^{e-}_{h,F} \right) (q_0, q_1), X_{0,1}(t) \rangle \; dt$$

where $X_{0,1}(t) = T_{(q_0,q_1)}(\tau_Q \circ \phi_t^{\Gamma_{(L,F)}} \circ R^{e-}_{h,F})(X_{q_0}, X_{q_1})$, for $(X_{q_0}, X_{q_1}) \in T_{q_0}Q \times T_{q_1}Q$.

Then, the exact discrete force maps are just $F^{e,+}_d : Q \times Q \to T^*Q$ and $F^{e,-}_d : Q \times Q \to T^*Q$ given by

$$\begin{aligned} \langle F^{e,+}_d(q_0, q_1), X_{q_1} \rangle &= \langle F^{e,h}_d(q_0, q_1), (0_{q_0}, X_{q_1}) \rangle \\ \langle F^{e,-}_d(q_0, q_1), X_{q_0} \rangle &= \langle F^{e,h}_d(q_0, q_1), (X_{q_0}, 0_{q_1}, \rangle. \end{aligned}$$

Now, denote by $q : Q \times Q \times [0, h] \to Q$ the function defined by

$$q(q_0, q_1, t) = q_{0,1}(t),$$

where $q_{0,1} : [0, h] \to Q$ is the solution of the forced Lagrangian system satisfying $q_{0,1}(0) = q_0$ and $q_{0,1}(h) = q_1$. Then it is clear that

$$q_{0,1}(t) = \left(\tau_Q \circ \phi_t^{\Gamma_{(L,F)}} \circ R^{e-}_{h,F} \right) (q_0, q_1).$$

So, with this notation, the maps $L^{e,h}_{d,F}$, $F^{e,+}_d$ and $F^{e,-}_d$ may be written as follows

$$L^{e,h}_{d,F}(q_0, q_1) = \int_0^h L(q_{0,1}(t), \dot{q}_{0,1}(t)) \; dt,$$

$$F_d^{e,+}(q_0, q_1) = \int_0^h \langle F(q_{0,1}(t), \dot{q}_{0,1}(t)), \frac{\partial q_{0,1}}{\partial q_1}\rangle \, dt$$

and

$$F_d^{e,-}(q_0, q_1) = \int_0^h \langle F(q_{0,1}(t), \dot{q}_{0,1}(t)), \frac{\partial q_{0,1}}{\partial q_0}\rangle \, dt,$$

where

$$\frac{\partial q_{0,1}}{\partial q_1} : T_{q_1}Q \to T_{q_{0,1}(t)}Q, \quad \text{and} \quad \frac{\partial q_{0,1}}{\partial q_0} : T_{q_0}Q \to T_{q_{0,1}(t)}Q$$

are given by

$$\langle \frac{\partial q_{0,1}}{\partial q_1}, X_{q_1}\rangle = T_{(q_0,q_1,t)}q(0_{q_0}, X_{q_1}, 0_t), \quad \langle \frac{\partial q_{0,1}}{\partial q_0}, X_{q_0}\rangle = T_{(q_0,q_1,t)}q(X_{q_0}, 0_{q_1}, 0_t),$$

for $X_{q_0} \in T_{q_0}Q$ and $X_{q_1} \in T_{q_1}Q$.

Using the previous definitions, one may prove a forced version of Theorem 3.7.11 (cf. [MW01]). Moreover, in [DA18], the authors give a forced version of the standard variational error theorem 3.7.14 using the variational error order of an associated unforced system with double dimension.

Finally, we will state a useful Lemma from [MW01], which is again a modified version of an existing theorem in unforced discrete mechanics.

Lemma 3.7.15. *Let (Q, L, F) be a forced Lagrangian problem with regular Lagrangian function L. The corresponding exact discrete Legendre transformations satisfy*

1. $\mathbb{F}^{f+}L_{d,F}^{e,h}(q_0, q_1) = \mathbb{F}L(q_{0,1}(h), \dot{q}_{0,1}(h))$;

2. $\mathbb{F}^{f-}L_{d,F}^{e,h}(q_0, q_1) = \mathbb{F}L(q_{0,1}(0), \dot{q}_{0,1}(0))$;

where $q_{0,1}(t)$ is the unique solution of the forced Euler-Lagrange equations verifying $q_{0,1}(0) = q_0$ and $q_{0,1}(h) = q_1$.

3.7.8 A prior version of nonholonomic discrete mechanics

There have been several attempts to capture the nature of nonholonomic mechanics in the discrete setting (cf. [CM01; MP06; FID08; BZ15; FBO12;

Cel+19; Igl+08; GNV21; GNJ17; FZ05]). We will not detain ourselves describing all the methods proposed so far because the literature is far too vast on this topic and our purpose is not to make a comparison between the different methods (see [MV19] for an excellent comparison) but rather find a method that agrees with the continuous dynamics, in a sense that we will specify later.

For the sake of completeness, we will describe here the formalism proposed in [CM01] because it is the numerical method that most resembles the steps followed in the unconstrained discrete Lagrangian theory. Moreover, it also has the advantage that its construction resembles the numerical method we will propose in Chapter 7. So, it serves our purposes for introducing our own advances in the subsequent chapters.

Indeed, the starting point of the formalism proposed in [CM01] (see also [Cor02]) is the discrete Lagrangian function $L_d^h : Q \times Q \to \mathbb{R}$ on the discrete velocity space $Q \times Q$. Let $\mathcal{D}$ be a distribution on Q and consider a discrete constraint space $\mathcal{M}_d \subseteq Q \times Q$ satisfying two requirements:

1. its dimension agrees with that of the distribution $\mathcal{D}$ as a submanifold of TQ, $\dim\mathcal{M}_d = \dim\mathcal{D}$;

2. The diagonal set of $Q \times Q$ is contained in the discrete constraint space, $(q, q) \in \mathcal{M}_d$ for all $q \in Q$.

The discrete constraint space will thereby impose a restriction of the admissible pairs of points, namely $(q_k, q_{k+1}) \in \mathcal{M}_d$. In summary, within this discrete formalism for nonholonomic mechanics, we need three ingredients: a discrete Lagrangian function, a distribution and a discrete constraint space.

Then, the discrete Lagrange-d'Alembert principle asserts that the discrete flow is a critical value of the discrete action map $S_d : C_d^N(Q) \to \mathbb{R}$, which is still given by

$$S_d(q_d) = \sum_{k=0}^{N-1} L_d^h(q_k, q_{k+1}),$$

but this time we impose the restriction $\delta q_k \in \mathcal{D}_{q_k}$, that is, the infinitesimal variation of the sequence must lie in the constraint distribution. More formally, the discrete Lagrange-d'Alembert principle states the following:

Definition 3.7.16 (Discrete Lagrange-d'Alembert principle)**.** The discrete flow of the discrete nonholonomic Lagrangian system determined by the discrete Lagrangian function L_d^h, the distribution $\mathcal{D}$ and the discrete constraint space $\mathcal{M}_d$ satisfies the constraint $(q_k, q_{k+1}) \in \mathcal{M}_d$ for all $k \in \{0, ..., N-1\}$ and is a critical value of the discrete action map S_d among all variations of sequences with fixed end-points whose infinitesimal generators satisfy $\delta q_k \in \mathcal{D}_{q_k}$.

As it happens with its continuous counterpart, the application of the discrete Lagrange-d'Alembert principle leads to a set of equations which will be the necessary and sufficient conditions to find critical values subordinated to the imposed restrictions. Assume in the following that $\mu^a \in \Omega^1(Q)$ with $a = 1, ..., n-k$ are 1-forms on Q defining the distribution $\mathcal{D}$

$$\mathcal{D} = \{v \in TQ \mid \mu^a(v) = 0\}$$

and μ_d^a are a set of $n-k$ functions on $Q \times Q$ whose zero set is the discrete constraint space $\mathcal{M}_d$.

Theorem 3.7.17. *A sequence $\{q_k\}_{k=1}^N$ of points in Q satisfies the discrete-Lagrange d'Alembert principle associated to the triple $(L_d^h, \mathcal{D}, \mathcal{M}_d)$ if and only if it satisfies the equations*

$$\begin{aligned} &D_2 L_d^d(q_k, q_{k+1}) + D_1 L_d^d(q_{k-1}, q_k) = \lambda_a \mu^a \\ &\mu_d^a(q_k, q_{k+1}) = 0. \end{aligned} \tag{3.7.8}$$

Proof. The proof follows along the same lines of the proof of Proposition 3.7.4. Indeed, given q_0 and q_N in Q, consider a *variation of sequences* in $C_d^N(q_0, q_N)$ of the form

$$q_d(s) = \{q_0(s), q_1(s), ..., q_N(s)\},$$

where $q_0(s) = q_0$ and $q_N(s) = q_N$ are fixed and let $q_i(0) = q_i \in Q$. Moreover, let the infinitesimal variations lie in the distribution $\mathcal{D}$, i.e.,

$$\delta q_i = \left.\frac{d}{ds}\right|_{s=0} q_i(s) \in \mathcal{D}_{q_i}.$$

Then, the computation of the differential of the discrete action functional holds

$$\langle dS_d(q_d), \delta q_d \rangle = \sum_{k=1}^{N-1} \langle D_1 L_d^h(q_k, q_{k+1}) + D_2 L_d^h(q_{k-1}, q_k), \delta q_k \rangle.$$

The sequence is a critical value if and only if the differential of the discrete action map kills the infinitesimal variations with $\delta q_i \in \mathcal{D}$. Thus, this is equivalent to

$$D_2 L_d^d(q_k, q_{k+1}) + D_1 L_d^d(q_{k-1}, q_k) = \lambda_a \mu^a.$$

The second equation in (3.7.8) is just the assertion that the pair (q_k, q_{k+1}) belongs to the discrete constraint space $\mathcal{M}_d$. □

This method is demonstrated to have a good long-time behaviour in most of the cases, as well as satisfying a special case of Noether's theorem, in the presence of *horizontal symmetries*, simulating the general properties of nonholonomic systems.

However, in contrast with the unconstrained case, there is no evidence that the choice of the exact discrete Lagrangian function to be the discrete Lagrangian leads to the exact discrete dynamics of nonholonomic systems. This is a big issue in what matters numerical analysis and estimation of the error order because we now lack the notion of how far are Lagrange-d'Alembert equations from the exact dynamics. This is the main question we will address in Chapter 7.

Chapter 4

The nonholonomic exponential map

The goal of this chapter is to define the *nonholonomic exponential map* associated with a regular nonholonomic system $(L, \mathcal{D})$ with configuration space Q and dynamical vector field $\Gamma_{(L,\mathcal{D})} \in \mathfrak{X}(\mathcal{D})$ and to prove that it is a diffeomorphism onto its image, at least restricted to a small open neighbourhood. This is one of the core concepts contained in this thesis and plays an important role in all the original results presented in the subsequent chapters. While this construction is very common in Riemannian geometry and has been applied to other fields such as sub-Riemannian geometry (cf. [ABR18]) or the study of second-order vector fields (cf. [HM20]), the concept of exponential map was never previously introduced in the context of nonholonomic mechanics, as far as we know.

In order to accomplish our goal, we need to prove the regular character of the nonholonomic exponential map in the most general situation. For that end, we will use an arbitrary SODE extension $\Gamma \in \mathfrak{X}(TQ)$ of the nonholonomic vector field $\Gamma_{(L,\mathcal{D})} \in \mathfrak{X}(\mathcal{D})$ and the SODE exponential map $\exp_h^\Gamma$ associated with Γ at time $h > 0$. We will prove in Section 4.2 that such SODE extension vector fields always exist. The fact that $\exp_h^\Gamma$ is a diffeomorphism at least locally has been either guessed or assumed by many authors but we have only found a formal proof very recently in [MDM21] (see also [MDM16]). Anyway, in order to make the thesis more self-contained, in Section 4.1 we include the proof of this fact.

Finally, in the last section, we define the nonholonomic exponential map, we prove that at a local level it is a diffeomorphism onto its image and we take a look into some examples.

4.1 Exponential map for SODE vector fields on the tangent bundle

Let $\Gamma \in \mathfrak{X}(TQ)$ be a SODE vector field on TQ. As we have seen in Section 2.4.2, a trajectory of Γ satisfies in each coordinate chart a system of second order differential equations of the type

$$\ddot{q}^i = f^i(q, \dot{q}), \tag{4.1.1}$$

for some smooth functions f^i on TQ. Our goal in this section is to prove that the equations above have a unique solution satisfying the boundary conditions

$$q(0) = q_0, \quad q(h) = q_1, \tag{4.1.2}$$

with $q_0 \in Q$, $h > 0$ a sufficiently small number and q_1 a point in a sufficiently small neighbourhood $U \subseteq Q$ of q_0, and that the map $\beta : U \to \mathcal{U}_0 \subseteq T_{q_0}Q$, assigning to each $q_1 \in U$ the initial velocity of the unique trajectory of (4.1.1) satisfying the boundary conditions (4.1.2), is a diffeomorphism.

Though the problem is local, we will try to use a geometric language as much as possible. We remark that we are also obtaining an analytical theorem as a by-product of our approach: we are deducing that if f^i are smooth then there exists a unique solution of (4.1.1) satisfying (4.1.2) and the solution is smooth with respect to all variables.

A preliminary convexity result for a SODE Γ may be deduced using the theory of explicit second order differential equations (see [Har02]).

Theorem 4.1.1. *Let Γ be a SODE in Q and q_0 be a point of Q. Then, one may find a sufficiently small positive number h_0, a family of tangent vectors of Q at q_0,*

$$v_{(h,q_0)} \in T_{q_0}Q, \quad \textit{for } 0 < h \leq h_0,$$

and two compact subsets C and $\bar{C}$ of Q and TQ, respectively, with $q_0 \in C$ and $v_{(h,q_0)} \in \bar{C}$, such that there exists a unique trajectory of Γ

$$\sigma_{q_0 q_0 h} : [0, h] \to C \subseteq Q$$

satisfying

$$\sigma_{q_0 q_0 h}(0) = q_0, \quad \sigma_{q_0 q_0 h}(h) = q_0,$$

and

$$\dot{\sigma}_{q_0 q_0 h}(t) \in \bar{C}, \quad \textit{for every } t \in [0, h].$$

Proof. Let $(U, \varphi \equiv (q^i))$ be a local chart on Q such that

$$\varphi(U) = B(0; \epsilon) \text{ and } \varphi(q_0) = (0, \ldots, 0),$$

where $B(0; \epsilon)$ is the open ball in $\mathbb{R}^n$ with centre the origin and radius $\epsilon > 0$.

We consider the corresponding local coordinates $(\tau_Q^{-1}(U), \bar{\varphi} \equiv (q^i, \dot{q}^i))$ on TQ. Note that $\bar{\varphi}(\tau_Q^{-1}(U)) = \varphi(U) \times \mathbb{R}^n$. Since Γ is a SODE, we also have that

$$\Gamma = \dot{q}^i \frac{\partial}{\partial q^i} + \xi^i(q, \dot{q}) \frac{\partial}{\partial \dot{q}^i}.$$

Then, the trajectories of Γ in U are the solutions of the system of second order differential equations

$$\frac{d^2 q^i}{dt^2} = \xi^i(q, \frac{dq}{dt}), \quad \text{for all } i.$$

Now, if we take

$$0 < R < \epsilon \quad \text{and} \quad 0 < R'$$

then, using that ξ^i is a real C^∞-function on $B(0; \epsilon) \times \mathbb{R}^n$, we deduce that there exist positive constants κ, $\kappa' > 0$ satisfying

$$\|D_1 \xi(q, \dot{q})\| \leq \kappa, \quad \|D_2 \xi(q, \dot{q})\| \leq \kappa', \quad \text{for } (q, \dot{q}) \in \overline{B(0; R)} \times \overline{B(0; R')},$$

where $\overline{B(0; R)}$ and $\overline{B(0; R')}$ are the closed balls in $\mathbb{R}^n$ centred at the origin and with radius R and R', respectively. Thus, from Proposition B.0.3 (see Appendix B), it follows that

$$\begin{aligned} \|\xi^i(q_1^j, \dot{q}_1^j) - \xi^i(q_2^j, \dot{q}_2^j)\| &\leq \|\xi^i(q_1^j, \dot{q}_1^j) - \xi^i(q_2^j, \dot{q}_1^j)\| + \|\xi^i(q_2^j, \dot{q}_1^j) - \xi^i(q_2^j, \dot{q}_2^j)\| \\ &\leq \kappa \|q_2 - q_1\| + \kappa' \|\dot{q}_2 - \dot{q}_1\| \end{aligned}$$

for $(q_1^j, \dot{q}_1^j), (q_2^j, \dot{q}_2^j) \in B(0; R) \times B(0; R')$.

Moreover, it is clear that there exists a positive constant $M > 0$ such that

$$\|\xi(q^j, \dot{q}^j)\| \leq M, \quad \forall (q, \dot{q}) \in \overline{B(0; R)} \times \overline{B(0, R')}.$$

Next, we choose a sufficiently small positive number h_0 satisfying

$$\frac{\kappa h_0^2}{8} + \frac{\kappa' h_0}{2} < 1, \quad \frac{M h_0^2}{8} \leq R, \quad \frac{M h_0}{2} \leq R'.$$

Now, if we take $h \in \mathbb{R}$, $0 < h \leq h_0$ and the compact subsets C and $\bar{C}$ of Q and TQ, respectively, given by

$$C = \varphi^{-1}(\overline{B(0;R)}), \quad \bar{C} = \bar{\varphi}^{-1}(\overline{B(0;R)} \times \overline{B(0;R')})$$

then, using Theorem B.0.1 (see Appendix B), we conclude that there exists a unique trajectory $\sigma_{q_0 q_0 h} : [0,h] \to C \subseteq Q$ of Γ such that

$$\sigma_{q_0 q_0 h}(0) = q_0, \quad \sigma_{q_0 q_0 h}(h) = q_0,$$

and

$$\dot{\sigma}_{q_0 q_0 h}(t) \in \bar{C}, \quad \text{for } t \in [0,h].$$

Therefore, if we take $v_{(h,q_0)} = \dot{\sigma}_{q_0 q_0 h}(0)$, we end the proof of the result. □

Now, we will denote by ϕ^Γ the flow of the SODE Γ

$$\phi^\Gamma : M^\Gamma \subseteq \mathbb{R} \times TQ \to TQ.$$

Here, M^Γ is the open subset of $\mathbb{R} \times TQ$ given by

$$M^\Gamma = \{(t,v) \in \mathbb{R} \times TQ \mid \phi^\Gamma(\cdot, v) \text{ is defined at least in } [0,t]\}.$$

Now, if q_0 is a point of Q and $h \geq 0$, we may consider the open subset $M^\Gamma_{(h,q_0)}$ of $T_{q_0}Q$ given by

$$M^\Gamma_{(h,q_0)} = \{v \in T_{q_0}Q \mid (h,v) \in M^\Gamma\}.$$

Note that if $h > 0$ is sufficiently small then it is clear that $M^\Gamma_{(h,q_0)} \neq \emptyset$. We recall that the definition of the exponential map associated with Γ at q_0 for the time h is

$$\exp^\Gamma_{(h,q_0)}(v) = (\tau_Q \circ \phi^\Gamma_h)(v), \quad \text{for } v \in M^\Gamma_{(h,q_0)}. \tag{4.1.3}$$

We remark that the map $\exp^\Gamma_{(0,q_0)}$ is constant. However, we have the following result.

Theorem 4.1.2. *Let Γ be a SODE in Q and q_0 a point in Q. We take a sufficiently small positive real number h and $v_{(h,q_0)} \in T_{q_0}Q$ as in Theorem 4.1.1. Then,*

$$v_{(h,q_0)} \in M^\Gamma_{(h,q_0)}, \quad \exp^\Gamma_{(h,q_0)}(v_{(h,q_0)}) = q_0,$$

and

$$T_{v_{(h,q_0)}} \exp^\Gamma_{(h,q_0)} : T_{v_{(h,q_0)}} M^\Gamma_{(h,q_0)} \to T_{q_0}Q$$

is an isomorphism.

Proof. From Theorem 4.1.1, it follows that

$$v_{(h,q_0)} \in M^{\Gamma}_{(h,q_0)} \quad \text{and} \quad \exp^{\Gamma}_{(h,q_0)}(v_{(h,q_0)}) = q_0.$$

Moreover, it is clear that the map

$$\exp^{\Gamma}_{(h,q_0)} : M^{\Gamma}_{(h,q_0)} \subseteq T_{q_0}Q \to Q$$

is smooth.

Next, we will proceed locally. So, we will denote by

$$(t, q^i, \dot{q}^i) \to (x^j(t, q^i, \dot{q}^i), \dot{x}^j(t, q^i, \dot{q}^i))$$

the flow of the SODE Γ given by

$$\Gamma(q^j, \dot{q}^j) = \dot{q}^i \frac{\partial}{\partial q^i} + \xi^i(q^j, \dot{q}^j) \frac{\partial}{\partial \dot{q}^i},$$

so that the second order equations

$$\ddot{x}^i(t, q^j, \dot{q}^j) = \xi^i(x^k(t, q^j, \dot{q}^j), \dot{x}^k(t, q^j, \dot{q}^j)) \tag{4.1.4}$$

are satisfied as well as the following boundary conditions

$$x^i(0, q^j, \dot{q}^j) = q^i, \quad \dot{x}^i(0, q^j, \dot{q}^j) = \dot{q}^i. \tag{4.1.5}$$

The local expression of the map $\exp^{\Gamma}_{(h,q_0)}$ is

$$\dot{q}^i \to \exp^{\Gamma}_{(h,q_0)}(\dot{q}^i) = \left(x^j(h, q_0, \dot{q}^i)\right).$$

Denote by $\dot{q}_{0h}$ the tangent vector $v_{(h,q_0)} \in T_{q_0}Q$. We must prove that the Jacobian matrix of $\exp^{\Gamma}_{(h,q_0)}$ at $\dot{q}_{0h}$

$$(D_{\dot{q}} \exp^{\Gamma}_{(h,q_0)})(\dot{q}_{0h}) = (D_{\dot{q}} x)(h, q_0, \dot{q}_{0h})$$

is non-singular which, by the inverse function theorem, automatically implies that the map $\exp^{\Gamma}_{(h,q_0)}$ is a diffeomorphism on a local neighbourhood of $\dot{q}_{0h}$.

Denote by $U_{(q_0,\dot{q}_{0h})}(t)$ the Jacobian matrix of the smooth map $\exp^{\Gamma}_{(t,q_0)}$ at $\dot{q}_{0h}$, that is,

$$U_{(q_0,\dot{q}_{0h})}(t) = (D_{\dot{q}} \exp^{\Gamma}_{(t,q_0)})(\dot{q}_{0h}) = (D_{\dot{q}} x)(t, q_0, \dot{q}_{0h}).$$

Then from the second order system of equations (4.1.4), using a standard argument on the differentiability of solutions with respect to initial conditions, we may prove that

$$\begin{aligned}\ddot{U}_{(q_0,\dot{q}_{0h})}(t) &= (D_q\xi)(x^i(t,q_0,\dot{q}_{0h}),\dot{x}^i(t,q_0,\dot{q}_{0h}))U_{(q_0,\dot{q}_{0h})}(t)\\ &+ (D_{\dot{q}}\xi)(x^i(t,q_0,\dot{q}_{0h}),\dot{x}^i(t,q_0,\dot{q}_{0h}))\dot{U}_{(q_0,\dot{q}_{0h})}(t)\end{aligned}$$

and, in a similar way, using (4.1.5) we also deduce that

$$U_{(q_0,\dot{q}_{0h})}(0) = 0, \quad \dot{U}_{(q_0,\dot{q}_{0h})}(0) = Id.$$

So, if we denote by $B_{(q_0,\dot{q}_{0h})}(t)$ and $F_{(q_0,\dot{q}_{0h})}(t)$ the matrices

$$(D_q\xi)(x^i(t,q_0,\dot{q}_{0h}),\dot{x}^i(t,q_0,\dot{q}_{0h})) \text{ and } (D_{\dot{q}}\xi)(x^i(t,q_0,\dot{q}_{0h}),\dot{x}^i(t,q_0,\dot{q}_{0h})),$$

respectively, it follows that

$$\ddot{U}_{(q_0,\dot{q}_{0h})}(t) = B_{(q_0,\dot{q}_{0h})}(t)U_{(q_0,\dot{q}_{0h})}(t) + F_{(q_0,\dot{q}_{0h})}(t)\dot{U}_{(q_0,\dot{q}_{0h})}(t).$$

Now, we consider the linear system of second order differential equations

$$\ddot{y}(t) = B_{(q_0,\dot{q}_{0h})}(t)y(t) + F_{(q_0,\dot{q}_{0h})}(t)\dot{y}(t). \tag{4.1.6}$$

Note that $B_{(q_0,\dot{q}_{0h})}$ and $F_{(q_0,\dot{q}_{0h})}$ are C^∞-matrices, for every sufficiently small positive number h.

So, taking into account that there exists a compact subset $\bar{C} \subseteq TQ$ such that $v_{(h,q_0)} \in \bar{C}$ (for every h), using Theorem B.0.1 and proceeding as in the proof of Theorem 4.1.1, we conclude that there exists a sufficiently small positive number $p_0 > 0$ such that for all h the unique solution

$$t \to y_{(q_0,\dot{q}_{0h})}(t)$$

of the system (4.1.6) satisfying the boundary conditions

$$y_{(q_0,\dot{q}_{0h})}(0) = 0, \;\; y_{(q_0,\dot{q}_{0h})}(p) = 0, \quad \text{with } 0 < p \leq p_0,$$

is the trivial solution.

Thus, from Lemma B.0.2 in Appendix B, we deduce that the matrix

$$U_{(q_0,\dot{q}_{0h})}(p), \quad \text{with } 0 < p \leq p_0,$$

is regular, for every h.

Therefore, it is sufficient to take $h = p$, with $0 < p \leq p_0$, and the result is proved. □

From Theorem 4.1.2, we have that there exist open subsets $\mathcal{U}_0$ and U in $M^\Gamma_{(h,q_0)}$ and Q, respectively, with $v_{(h,q_0)} \in \mathcal{U}_0$ and $q_0 \in U$, such that the map

$$\exp^\Gamma_{(h,q_0)} : \mathcal{U}_0 \subseteq M^\Gamma_{(h,q_0)} \to U \subseteq Q$$

is a diffeomorphism.

Next, we will consider the open subset M^Γ_h of TQ given by

$$M^\Gamma_h = \{v \in TQ \mid (h,v) \in M^\Gamma\}.$$

Note that

$$v \in M^\Gamma_h \Longrightarrow M^\Gamma_{(h,\tau_Q(v))} = M^\Gamma_h \cap T_{\tau_Q(v)}Q \subseteq M^\Gamma_h.$$

Thus, since $\tau_Q : TQ \to Q$ is an open map, it follows that $\tau_Q(M^\Gamma_h)$ is an open subset of Q and

$$M^\Gamma_h = \bigcup_{q \in \tau_Q(M^\Gamma_h)} M^\Gamma_{(h,q)}.$$

Recall that the smooth map $\exp^\Gamma_h : M^\Gamma_h \subseteq TQ \to Q \times Q$ defined as

$$\exp^\Gamma_h(v) = (\tau_Q(v), \exp^\Gamma_{(h,\tau_Q(v))}(v)), \quad \text{for } v \in M^\Gamma_h$$

is the extended exponential map associated with the SODE Γ at time h.

Now, we deduce that

Lemma 4.1.3. *Let v be an element of M^Γ_h such that $exp^\Gamma_{(h,\tau_Q(v))}$ is non-singular at v. Then, exp^Γ_h is also non-singular at v.*

Proof. We must prove that the map

$$T_v(\exp^\Gamma_h) : T_v(M^\Gamma_h) \simeq T_v(TQ) \to T_{\tau_Q(v)}Q \times T_{\exp^\Gamma_{(h,\tau_Q(v))}(v)}Q$$

is a linear isomorphism.

Suppose that

$$0 = (T_v(\exp^\Gamma_h))(X_v), \quad \text{with } X_v \in T_v(M^\Gamma_h).$$

Then, we have that

$$0 = (T_v\tau_Q)(X_v) \quad \text{and} \quad 0 = (T_v\exp^\Gamma_{(h,\tau_Q(v))})(X_v).$$

The first condition implies that

$$X_v \in T_v(M_h^\Gamma \cap T_{\tau_Q(v)}Q) = T_v(M^\Gamma_{(h,\tau_Q(v))})$$

and thus, using the second one, we conclude that

$$X_v = 0.$$

□

As we know, if $h > 0$ is sufficiently small and $q_0 \in Q$ then the map $\exp^\Gamma_{(h,q_0)} : M^\Gamma_{(h,q_0)} \to Q$ is non-singular at the point $v_{(h,q_0)} \in M^\Gamma_{(h,q_0)}$. Therefore, using Lemma 4.1.3, we deduce the following result

Theorem 4.1.4. *Let Γ be a SODE in TQ and q_0 be a point of Q. Then, one may find a sufficiently small positive number h, an open subset $\mathcal{U}_h \subseteq M_h^\Gamma \subseteq TQ$, with $v_{(h,q_0)} \in \mathcal{U}_h$, and an open subset U of Q, with $q_0 \in U$, such that:*

1. *The map*
$$exp_h^\Gamma : \mathcal{U}_h \subseteq M_h^\Gamma \to U \times U \subseteq Q \times Q$$
is a diffeomorphism.

2. *For every couple $(q, q') \in U \times U$ there exists a unique trajectory of Γ*
$$\sigma_{qq'h} : [0, h] \to Q$$
satisfying
$$\sigma_{qq'h}(0) = q, \quad \sigma_{qq'h}(h) = q' \quad and \quad \dot{\sigma}_{qq'h}(0) \in \mathcal{U}_h.$$

We will denote by $R_h^{e^-} : U \times U \to \mathcal{U}_h$ (respectively, $R_h^{e+} : U \times U \to \mathcal{U}_h$) the inverse map of the diffeomorphism $\exp_h^\Gamma : \mathcal{U}_h \to U \times U$ (respectively, $\exp_h^\Gamma \circ \phi^\Gamma_{-h} : \phi^\Gamma_h(\mathcal{U}_h) \to U \times U$).

The maps

$$R_h^{e^-} : U \times U \subseteq Q \times Q \to \mathcal{U}_h \subseteq TQ \text{ and } R_h^{e^+} : U \times U \subseteq Q \times Q \to \phi_h^\Gamma(\mathcal{U}_h) \subseteq TQ$$

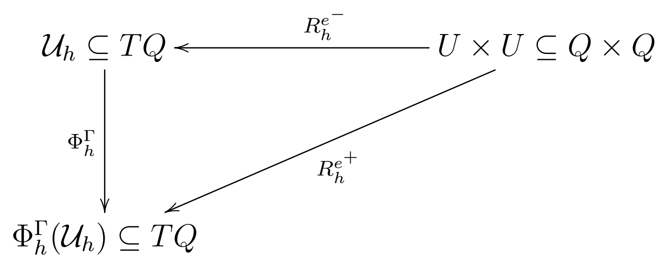
are called the *exact inverse retraction maps* associated with Γ. We have that

$$R_h^{e^-}(q, q') = \dot{\sigma}_{qq'h}(0), \quad R_h^{e^+}(q, q') = \dot{\sigma}_{qq'h}(h).$$

Note that

$$R_h^{e^+} = \phi_h^\Gamma \circ R_h^{e^-},$$

that is, the following diagram



is commutative.

In [MDM21] (see also [MDM16]), the authors give a generalized version of the previous theorem in the scope of SODE vector fields on Lie algebroids.

4.2 SODE extensions for the nonholonomic dynamics

We will see that *global SODE extensions* of the nonholonomic dynamics $\Gamma_{(L,\mathcal{D})} \in \mathfrak{X}(\mathcal{D})$, associated with a regular nonholonomic system $(L, \mathcal{D})$ with configuration space Q, always exist.

For this purpose, we will consider a Riemannian metric g on Q. Then, we have the orthogonal projectors

$$\mathcal{P} : TQ \to \mathcal{D}, \quad \mathcal{Q} : TQ \to \mathcal{D}^\perp$$

where $\mathcal{D}^\perp$ is the orthogonal complement to $\mathcal{D}$. Thus, we can define a vector bundle isomorphism

$$(\mathcal{P}, \mathcal{Q}) : TQ \to \mathcal{D} \oplus_Q \mathcal{D}^\perp$$

over the identity of Q. So, the tangent map to $(\mathcal{P}, \mathcal{Q})$

$$T(\mathcal{P}, \mathcal{Q}) : TTQ \to T(\mathcal{D} \oplus_Q \mathcal{D}^\perp) = T\mathcal{D} \oplus_{TQ} T\mathcal{D}^\perp$$

induces a vector bundle isomorphism (over the identity of TQ) between TTQ and

$$T\mathcal{D} \oplus_{TQ} T\mathcal{D}^\perp = \{(X, Y) \in T\mathcal{D} \times T\mathcal{D}^\perp \;/\; (T\tau)(X) = (T\tau^\perp)(Y)\}$$

where $\tau : \mathcal{D} \to Q$ and $\tau^\perp : \mathcal{D}^\perp \to Q$ are the canonical vector bundle projections. In fact, if $v_q \in T_qQ$ then

$$T_{v_q}(TQ) \simeq \{(X,Y) \in T_{\mathcal{P}(v_q)}\mathcal{D} \times T_{\mathcal{Q}(v_q)}\mathcal{D}^\perp \;/\; (T_{\mathcal{P}(v_q)}\tau)(X) = (T_{\mathcal{Q}(v_q)}\tau^\perp)(Y)\}.$$

Under the previous identifications, the canonical inclusion of $\mathcal{D}$ in TQ and the nonholonomic dynamics

$$i_{\mathcal{D}} : \mathcal{D} \to TQ \simeq \mathcal{D} \oplus_Q \mathcal{D}^\perp,\ \Gamma_{(L,\mathcal{D})} : \mathcal{D} \to T\mathcal{D} \subseteq TTQ \simeq T\mathcal{D} \oplus_{TQ} T\mathcal{D}^\perp$$

are given by

$$i_{\mathcal{D}}(u_q) = (u_q, 0^\perp(q)),\ \Gamma_{(L,\mathcal{D})}(u_q) = (\Gamma_{(L,\mathcal{D})}(u_q), (T_q 0^\perp)(u_q)),$$

for $u_q \in \mathcal{D}_q$, with $0^\perp : Q \to \mathcal{D}^\perp$ the zero section in $\mathcal{D}^\perp$. Moreover, a vector field $\Gamma : TQ \simeq \mathcal{D} \oplus_Q \mathcal{D}^\perp \to TTQ \simeq T\mathcal{D} \oplus_{TQ} T\mathcal{D}^\perp$ on $TQ \simeq \mathcal{D} \oplus_Q \mathcal{D}^\perp$

$$\Gamma(u_q, v_q) = (\Gamma_1(u_q, v_q), \Gamma_2(u_q, v_q)) \in T_{u_q}\mathcal{D} \times_{T_qQ} T_{v_q}\mathcal{D}^\perp,\ \ (u_q, v_q) \in \mathcal{D}_q \times \mathcal{D}_q^\perp$$

is a SODE if and only if

$$(T_{(u_q,v_q)}\pi)(\Gamma(u_q, v_q)) = u_q + v_q$$

with $\pi : \mathcal{D} \oplus_Q \mathcal{D}^\perp \to Q$ the canonical projection. So, Γ is a SODE if and only if

$$(T_{(u_q,v_q)}\tau)(\Gamma_1(u_q, v_q)) = u_q + v_q = (T_{(u_q,v_q)}\tau^\perp)(\Gamma_2(u_q, v_q)).$$

Now, using again the previous identifications, we can introduce a SODE $\Gamma : \mathcal{D} \oplus_Q \mathcal{D}^\perp \to T\mathcal{D} \oplus_{TQ} T\mathcal{D}^\perp$ on $\mathcal{D} \oplus_Q \mathcal{D}^\perp$, which extends $\Gamma_{(L,\mathcal{D})}$, given by

$$\Gamma(u_q, v_q) = (\Gamma_{\mathcal{D}}(u_q, v_q), \Gamma_{\mathcal{D}^\perp}(u_q, v_q))$$

with $\Gamma_{\mathcal{D}}$ and $\Gamma_{\mathcal{D}^\perp}$ defined as follows.

Definition of $\Gamma_{\mathcal{D}}$

If $f \in C^\infty(Q)$, $\alpha : Q \to \mathcal{D}^*$ is a section of the vector bundle $\tau^* : \mathcal{D}^* \to Q$ and $\hat{\alpha} : \mathcal{D} \to \mathbb{R}$ is the fiberwise linear function induced by α then there exists a unique tangent vector $\Gamma_{\mathcal{D}}(u_q, v_q) \in T_{u_q}\mathcal{D}$ which satisfies

$$\Gamma_{\mathcal{D}}(u_q, v_q)(f \circ \tau) = u_q(f) + v_q(f), \tag{4.2.1}$$

and

$$\Gamma_{\mathcal{D}}(u_q, v_q)(\hat{\alpha}) = \Gamma_{(L,\mathcal{D})}(u_q)(\hat{\alpha}) + T_{\mathcal{D}}(\alpha)(u_q, v_q), \tag{4.2.2}$$

where $T_{\mathcal{D}} : \Gamma(\mathcal{D}^*) \to \Gamma(\mathcal{D}^* \otimes_Q (\mathcal{D}^\perp)^*)$ is a $\mathbb{R}$-linear map and

$$T_{\mathcal{D}}(f\alpha)(u_q, v_q) = v_q(f)\alpha(q)(u_q) + f(q)T_{\mathcal{D}}(\alpha)(u_q, v_q). \tag{4.2.3}$$

For instance, we can take

$$T_{\mathcal{D}}(\alpha)(u_q, v_q) = (\nabla_{v_q}\alpha)(u_q),$$

with ∇ the Levi-Civita connection of g. Note that α may be considered as a 1-form on Q. In fact, α may be considered as a section of the annihilator $(\mathcal{D}^\perp)^0 \to Q$ of $\mathcal{D}^\perp$ (a vector subbundle of T^*Q).

We remark that (4.2.1) and (4.2.2) are compatible. In fact, using (4.2.1), (4.2.2), (4.2.3) and the fact that $(T_{u_q}\tau)(\Gamma_{(L,\mathcal{D})}(u_q)) = u_q$, it follows that

$$\begin{aligned}
\Gamma_{\mathcal{D}}(u_q, v_q)(\widehat{f\alpha}) &= \Gamma_{\mathcal{D}}(u_q, v_q)((f \circ \tau)\hat{\alpha}) \\
&= \Gamma_{(L,\mathcal{D})}(u_q)((f \circ \tau)\hat{\alpha}) + T_{\mathcal{D}}(f\alpha)(u_q, v_q) \\
&= (u_q(f) + v_q(f))\alpha(q)(u_q) \\
&+ f(q)(\Gamma_{(L,\mathcal{D})}(u_q)(\hat{\alpha}) + T_{\mathcal{D}}(\alpha)(u_q, v_q)) \\
&= \Gamma_{\mathcal{D}}(u_q, v_q)(f \circ \tau)\alpha(q)(u_q) + f(q)\Gamma_{\mathcal{D}}(u_q, v_q)(\hat{\alpha}).
\end{aligned}$$

So, there exists a unique tangent vector $\Gamma_{\mathcal{D}}(u_q, v_q) \in T_{u_q}\mathcal{D}$ which satisfies (4.2.1) and (4.2.2). In addition, from (4.2.1) and (4.2.2), we have that

$$(T_{u_q}\tau)(\Gamma_{\mathcal{D}}(u_q, v_q)) = u_q + v_q, \quad \Gamma_{\mathcal{D}}(u_q, 0^\perp(q)) = \Gamma_{(L,\mathcal{D})}(u_q). \tag{4.2.4}$$

Definition of $\Gamma_{\mathcal{D}^\perp}$

If $(u_q, v_q) \in \mathcal{D}_q \times \mathcal{D}_q^\perp$, $f \in C^\infty(Q)$, $\alpha^\perp : Q \to (\mathcal{D}^\perp)^*$ is a section of the vector bundle $(\tau^\perp)^* : (\mathcal{D}^\perp)^* \to Q$ and $\widehat{\alpha^\perp} : \mathcal{D}^\perp \to \mathbb{R}$ is the fiberwise linear function induced by $\alpha^\perp$ then there exists a unique tangent vector $\Gamma_{\mathcal{D}^\perp}(u_q, v_q) \in T_{v_q}\mathcal{D}^\perp$ which satisfies

$$\Gamma_{\mathcal{D}^\perp}(u_q, v_q)(f \circ \tau^\perp) = u_q(f) + v_q(f), \tag{4.2.5}$$

and

$$\Gamma_{\mathcal{D}^\perp}(u_q, v_q)(\widehat{\alpha^\perp}) = \mathcal{T}_{\mathcal{D}^\perp}(\alpha^\perp)(u_q, v_q), \tag{4.2.6}$$

where $\mathcal{T}_{\mathcal{D}^\perp} : \Gamma((\mathcal{D}^\perp)^*) \to \Gamma(\mathcal{D}^* \otimes_Q (\mathcal{D}^\perp)^*)$ is a $\mathbb{R}$-linear map and

$$\mathcal{T}_{\mathcal{D}^\perp}(f\alpha^\perp)(u_q, v_q) = (u_q + v_q)(f)\alpha^\perp(q)(v_q) + f(q)\mathcal{T}_{\mathcal{D}^\perp}(\alpha^\perp)(u_q, v_q). \quad (4.2.7)$$

That relations (4.2.5) and (4.2.6) are compatible may be proved using (4.2.7) and proceeding as in the definition of $\Gamma_{\mathcal{D}}$. Furthermore, from (4.2.5) and (4.2.6), it follows that

$$(T_{v_q}\tau)(\Gamma_{\mathcal{D}^\perp}(u_q, v_q)) = u_q + v_q, \;\; \Gamma_{\mathcal{D}^\perp}(u_q, 0^\perp(q)) = (T_q 0^\perp)(u_q). \quad (4.2.8)$$

Now, using the first relations in equations (4.2.4) and (4.2.8), we obtain that

$$\Gamma(u_q, v_q) \in T_{(u_q,v_q)}(\mathcal{D} \oplus_Q \mathcal{D}^\perp),$$

and that Γ is a SODE. On the other hand, using the second relations in Equations (4.2.4) and (4.2.8), we conclude that

$$\Gamma(u_q, 0^\perp(q)) = (\Gamma_{(L,\mathcal{D})}(u_q), (T_q 0^\perp)(u_q)),$$

and, thus, Γ is a SODE extension of the nonholonomic dynamics $\Gamma_{(L,\mathcal{D})}$.

An example of a map $\mathcal{T}_{\mathcal{D}^\perp} : \Gamma((\mathcal{D}^\perp)^*) \to \Gamma(\mathcal{D}^* \otimes_Q (\mathcal{D}^\perp)^*)$ satisfying the previous conditions may be obtained as follows. First of all, we will consider a vector field $\Gamma^\perp$ in $\mathcal{D}^\perp$ which vanishes on the zero section $0^\perp : Q \to \mathcal{D}^\perp$ and, in addition, it is a SODE along $\mathcal{D}^\perp$. For instance,

$$\Gamma^\perp(v_q) = (T_{v_q}\mathcal{Q})(\Gamma_g(v_q)), \quad \text{for } v_q \in \mathcal{D}^\perp_q, \quad (4.2.9)$$

where Γ_g is the geodesic flow associated with the Riemannian metric g. In fact, using that Γ_g is a SODE in TQ and that it vanishes on the zero section $0 : Q \to TQ$, we deduce that $\Gamma^\perp$, defined as in (4.2.9), is a SODE along $\mathcal{D}^\perp$ and it vanishes on the zero section $0^\perp : Q \to \mathcal{D}^\perp$. Next, we can take the $\mathbb{R}$-linear map $T_{\mathcal{D}^\perp} : \Gamma((\mathcal{D}^\perp)^*) \to \Gamma(\mathcal{D}^* \otimes_Q (\mathcal{D}^\perp)^*)$ given by

$$T_{\mathcal{D}^\perp}(\alpha^\perp)(u_q, v_q) = (\nabla_{u_q}\alpha^\perp)(v_q).$$

Then, our map $\mathcal{T}_{\mathcal{D}^\perp} : \Gamma((\mathcal{D}^\perp)^*) \to \Gamma(\mathcal{D}^* \otimes_Q (\mathcal{D}^\perp)^*)$ may be defined by

$$\mathcal{T}_{\mathcal{D}^\perp}(\alpha^\perp)(u_q, v_q) = \Gamma^\perp(v_q)(\widehat{\alpha^\perp}) + T_{\mathcal{D}^\perp}(\alpha^\perp)(u_q, v_q).$$

4.3 Exponential map and the exact discrete submanifold for the nonholonomic dynamics

Let $L : TQ \to \mathbb{R}$ be a regular Lagrangian function and $\mathcal{D}$ a regular distribution on Q such that the non-holonomic system $(L, \mathcal{D})$ is also regular and let $\Gamma_{(L,\mathcal{D})}$ be the SODE on $\mathcal{D}$ which is solution of the non-holonomic dynamics. Denote by $\phi_t^{\Gamma_{(L,\mathcal{D})}} : \mathcal{D} \to \mathcal{D}$ the flow of $\Gamma_{(L,\mathcal{D})}$ and for h a sufficiently small positive number, we consider the open subset of $\mathcal{D}$ given by

$$M_h^{\Gamma_{(L,\mathcal{D})}} = \{v \in \mathcal{D} \mid \phi_t^{\Gamma_{(L,\mathcal{D})}}(v) \text{ is defined for } t \in [0, h]\}.$$

Note that, if $\Gamma \in \mathfrak{X}(TQ)$ is a SODE extension of $\Gamma_{(L,\mathcal{D})}$ then

$$M_h^{\Gamma_{(L,\mathcal{D})}} = M_h^{\Gamma} \cap \mathcal{D}.$$

Definition 4.3.1. The map

$$\begin{aligned} \exp_h^{\Gamma_{(L,\mathcal{D})}} : M_h^{\Gamma_{(L,\mathcal{D})}} \subseteq \mathcal{D} &\to Q \times Q \\ v \in \mathcal{D} &\mapsto (\tau_Q(v), (\tau_Q \circ \phi_h^{\Gamma_{(L,\mathcal{D})}})(v)) \in Q \times Q \end{aligned} \tag{4.3.1}$$

is called the *nonholonomic exponential map of* $\Gamma_{(L,\mathcal{D})}$ *at time* h.

Now, we may prove the following result

Theorem 4.3.2. *Let $(L, \mathcal{D})$ be a regular nonholonomic system with configuration space Q and q_0 a point in Q. Then, one may find a sufficiently small positive number h, an open subset $\mathcal{U}_h^{nh} \subseteq M_h^{\Gamma_{(L,\mathcal{D})}} \subseteq \mathcal{D}$ and an open subset $U \subseteq Q$, with $q_0 \in U$, such that the map*

$$\exp_h^{\Gamma_{(L,\mathcal{D})}} : \mathcal{U}_h^{nh} \subseteq M_h^{\Gamma_{(L,\mathcal{D})}} \to U \times U$$

is an embedding.

Proof. Let Γ be a SODE in TQ such that $\Gamma_{|\mathcal{D}} = \Gamma_{(L,\mathcal{D})}$ (see Section 4.2). Then, using Theorem 4.1.4, we may find a sufficiently small positive number h, an open subset $\mathcal{U}_h \subseteq M_h^{\Gamma} \subseteq TQ$, with $v_{(h,q_0)} \in \mathcal{U}_h$, and an open subset U of Q, with $q_0 \in U$, such that the map

$$\exp_h^{\Gamma} : \mathcal{U}_h \subseteq M_h^{\Gamma} \to U \times U \subseteq Q \times Q$$

is a diffeomorphism.

Now, since $\Gamma_{|\mathcal{D}} = \Gamma_{(L,\mathcal{D})}$, it is clear that

$$(\exp_h^{\Gamma})_{|\mathcal{U}_h \cap M_h^{\Gamma_{(L,\mathcal{D})}}} = \exp_h^{\Gamma_{(L,\mathcal{D})}}.$$

So, if we take the open subset of $\mathcal{D}$

$$\mathcal{U}_h^{nh} = \mathcal{U}_h \cap M_h^{\Gamma_{(L,\mathcal{D})}}$$

then, using that every immersion is a local embedding, we can suppose (without the loss of generality) that the map $\exp_h^{\Gamma_{(L,\mathcal{D})}} : \mathcal{U}_h^{nh} \to U \times U$ is an embedding. $\square$

Now, we may introduce the following definition.

Definition 4.3.3. *The exact discrete nonholonomic constraint submanifold* is the submanifold of $Q \times Q$ given by

$$\mathcal{M}_h^{e,nh} = \exp_h^{\Gamma_{(L,\mathcal{D})}}(\mathcal{U}_h^{nh}) \, .$$

In view of Theorem 4.3.2, the map

$$\exp_h^{\Gamma_{(L,\mathcal{D})}} : \mathcal{U}_h^{nh} \to \mathcal{M}_h^{e,nh}$$

is a diffeomorphism and we can define its inverse diffeomorphism, called the *nonholonomic exact inverse retraction* map

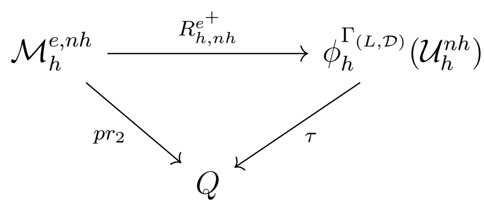
$$R_{h,nh}^{e^-} : \mathcal{M}_h^{e,nh} \longrightarrow \mathcal{U}_h^{nh}.$$

The following diagrams are commutative:

$$\begin{array}{ccc} \mathcal{U}_h^{nh} & \xrightarrow{\exp_h^{\Gamma_{(L,\mathcal{D})}}} & \mathcal{M}_h^{e,nh} \\ & {\scriptstyle\tau}\searrow \quad \swarrow{\scriptstyle pr_1} & \\ & Q & \end{array} \qquad \begin{array}{ccc} \mathcal{M}_h^{e,nh} & \xrightarrow{R_{h,nh}^{e^-}} & \mathcal{U}_h^{nh} \\ & {\scriptstyle pr_1}\searrow \quad \swarrow{\scriptstyle\tau} & \\ & Q & \end{array}$$

We will also use the map: $R_{h,nh}^{e^+} : \mathcal{M}_h^{e,nh} \longrightarrow \phi_h^{\Gamma_{(L,\mathcal{D})}}(\mathcal{U}_h^{nh})$ defined by

$$R_{h,nh}^{e^+} = \phi_h^{\Gamma_{(L,\mathcal{D})}} \circ R_{h,nh}^{e^-}$$



We will compute and analyse the trajectories of two of the nonholonomic systems we have seen before, namely, the nonholonomic particle and the vertical rolling disk, in order to have a practical understanding of the objects we have just defined.

Example 4.3.4. Consider the nonholonomic particle in Example 3.6.5. The nonholonomic vector field is given by

$$\Gamma_{(L,\mathcal{D})} = \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + y\dot{x}\frac{\partial}{\partial z} - y\frac{\dot{x}\dot{y}}{1+y^2}\frac{\partial}{\partial \dot{x}} + \frac{\dot{x}\dot{y}}{1+y^2}\frac{\partial}{\partial \dot{z}}$$

and the equations for this system have an explicit solution given by

$$\begin{cases} x_{nh}(t) = \frac{\dot{x}_0}{\dot{y}_0}\sqrt{y_0^2+1}(\operatorname{arcsinh}(\dot{y}_0 t + y_0) - \operatorname{arcsinh}(y_0)) + x_0 \\ y_{nh}(t) = \dot{y}_0 t + y_0 \\ z_{nh}(t) = \frac{\dot{x}_0}{\dot{y}_0}\sqrt{y_0^2+1}(\sqrt{(\dot{y}_0 t + y_0)^2+1} - \sqrt{y_0^2+1}) + z_0, \quad \text{if } \ \dot{y}_0 \neq 0, \end{cases} \tag{4.3.2}$$

or

$$\begin{cases} x_{nh}(t) = \dot{x}_0 t + x_0 \\ y_{nh}(t) = y_0 \\ z_{nh}(t) = y_0\dot{x}_0 t + z_0, \quad \text{if } \ \dot{y}_0 = 0. \end{cases} \tag{4.3.3}$$

From (4.3.2) and (4.3.3), we construct its corresponding flow and nonholonomic exponential map

$$\phi_t^{\Gamma_{(L,\mathcal{D})}}(x_0, y_0, z_0, \dot{x}_0, \dot{y}_0, y_0\dot{x}_0) = (x_{nh}, y_{nh}, z_{nh}, \dot{x}_{nh}, \dot{y}_{nh}, \dot{z}_{nh}),$$

$$\exp_h^{\Gamma_{(L,\mathcal{D})}}(x_0, y_0, z_0, \dot{x}_0, \dot{y}_0, y_0\dot{x}_0) = (x_0, y_0, z_0, x_{nh}(h), y_{nh}(h), z_{nh}(h)).$$

We see that this is an invertible map, when we restrict the co-domain to its image, and we may explicitly compute the inverse to be

$$R_{h,nh}^{e^-}(x_0, y_0, z_0, x_1, y_1, z_1) = \Bigg(x_0, y_0, z_0, \frac{(x_1-x_0)(y_1-y_0)}{h\sqrt{y_0^2+1}(\operatorname{arcsinh}(y_1) - \operatorname{arcsinh}(y_0))},$$
$$\frac{y_1-y_0}{h}, \frac{y_0(x_1-x_0)(y_1-y_0)}{h\sqrt{y_0^2+1}(\operatorname{arcsinh}(y_1) - \operatorname{arcsinh}(y_0))}\Bigg),$$

in the case where $y_1 \neq y_0$. Note that the domain of the map $R^{e^-}_{h,nh}$ is not $\mathbb{R}^3 \times \mathbb{R}^3$, it is restricted to $\mathcal{M}_h^{e,nh}$, which explicitly means that

$$\frac{z_1 - z_0}{h} - \frac{(x_1 - x_0)\left(\sqrt{y_1^2+1} - \sqrt{y_0^2+1}\right)}{h(\operatorname{arcsinh}(y_1) - \operatorname{arcsinh}(y_0))} = 0. \tag{4.3.4}$$

In fact, let the left-hand side of equation (4.3.4) be denoted by $\mu_d : Q \times Q \to \mathbb{R}$. It is a constraint function whose annihilation gives the discrete space $\mathcal{M}_h^{e,nh}$. △

Example 4.3.5. The Lagrange-d'Alembert equations for the vertical rolling disk are easily integrated. Indeed, in equations (3.6.4) we may substitute the value of λ_1 and λ_2 in the third equation, by the first and second equations. Moreover, the values of $\ddot{x}$ and $\ddot{y}$ are determined by the constraints. In this way, the solution is completely described by the four equations

$$\begin{cases} (I + mR^2)\ddot{\theta} = 0 \\ J\ddot{\varphi} = 0 \\ \dot{x} = R\cos\varphi\ \dot{\theta} \\ \dot{y} = R\sin\varphi\ \dot{\theta}. \end{cases}$$

Thus, the solution of these system is given by

$$\begin{cases} \theta(t) = \Omega t + \theta_0 \\ \varphi(t) = \omega t + \varphi_0 \\ x(t) = \frac{\Omega}{\omega}R\sin(\omega t + \varphi_0) + \left(x_0 - \frac{\Omega}{\omega}R\sin\varphi_0\right) \\ y(t) = -\frac{\Omega}{\omega}R\cos(\omega t + \varphi_0) + \left(y_0 + \frac{\Omega}{\omega}R\cos\varphi_0\right), \quad \text{if} \ \ \omega \neq 0, \end{cases} \tag{4.3.5}$$

and

$$\begin{cases} \theta(t) = \Omega t + \theta_0 \\ \varphi(t) = \varphi_0 \\ x(t) = \Omega R\cos(\varphi_0)t + x_0 \\ y(t) = \Omega R\sin(\varphi_0)t + y_0, \quad \text{if} \ \ \omega = 0, \end{cases} \tag{4.3.6}$$

Hence, if $(x_0, y_0, \theta_0, \varphi_0, \Omega, \omega)$ are coordinates on $\mathcal{D}$, the nonholonomic exponential map is given by the map

$$\exp_h^{\Gamma_{(L,\mathcal{D})}}(x_0, y_0, \theta_0, \varphi_0, \Omega, \omega) = (x_0, y_0, \theta_0, \varphi_0, x(h), y(h), \theta(h), \varphi(h))$$

which is a diffeomorphism between $\mathcal{D}$ and the submanifold $\mathcal{M}_h^{e,nh}$ of $(\mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1) \times (\mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1)$ determined by the local expressions

$$x_1 - x_0 = \frac{\theta_1 - \theta_0}{\varphi_1 - \varphi_0} R(\sin\varphi_1 - \sin\varphi_0), \quad y_1 - y_0 = \frac{\theta_1 - \theta_0}{\varphi_1 - \varphi_0} R(\cos\varphi_0 - \cos\varphi_1).$$

Hence, using the coordinates $(x_0, y_0, \theta_0, \varphi_0, \theta_1, \varphi_1)$ on $\mathcal{M}_h^{e,nh}$ and the previously mentioned coordinates on $\mathcal{D}$, the exact retraction reads

$$R^{e^-}_{h,nh}(x_0, y_0, \theta_0, \varphi_0, \theta_1, \varphi_1) = \left(x_0, y_0, \theta_0, \varphi_0, \frac{\theta_1 - \theta_0}{h}, \frac{\varphi_1 - \varphi_0}{h}\right).$$

△

Chapter 5

Radial trajectories of nonholonomic mechanical systems

In this chapter, we will start a program relating mechanical nonholonomic systems, that is, nonholonomic systems $(L, \mathcal{D})$ with configuration space Q and where the Lagrangian is of mechanical type

$$L(v_q) = \frac{1}{2} g(v_q, v_q) - V(q), \quad q \in Q, \ v_q \in T_q Q,$$

with g a Riemannian metric on Q and $V \in C^\infty(Q)$ the potential energy, with typical constructions from Riemannian geometry. Our main goal in this chapter will be to characterize the nonholonomic trajectories with fixed starting point as length minimizing geodesics in a prescribed Riemannian manifold. Together with the fact that the nonholonomic exponential map is a diffeomorphism onto its image at a local level, this collection of results leaves the door open for further developments in this direction. The fact that the exponential map is regular has a clearer simpler proof when the Lagrangian is kinetic.

In the first section, we will review the definition of nonholonomic connection proposed by [Lew98] which allows us to see the nonholonomic trajectories as geodesics of a non-Levi-Civita connection. In the second section, we will prove that nonholonomic trajectories are indeed geodesics with respect to a proper family of Riemannian metrics characterized by the Gauss condition. Then, on the next section, we will generalize the results to mechanical systems, using a nonholonomic version of the Maupertuis principle.

5.1 The nonholonomic connection

We have seen that trajectories of unconstrained mechanical Lagrangian systems are the geodesics of the Levi-Civita connection in the absence of a potential energy (see Section 3.2).

When a mechanical Lagrangian system is subjected to nonholonomic constraints, a similar result holds. The first to observe this was Synge ([Syn28]), who observed that nonholonomic trajectories satisfy a geodesic equation, though the associated connection, for which nonholonomic trajectories are geodesics, was not introduced until [Lew98], where the author studies further properties of this connection, to which we will call the *nonholonomic connection* (see also [BC95; BC98]). We recall the definition of the nonholonomic connection with a slight modification: we will start with a semi-Riemannian metric instead of a purely Riemannian metric. This modification will be useful to our purposes later.

If h is a semi-Riemannian metric and $\mathcal{D}$ is a distribution on Q, then $\mathcal{D}^\perp$ will denote the *orthogonal distribution*, i.e., at each point $q \in Q$, $\mathcal{D}^\perp_q$ is the orthogonal complementary subspace to $\mathcal{D}_q$ with respect to the inner product h_q.

Definition 5.1.1. Let (Q, h) be a semi-Riemannian manifold. A distribution $D \subseteq TQ$ of the smooth manifold Q is said to be *regular with respect to* h if $\mathcal{D} \cap \mathcal{D}^\perp = \{0\}$.

Note that if the semi-Riemannian metric was in fact a Riemannian metric, i.e., h is a positive definite tensor, then the distribution $\mathcal{D}$ would automatically be regular. However, in general, there is no reason why we should have $\mathcal{D} \cap \mathcal{D}^\perp = \{0\}$. For example, consider the semi-Riemannian metric h induced by the non-degenerate symmetric bilinear form on $\mathbb{R}^2$ also denoted by h given by

$$\begin{aligned} h : \mathbb{R}^2 \times \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (v, w) &\mapsto v_1 w_1 - v_2 w_2. \end{aligned}$$

h is obviously a semi-Riemannian metric in $\mathbb{R}^2$ but if $\mathcal{D}$ is the distribution spanned by the vector field $\frac{\partial}{\partial x} + \frac{\partial}{\partial y}$, then its orthogonal complement is exactly $\mathcal{D}$!

Lemma 5.1.2. *The distribution $\mathcal{D}$ is regular with respect to a semi-Riemannian metric h if and only if $TQ = \mathcal{D} \oplus_Q \mathcal{D}^\perp$.*

Proof. We may prove that (see Lemma 22, Chapter 2 in [O'N83])

$$\dim(\mathcal{D}_q) + \dim(\mathcal{D}_q^\perp) = \dim(T_qQ) = n, \quad \forall q \in Q.$$

Thus, applying the equation

$$\dim(\mathcal{D}_q + \mathcal{D}_q^\perp) = \dim(\mathcal{D}_q) + \dim(\mathcal{D}_q^\perp) - \dim(\mathcal{D}_q \cap \mathcal{D}_q^\perp)$$

we deduce that $T_qQ = \mathcal{D}_q \oplus \mathcal{D}_q^\perp$ if and only if the distribution is regular with respect to h. □

Hence, if $\mathcal{D}$ is regular with respect to h, we may consider the orthogonal projectors $P : TQ \to \mathcal{D}$ and $P' : TQ \to \mathcal{D}^\perp$.

Definition 5.1.3. Let h be a semi-Riemannian metric and ∇^h the associated Levi-Civita connection. Let $\mathcal{D}$ be a constraint distribution on Q, regular with respect to h. The *nonholonomic connection* is the linear connection $\nabla^{nh} : \mathfrak{X}(Q) \times \mathfrak{X}(Q) \to \mathfrak{X}(Q)$ given by

$$\nabla_X^{nh} Y = \nabla_X^h Y + (\nabla_X^h P')(Y). \tag{5.1.1}$$

Remark 5.1.4. We will often use the equivalent expression for the nonholonomic connection given by

$$\nabla_X^{nh} Y = P(\nabla_X^h Y) + \nabla_X^h (P'(Y)). \tag{5.1.2}$$

Remark 5.1.5. If h is a Riemannian metric, then the regularity condition is automatically satisfied and the nonholonomic connection is well-defined for every distribution.

This connection is not symmetric, in general, neither it is compatible with the metric. Nevertheless, it satisfies the more restrictive condition of compatibility with the metric over sections of $\mathcal{D}$ (see [Lew98]), i.e.,

$$X(h(Y, Z)) = h(\nabla_X^{nh} Y, Z) + h(Y, \nabla_X^{nh} Z), \quad \forall X, Y, Z \in \Gamma(\mathcal{D}). \tag{5.1.3}$$

Another important property that we will use is the following one: if $Y \in \Gamma(\mathcal{D})$, then $\nabla_X^{nh} Y = P(\nabla_X^h Y)$ for any vector field $X \in \mathfrak{X}(Q)$, as a consequence of (5.1.2).

If h is a pseudo-Riemannian metric on Q, recall that $L_h : TQ \to \mathbb{R}$ is the kinetic Lagrangian function associated with h defined by

$$L_h(v) = \frac{1}{2} h(v, v), \quad v \in TQ.$$

Note that, using the fact that h is non-degenerate, we deduce that the Poincaré-Cartan 2-form $\omega_{L_h} = -d\theta_{L_h}$ is symplectic (see Section 3.1).

We will show first a useful property relating the symplectic and the metric structures, which we will use in the proof of the theorem below.

Lemma 5.1.6. *Let h be a pseudo-Riemannian metric on Q, L_h the associated kinetic Lagrangian and ω_{L_h} the associated symplectic form on TQ. Denote by $\sharp_h : T^*Q \to TQ$ and $\sharp_{\omega_{L_h}} : T^*TQ \to TTQ$ the musical isomorphisms with respect to the metric and symplectic form, respectively. Then for any $\alpha \in \Omega^1(Q)$ we have*

$$\sharp_{\omega_{L_h}} \circ \alpha^{\mathbf{v}} = -(\sharp_h \circ \alpha)^{\mathbf{v}},$$

where $\alpha^{\mathbf{v}} \in \Omega^1(TQ)$ and $(\sharp_h \circ \alpha)^{\mathbf{v}} \in \mathfrak{X}(TQ)$ are the vertical lifts to TQ of the 1-form α and the vector field $\sharp_h \circ \alpha$, respectively (see Appendix A).

Proof. It follows using the first relation in (A.0.5) (see Appendix A). □

Now we will see that a distribution is regular with respect to a pseudo-Riemannian metric if and only if the induced nonholonomic system is regular.

Theorem 5.1.7. *Given a pseudo-Riemannian metric h on a manifold Q and a distribution $\mathcal{D}$, the following are equivalent:*

1. *$\mathcal{D} \cap \mathcal{D}^\perp = \{0\}$, where $\mathcal{D}^\perp$ is the orthogonal distribution with respect to h;*
2. *The nonholonomic system $(L_h, \mathcal{D})$ is regular.*

Proof. We recall that the nonholonomic system $(L, \mathcal{D})$ is regular if and only if

$$T_v\mathcal{D} \cap G_v = \{0\}, \quad \forall v \in \mathcal{D},$$

where G is the distribution on TQ along $\mathcal{D}$ defined by $G = \sharp_{\omega_{L_h}}((\mathcal{D}^o)^{\mathbf{v}})$ (see Section 3.6.2).

So, suppose first that $\mathcal{D} \cap \mathcal{D}^\perp = \{0\}$ and take $X_u \in T_u\mathcal{D} \cap G_u$.

Since $X_u \in G_u$, then there exists $\alpha \in \mathcal{D}^o$ such that

$$X_u = \sharp_{\omega_{L_h}}(\alpha_u^{\mathbf{v}}) = -(\sharp_h(\alpha))_u^{\mathbf{v}} \in T_u\mathcal{D},$$

where the last equation follows from Lemma 5.1.6.

Therefore, $\sharp_h(\alpha)$ is in $\mathcal{D}$, but since $\sharp_h(\mathcal{D}^o) = \mathcal{D}^\perp$, it must be the zero vector. Hence $X_u = -(\sharp_h(\alpha))_u^{\mathbf{v}} = 0$.

Conversely, if $u \in \mathcal{D} \cap \mathcal{D}^\perp$, there exists $\alpha \in \mathcal{D}^o$ such that $u = \sharp_h(\alpha)$. Now, using Lemma 5.1.6, the vector

$$(\sharp_h(\alpha))_u^{\mathbf{v}} \in T_u\mathcal{D} \cap G_u = \{0\},$$

then $\sharp_h(\alpha) = 0$ or, in other words, $u = 0$. □

By the theorem above, if the distribution $\mathcal{D}$ is regular with respect to h then the nonholonomic system $(L_h, \mathcal{D})$ is regular and we may consider the nonholonomic SODE $\Gamma_{(L_h,\mathcal{D})}$ on $\mathcal{D}$, as a consequence of Theorem 3.6.8. We will show next that the trajectories of $\Gamma_{(L_h,\mathcal{D})}$ are just the geodesics of the associated nonholonomic connection ∇^{nh} with initial velocities in $\mathcal{D}$.

Theorem 5.1.8. *Let h be a pseudo-Riemannian metric and $L_h : TQ \to \mathbb{R}$ its associated Lagrangian. If $\mathcal{D}$ is a nonholonomic distribution satisfying $\mathcal{D} \cap \mathcal{D}^\perp = \{0\}$, then the trajectories of $\Gamma_{(L_h,\mathcal{D})}$ are the geodesics of the connection ∇^{nh} with initial velocities in $\mathcal{D}$.*

Proof. Let $c_v : I \to Q$ be a trajectory of $\Gamma_{(L_h,\mathcal{D})}$ with initial velocity $\dot{c}_v(0) = v \in \mathcal{D}$. We must prove that

$$\nabla^{nh}_{\dot{c}_v}\dot{c}_v = 0.$$

Given any $X \in \Gamma(\mathcal{D})$, we will apply the geometric equation which defines $\Gamma_{(L_h,\mathcal{D})}$ to the complete lift $X^{\mathbf{c}}$ of X (see Section 2.4.1) at points in $\mathcal{D}$. In fact, given $u \in \mathcal{D}$, there exists $\mu \in (TD)^o$ such that

$$\langle \left(i_{\Gamma_{(L_h,\mathcal{D})}}\omega_{L_h} - dL_h \right)(u), X^{\mathbf{c}}(u)\rangle = \langle S^*(\mu)(u), X^{\mathbf{c}}(u)\rangle.$$

Using the skew-symmetry of ω_{L_h} and the fact that $SX^{\mathbf{c}} = X^{\mathbf{v}}$ (see Section 2.4.1) we get

$$-\langle i_{X^{\mathbf{c}}}\omega_{L_h}, \Gamma_{(L_h,\mathcal{D})}\rangle - X^{\mathbf{c}}(L_h) = \langle \mu, X^{\mathbf{v}}\rangle.$$

Note that the right-hand side vanishes because $X^{\mathsf{v}} \in \mathfrak{X}(\mathcal{D})$ and $\mu \in (TD)^o$. Also, using Lemma A.0.1 and equation (A.0.5) from Appendix A on the left-hand side of the previous equations we deduce

$$-d(\widehat{\flat_h(X)})(\Gamma_{(L_h,\mathcal{D})}) + 2\theta_{L_{(\nabla^h X)}}(\Gamma_{(L_h,\mathcal{D})}) - L_{\mathcal{L}_X h} = 0,$$

where $(\nabla^h X)$ is the $(0,2)$-tensor field defined by

$$(\nabla^h X)(Y,Z) = h(\nabla^h_X Y, Z),$$

∇^h the Levi-Civita connection associated with h and $\widehat{\flat_h(X)}$ is the fiberwise linear function on TQ induced by the 1-form $\flat_h(X)$.

By Lemma A.0.2, we deduce that $L_{\mathcal{L}_X h} = 2L_{(\nabla^h X)}$. Moreover, since the vector field $\Gamma_{(L_h,\mathcal{D})}$ is a SODE along $\mathcal{D}$, it follows that $S\Gamma_{(L_h,\mathcal{D})} = \Delta|_{\mathcal{D}}$, with Δ being the Liouville vector field on TQ and

$$\theta_{L_{(\nabla^h X)}}(\Gamma_{(L_h,\mathcal{D})}) = \Delta(L_{(\nabla^h X)})|_{\mathcal{D}} = 2L_{(\nabla^h X)}|_{\mathcal{D}}.$$

So the equation boils down to

$$\Gamma_{(L_h,\mathcal{D})}(u)(\widehat{\flat_h(X)}) = 2L_{\nabla^h X}(u) = h(\nabla^h_u X, u).$$

Evaluating the last equation over the curve $\dot{c}_v$ and noting that $\Gamma_{(L_h,\mathcal{D})}(\dot{c}_v)$ is just $\ddot{c}_v$, we deduce

$$\ddot{c}_v\left(\widehat{\flat_h(X)}\right) = h(\nabla^h_{\dot{c}_v} X, \dot{c}_v).$$

Then, of course,

$$\frac{d}{dt}\left(\widehat{\flat_h(X)}(\dot{c}_v)\right) = h(\nabla^h_{\dot{c}_v} X, \dot{c}_v),$$

which is by definition

$$\frac{d}{dt}\left(h(X \circ c_v, \dot{c}_v)\right) = h(\nabla^h_{\dot{c}_v} X, \dot{c}_v).$$

Using the fact that the connection is compatible with the metric, the previous equation reduces to

$$h(\nabla^h_{\dot{c}_v} X, \dot{c}_v) + h(X \circ c_v, \nabla^h_{\dot{c}_v} \dot{c}_v) = h(\nabla^h_{\dot{c}_v} X, \dot{c}_v)$$

where the first term on the left-hand side cancels with the term on the right-hand side, giving

$$h(X \circ c_v, \nabla^h_{\dot{c}_v} \dot{c}_v) = 0.$$

Since X is an arbitrary section in $\Gamma(\mathcal{D})$ we conclude that $P(\nabla^h_{\dot{c}_v}\dot{c}_v) = 0$. But, since $\dot{c}_v \in \mathcal{D}$, the connection is forced to satisfy $P(\nabla^h_{\dot{c}_v}\dot{c}_v) = \nabla^{nh}_{\dot{c}_v}\dot{c}_v$. Hence, we conclude

$$\nabla^{nh}_{\dot{c}_v}\dot{c}_v = 0.$$

□

Using Theorem 5.1.8, we will describe the action of the nonholonomic SODE $\Gamma_{(L_h,\mathcal{D})}$ on basic and fiberwise linear functions on $\mathcal{D}$.

Note that a basic function on $\mathcal{D}$ is of the form $f \circ \tau_{\mathcal{D}}$, with $f \in C^\infty(Q)$ and $\tau_{\mathcal{D}} : \mathcal{D} \to Q$ the vector bundle projection. On the other hand, a fiberwise linear function on $\mathcal{D}$ is given by $\widehat{\alpha}$, with $\alpha \in \Gamma(\mathcal{D}^*)$.

In addition, a fiberwise quadratic function on $\mathcal{D}$ has the form $T^{\mathfrak{q}}$, with T a section of the vector bundle $\mathcal{D}^* \otimes \mathcal{D}^* \to Q$ and

$$T^{\mathfrak{q}}(v) = T(v, v), \quad v \in \mathcal{D}. \tag{5.1.4}$$

Remark 5.1.9. If U is an open subset U of Q with local coordinates (q^i), $\{e_a\}$ is a local basis of sections of $\Gamma(\mathcal{D})$, $\{e^a\}$ the dual basis of $\Gamma(\mathcal{D}^*)$ and

$$\alpha = \alpha_a(q)e^a, \quad T = T_{ab}(q)e^a \otimes e^b,$$

then

$$\widehat{\alpha}(q^i, v^a) = \alpha_b(q)v^b, \quad T^{\mathfrak{q}}(q^i, v^a) = T_{ab}(q)v^a v^b,$$

where (q^i, v^a) are the local coordinates in $\mathcal{D}$ induced by the local coordinates (q^i) on Q and the local basis of sections of $\Gamma(\mathcal{D})$.

Theorem 5.1.10. *Let h be a pseudo-Riemannian metric and $\mathcal{D}$ be a distribution in the same conditions as in the previous theorem. If $\Gamma_{(L_h,\mathcal{D})}$ is the nonholonomic SODE associated to the problem then it acts on basic functions and on fiberwise linear functions on $\mathcal{D}$ in the following way*

$$\Gamma_{(L_h,\mathcal{D})}(f \circ \tau_{\mathcal{D}}) = \widehat{df}|_{\mathcal{D}}, \quad \Gamma_{(L_h,\mathcal{D})}(\widehat{\alpha}) = (\nabla^{nh}\alpha)^{\mathfrak{q}}, \tag{5.1.5}$$

for $f \in C^\infty(Q)$ and $\alpha \in \Gamma(\mathcal{D}^*)$, where ∇^{nh} is the nonholonomic connection and $\nabla^{nh}\alpha$ is the section of the vector bundle $\mathcal{D}^* \otimes \mathcal{D}^* \to Q$ given by

$$(\nabla^{nh}\alpha)(X, Y) = (\nabla^{nh}_X\alpha)(Y) = X(\alpha(Y)) - \alpha(\nabla^{nh}_X Y), \text{ for } X, Y \in \Gamma(\mathcal{D}).$$

Proof. Take $f \in C^\infty(Q)$ and $v \in \mathcal{D}$. Evaluating the vector field $\Gamma_{(L_h,\mathcal{D})}$ at v and then applying it to the basic function $f \circ \tau_{\mathcal{D}}$ is equivalent to apply the vector $(T_v\tau_{\mathcal{D}})(\Gamma_{(L_h,\mathcal{D})}(v))$ to the function f.

Since $\Gamma_{(L_h,\mathcal{D})}$ is a SODE on $\mathcal{D}$, its projection to the tangent bundle TQ is the identity on $\mathcal{D}$. Therefore, we obtain

$$\Gamma_{(L_h,\mathcal{D})}(v)(f \circ \tau_{\mathcal{D}}) = v(f),$$

which is exactly $\widehat{df}|_{\mathcal{D}}(v)$.

As for the second expression, let α be a section of $\mathcal{D}^*$ and take a trajectory of $\Gamma_{(L_h,\mathcal{D})}$ denoted by $c_v : I \to Q$, where the subscript means that $\dot{c}_v(0) = v$.

Let $\sharp_h(\alpha) : Q \to \mathcal{D}$ be the section of $\mathcal{D}$ given by

$$h(\sharp_h(\alpha), X) = \alpha(X), \quad \forall X \in \Gamma(\mathcal{D}).$$

Applying $\Gamma_{(L_h,\mathcal{D})}(v)$ to the fiberwise linear function $\widehat{\alpha}$ is equivalent to

$$\Gamma_{(L_h,\mathcal{D})}(v)(\widehat{\alpha}) = \ddot{c}_v(0)(\widehat{\alpha}).$$

Using the definition of derivative along a curve, the last line is equivalent to

$$\Gamma_{(L_h,\mathcal{D})}(v)(\widehat{\alpha}) = \left.\frac{d}{dt}\right|_{t=0} (\widehat{\alpha} \circ \dot{c}_v(t)),$$

and thus, using the notation we have just introduced we write

$$\Gamma_{(L_h,\mathcal{D})}(v)(\widehat{\alpha}) = \left.\frac{d}{dt}\right|_{t=0} h(\sharp_h(\alpha) \circ c_v(t), \dot{c}_v(t)),$$

Using the compatibility condition (5.1.3), this is equivalent to

$$\Gamma_{(L_h,\mathcal{D})}(v)(\widehat{\alpha}) = h((\nabla^{nh}_v \sharp_h(\alpha))(c_v(0)), v) + h(\sharp_h(\alpha)(c_v(0)), \nabla^{nh}_v \dot{c}_v(0)).$$

Since by Theorem 5.1.8, c_v is a geodesic of the connection ∇^{nh}, the last term above vanishes.

Suppose now that X is a section of $\mathcal{D}$ extending v, i.e., $X(q) = v$. With this new ingredient the last equation may be rewritten as

$$\Gamma_{(L_h,\mathcal{D})}(v)(\widehat{\alpha}) = h(\nabla^{nh}_X \sharp_h \alpha(q), X(q)).$$

By adding and subtracting the term $h(\sharp_h\alpha(q), \nabla^{nh}_X X(q))$ in the previous equation we may apply (5.1.3) and get

$$\Gamma_{(L_h,\mathcal{D})}(v)(\widehat{\alpha}) = X(q)h(\sharp_h\alpha, X) - h(\sharp_h\alpha(q), \nabla^{nh}_X X(q)),$$

and finally unwinding the definition of $\sharp_h(\alpha)$ we deduce

$$\Gamma_{(L_h,\mathcal{D})}(v)(\widehat{\alpha}) = [(\nabla^{nh}_X\alpha)(X)](q).$$

The right-hand side of the last equation is a $(0,2)$-tensor, as such, its value does not depend on the whole section and thus $\nabla^{nh}\alpha(v,v)$ is well-defined. Therefore, using the notation introduced before the theorem, it can be rewritten as $(\nabla^{nh}\alpha)^{\mathsf{q}}(v)$. □

In the following, we will see how the trajectories of Lagrangian systems with mechanical Lagrangian function are related with the nonholonomic connection. Suppose we are given a nonholonomic system $(L_{(h,V)}, \mathcal{D})$ with Lagrangian function of the form

$$L_{(h,V)}(v_q) = L_h(v_q) - V \circ \tau_Q(v_q), \tag{5.1.6}$$

where h is a pseudo-Riemannian metric on Q, $V : Q \to \mathbb{R}$ is the potential energy (with $V \in C^\infty(Q)$) and $\tau_Q : TQ \to Q$ is the canonical projection.

Given a pseudo-Riemannian metric h, define the *gradient vector field with respect to h*, $\mathrm{grad}_h V$, as

$$\mathrm{grad}_h V = \sharp_h(dV), \tag{5.1.7}$$

where $\sharp_h : \Omega^1(Q) \to \mathfrak{X}(Q)$ is the inverse isomorphism of $\flat_h$.

Then, we will prove a Lemma that will allow us to extend some results we have already proved for kinetic type Lagrangian functions to the mechanical type Lagrangian functions with little effort.

Lemma 5.1.11. *Let h be a pseudo-Riemannian metric on Q, V a potential energy on Q and $L_{(h,V)}$ a mechanical Lagrangian associated with h and V defined as in* (5.1.6)*. If $\theta_{L_{(h,V)}}$, $\omega_{L_{(h,V)}}$ and $E_{L_{(h,V)}}$ are the Poincaré-Cartan 1-form, the Poincaré-Cartan 2-form and the Lagrangian energy with respect to $L_{(h,V)}$, respectively, then we have that*

$$\theta_{L_{(h,V)}} = \theta_{L_h}, \quad \omega_{L_{(h,V)}} = \omega_{L_h}, \quad E_{L_{(h,V)}} = L_h + V \circ \tau_Q.$$

Proof. By the definition of the Poincaré-Cartan 1-form we have that

$$\theta_{L_{(h,V)}} = S^*(dL_{(h,V)}) = S^*(dL_h - dV^{\mathbf{v}}).$$

Note that we used that $V^{\mathbf{v}} = V \circ \tau_Q$. By (2.4.10) and since the pullback commutes with the differential one has that

$$(dV)^{\mathbf{v}} = d(V^{\mathbf{v}}),$$

for every $V \in C^\infty(Q)$. Then, applying (2.4.21) we deduce

$$\theta_{L_{(h,V)}} = S^*(dL_{(h,V)}) = S^*(dL_h) = \theta_{L_h}.$$

Moreover the equality for the Poincaré-Cartan 2-form follows directly from the above.

As for the Lagrangian energy just observe that

$$\Delta(V^{\mathbf{v}}) = 0, \quad \text{and} \quad \Delta(L_h) = 2L_h,$$

where Δ is the Liouville vector field on TQ. Then it is clear that

$$E_{L_{(h,V)}} = \Delta(L_{(h,V)}) - L_{(h,V)} = L_h + V \circ \tau_Q.$$

□

Now we will prove a result which is analogous to Theorem 5.1.7.

Theorem 5.1.12. *Given $L_{(h,V)}$ a mechanical Lagrangian on TQ associated with a pseudo-Riemannian metric h and a potential energy V, defined as in* (5.1.6) *and a distribution $\mathcal{D}$, the nonholonomic system $(L_{(h,V)}, \mathcal{D})$ is regular if and only if the distribution $\mathcal{D}$ is non-degenerate, in the sense of Theorem 5.1.7.*

Proof. Note that Theorem 5.1.7 is a consequence of the Lemma preceding it, which in turn is a consequence of the first relation in (A.0.5), which by Lemma 5.1.11 remains unchanged for mechanical type Lagrangian functions. Therefore, the proof of Theorem 5.1.7 extends to this case. □

Now we prove a result analogous to Theorem 5.1.8. In fact, we characterize the trajectories of $\Gamma_{(L_{(h,V)},\mathcal{D})}$ as the solution of an equation involving the nonholonomic connection ∇^{nh}.

Theorem 5.1.13. *Let $L_{(h,V)}$ be a mechanical Lagrangian on TQ associated with a pseudo-Riemannian metric h and a potential energy V defined as in* (5.1.6)*. If $\mathcal{D}$ is a non-degenerate distribution then the trajectories of $\Gamma_{(L_{(h,V)},\mathcal{D})}$ are solutions of*

$$\nabla^{nh}_{\dot{c}_v}\dot{c}_v = -P(\mathrm{grad}_h V \circ c_v), \quad \dot{c}_v(t) \in \mathcal{D}_{c_v(t)}, \tag{5.1.8}$$

where ∇^{nh} is the nonholonomic connection associated to h and $P : TQ \to \mathcal{D}$ is the orthogonal projection onto $\mathcal{D}$.

Remark 5.1.14. We remark that, in the absence of constraints (where $\mathcal{D} = TQ$), then the vector field $\Gamma_{(L_{(h,V)},TQ)}$ is exactly $\Gamma_{L_{(h,V)}}$,i.e., the Lagrangian vector field associated to the Lagrangian $L_{(h,V)}$. Moreover, a curve $c_v : I \to Q$ is a trajectory of $\Gamma_{L_{(h,V)}}$ if and only if it satisfies

$$\nabla^h_{\dot{c}_v}\dot{c}_v = -\mathrm{grad}_h V \circ c_v,$$

where ∇^h is the Levi-Civita connection associated to h (see Theorem 3.2.2).

Proof. We may follow the proof of Theorem 5.1.8 making the appropriate changes. The first change is to substitute L_h by $E_{L_{(h,V)}}$ on the geometric equation. Then, following the same arguments we will eventually get

$$\Gamma_{(L_{(h,V)},\mathcal{D})}(u)(\widehat{\flat_h(X)}) = h(\nabla^h_u X, u) - X^{\mathbf{c}}(u)(V \circ \tau_Q),$$

which by the definition of the complete lift (see (2.4.2)) is equivalent to

$$\Gamma_{(L_{(h,V)},\mathcal{D})}(u)(\widehat{\flat_h(X)}) = h(\nabla^h_u X, u) - dV(X)(\tau_Q(u)).$$

Then, evaluating the last equation over the curve $\dot{c}_v$, noting that $\Gamma_{(L_{(h,V)},\mathcal{D})}(\dot{c}_v)$ is just $\ddot{c}_v$ and using (5.1.7), we deduce

$$\ddot{c}_v\left(\widehat{\flat_h(X)}\right) = h(\nabla^h_{\dot{c}_v}X, \dot{c}_v) - h(\mathrm{grad}_h V \circ c_v, X \circ c_v).$$

As in the proof of Theorem 5.1.8, the last expression reduces to

$$h(X \circ c_v, \nabla^h_{\dot{c}_v}\dot{c}_v + \mathrm{grad}_h V \circ c_v) = 0$$

and since X is an arbitrary section of $\mathcal{D}$ we deduce that

$$P(\nabla^h_{\dot{c}_v}\dot{c}_v + \mathrm{grad}_h V \circ c_v) = 0,$$

which finishes the proof. □

5.2 Trajectories of kinetic nonholonomic systems

In this section, we will study the radial trajectories of a kinetic nonholonomic system starting from a fixed point. We will show that we may find a family of Riemannian metrics on the image submanifold of the nonholonomic exponential map, for which the radial nonholonomic trajectories are exactly the Riemannian geodesics. We will rely on special properties we find in the nonholonomic exponential map associated to a kinetic Lagrangian.

Suppose that we are given a Riemannian manifold (Q, g), a distribution $\mathcal{D}$ and consider the associated nonholonomic connection ∇^{nh}. Let c_v denote a geodesic with respect to the nonholonomic connection and initial velocity $v \in \mathcal{D}$. We will sometimes call it a nonholonomic geodesic, for simplicity. In the following lemma, we summarize some simple useful properties of nonholonomic geodesics.

Lemma 5.2.1. *Let $c_v : I \to Q$ be a nonholonomic geodesic with initial velocity $v \in \mathcal{D}_q$, i.e.*

$$c_v(t_0) = q \quad \textit{and} \quad \dot{c}_v(t_0) = v.$$

1. *We have that*
$$\|\dot{c}_v(t)\| = \|v\|, \quad \textit{for } t \in I. \tag{5.2.1}$$
2. *If $v = 0$ then $c_v(t) = q$, for every $t \in I$.*
3. *If $v \neq 0$ then a reparametrization of c_v,*
$$c_v \circ h : J \to Q, \quad s \mapsto c_v(h(s))$$
is a nonholonomic geodesic if and only if
$$h(s) = as + b, \quad \textit{with } a, b \in \mathbb{R}.$$

Proof. Using (5.1.3) and the fact that $\dot{c}_v(t) \in \mathcal{D}_{c(t)}$, for every $t \in I$, it follows that

$$\frac{d}{dt}\left(g(c_v(t))(\dot{c}_v(t), \dot{c}_v(t))\right) = 2g(c_v(t))(\nabla^{nh}_{\dot{c}_v(t)}\dot{c}_v(t), \dot{c}_v(t)) = 0.$$

Thus, we deduce that

$$\|\dot{c}_v(t)\| = \|\dot{c}_v(0)\| = \|v\|, \quad \text{for } t \in I.$$

This proves 1. 2 follows from 1.

On the other hand, if $v \neq 0$ then for the reparametrization $c_v \circ h : J \to Q$, we have that

$$\begin{array}{rcl} \nabla^{nh}_{\frac{d}{ds}(c_v\circ h)} \frac{d}{ds}(c_v \circ h) & = & \nabla^{nh}_{\frac{dh}{ds}(\frac{dc_v}{dt}\circ h)} \frac{dh}{ds}(\frac{dc_v}{dt} \circ h) \\ & = & \left(\frac{dh}{ds}\right)^2 \nabla^{nh}_{(\frac{dc_v}{dt}\circ h)}(\frac{dc_v}{dt} \circ h) + \frac{d^2h}{ds^2}(\frac{dc_v}{dt} \circ h). \end{array}$$

Now, from (5.2.1), we obtain that

$$(\frac{dc_v}{dt} \circ h)(s) \neq 0, \quad \forall s.$$

Therefore, since $\nabla^{nh}_{(\frac{dc_v}{dt}\circ h)}(\frac{dc_v}{dt} \circ h) = 0$, we conclude that

$$\nabla^{nh}_{\frac{d}{ds}(c_v\circ h)} \frac{d}{ds}(c_v \circ h) = 0 \Leftrightarrow \frac{d^2h}{ds^2} = 0 \Leftrightarrow h(s) = as + b,$$

with $a, b \in \mathbb{R}$. □

The tangent lift of the nonholonomic geodesics of a kinetic nonholonomic system $(Q, g, \mathcal{D})$ are the integral curves of the vector field $\Gamma_{(L_g,\mathcal{D})} \in \mathfrak{X}(\mathcal{D})$, which is a second order differential equation along the points of $\mathcal{D}$ considered as a vector subbundle of TQ (see Section 3.6.2).

Denote by $\phi_t^{\Gamma_{(L_g,\mathcal{D})}} : D \to D$ the flow of $\Gamma_{(L_g,\mathcal{D})}$ and for h a sufficiently small positive number, we consider the open subset of $\mathcal{D}$ given by

$$M_h^{\Gamma_{(L_g,\mathcal{D})}} = \{v \in \mathcal{D} \mid \phi_t^{\Gamma_{(L_g,\mathcal{D})}}(v) \text{ is defined for } t \in [0, h]\}.$$

Using the last part of Lemma 5.2.1 we can assume, without the loss of generality, that $h = 1$. Then, we will denote the open subset $M_1^{\Gamma_{(L_g,\mathcal{D})}}$ of $\mathcal{D}$ by $M^{\Gamma_{(L_g,\mathcal{D})}}$.

From the flow of $\Gamma_{(L_g,\mathcal{D})}$, we can define the nonholonomic exponential map

$$\begin{array}{l} \exp^{\Gamma_{(L_g,\mathcal{D})}} : M^{\Gamma_{(L_g,\mathcal{D})}} \subseteq \mathcal{D} \to Q \times Q \\ \qquad\qquad\qquad\qquad v \mapsto (\tau_Q(v), \tau_Q \circ \phi_1^{\Gamma_{(L_g,\mathcal{D})}}(v)) \end{array}$$

(see Section 4.3). We remark that if $c_v : [0, 1] \to Q$ is the nonholonomic geodesic with $\dot{c}_v(0) = v$ then

$$\exp^{\Gamma_{(L_g,\mathcal{D})}}(v) = (\tau_Q(v), c_v(1)) = (c_v(0), c_v(1)).$$

We will use in the sequel the restriction of this map to the open subset $M_q^{\Gamma_{(L_g,\mathcal{D})}} = M^{\Gamma_{(L_g,\mathcal{D})}} \cap \mathcal{D}_q$ of $\mathcal{D}_q$ with $q \in Q$ fixed, that is, we define

$$\exp_q^{nh} = pr_2 \circ \exp^{\Gamma_{(L_g,\mathcal{D})}}\Big|_{M_q^{\Gamma_{(L_g,\mathcal{D})}}} : M_q^{\Gamma_{(L_g,\mathcal{D})}} \subset \mathcal{D}_q \longrightarrow Q$$

So, if $v_q \in \mathcal{D}_q$ and $c_{v_q} : [0,1] \to Q$ is the nonholonomic geodesic with initial velocity v_q then

$$\exp_q^{nh}(v_q) = c_{v_q}(1).$$

The reader can compare this definition of $\exp_q^{nh}$ with the definition of the Riemannian exponential at q (see equation (2.1.7) in Section 2.1.2).

In fact, the nonholonomic exponential map conserves many of the properties we may find in Riemannian exponential maps. Thus,

$$\exp_q^{nh}(0_q) = q. \tag{5.2.2}$$

The second result we are going to prove is the *rescaling lemma*.

Lemma 5.2.2. *Let $c_{v_q} : [0,1] \to Q$ denote the nonholonomic geodesic with initial velocity $v_q \in M_q^{\Gamma_{(L_g,\mathcal{D})}}$. Then we have that*

$$c_{v_q}(t) = \exp_q^{nh}(tv_q), \quad \textit{for } t \in [0,1].$$

Proof. If $t = 0$ the result is obvious.

Suppose that $t \neq 0$. Then, we can consider the curve

$$s \to q(s) = c_{v_q}(ts).$$

Using the last part of Lemma 5.2.1, we deduce that the previous curve is a nonholonomic geodesic. Moreover, its initial velocity is tv_q.

Thus,

$$c_{tv_q}(s) = c_{v_q}(ts).$$

As a consequence of this and by the definition of exponential map, it follows that

$$\exp_q^{nh}(tv_q) = c_{tv_q}(1) = c_{v_q}(t).$$

□

Using the previous results, we may give an alternative simpler proof of the fact that the tangent map of the exponential map at 0_q is an isomorphism onto its image (for the general case see Section 4.3).

Proposition 5.2.3. *If $q \in Q$ then, under the canonical linear identification between $T_{0_q} M_q^{\Gamma_{(L_g,\mathcal{D})}}$ and $\mathcal{D}_q$, the linear map*

$$T_{0_q} exp_q^{nh} : T_{0_q} M_q^{\Gamma_{(L_g,\mathcal{D})}} \simeq D_q \to T_q Q$$

is just the canonical inclusion of D_q in T_qQ. So, there exists a star-shaped open subset $\mathcal{U}_0$ of $\mathcal{D}_q$ around $0_q \in \mathcal{U}_0$ such that the nonholonomic exponential map $\exp_q^{nh} : \mathcal{U}_0 \to Q$ is an embedding.

Proof. Observe that for $v_q \in D_q$,

$$\begin{aligned}(T_{0_q}\exp_q^{nh})(v_q) &= \left.\frac{d}{dt}\right|_{t=0} \exp_q^{nh}(tv_q) \\ &= \left.\frac{d}{dt}\right|_{t=0} c_{v_q}(t) = v_q\end{aligned}$$

using Lemma 5.2.2. Therefore, $T_{0_q}\exp_q^{nh} : \mathcal{D}_q \to T_qQ$ is just the canonical inclusion.

Thus, there exists a star-shaped open subset $\mathcal{U}_0$ of D_{q_0} around $0_{q_0} \in \mathcal{U}_0$, such that $\exp_q^{nh} : \mathcal{U}_0 \to Q$ is a diffeomorphism over its image. □

5.2.1 The geodesic character of radial trajectories

The goal of this section is to prove the following theorem:

Theorem 5.2.4. *Let $(Q, g, \mathcal{D})$ be a kinetic nonholonomic system and q a fixed point in Q. Then:*

i) There exists a submanifold $\mathcal{M}_q^{nh}$ of Q, with $q \in \mathcal{M}_q^{nh}$, and a diffeomorphism $exp_q^{nh} : \mathcal{U}_0 \subseteq D_q \to \mathcal{M}_q^{nh} \subseteq Q$, where $\mathcal{U}_0$ is a star-shaped open subset of $\mathcal{D}_q$ around $0_q \in \mathcal{U}_0$ and $exp_q^{nh}(0_q) = q$. exp_q^{nh} is the nonholonomic exponential map at q. Moreover, we have that:

(a) Under the canonical linear identification between $\mathcal{D}_q$ and $T_{0_q}\mathcal{U}_0$, the linear monomorphism

$$T_{0_q} exp_q^{nh} : T_{0_q}\mathcal{U}_0 \simeq D_q \to T_qQ$$

is just the canonical inclusion of D_q in T_qQ.

(b) For every $v_q \in \mathcal{U}_0$,

$$exp_q^{nh}(tv_q) = c_{v_q}(t), \quad t \in [0,1],$$

with $c_{v_q} : [0,1] \to \mathcal{M}_q^{nh} \subseteq Q$ the (unique) nonholonomic trajectory satisfying $c_{v_q}(0) = q, \dot{c}_{v_q}(0) = v_q$.

ii) All the radial kinetic nonholonomic trajectories from the fixed point $q \in Q$ are a homothetic reparametrization of nonholonomic trajectories as in (b). In addition, they are minimizing geodesics for a Riemannian metric g_q^{nh} on $\mathcal{M}_q^{nh}$ if and if only if the Riemannian metric $\mathcal{G}_0 = (exp_q^{nh})^*(g_q^{nh})$ on $\mathcal{U}_0$ satisfies the Gauss condition, that is,

$$\mathcal{G}_0(v_q)(v_q, w_q) = \mathcal{G}_0(0_q)(v_q, w_q), \quad \text{for } v_q \in \mathcal{U}_0 \text{ and } w_q \in D_q.$$

iii) Such Riemannian metrics on $\mathcal{M}_q^{nh}$ always exist and if g_q^{nh} is one of them then the Riemannian exponential associated with g_q^{nh} at q is just exp_q^{nh}.

In order to prove it, we will need to make a small digression through Riemannian structures on vector spaces. In particular, we will present a motivation for using the terminology "Gauss condition".

Let g be a Riemannian metric on a manifold Q and q a point in Q. Denote by $\exp_q^g : T_qQ \to Q$ the Riemannian exponential at the point q. As we know (see Section 2.1.3 and also [Car92; O'N83]),

$$\exp_q^g(v_q) = c_{v_q}(1), \tag{5.2.3}$$

for $v_q \in T_qQ$, where $c_{v_q} : [0,1] \to Q$ is the unique geodesic in Q with initial velocity v_q, that is, $c_{v_q}(0) = q$ and $\dot{c}_{v_q}(0) = v_q$. Note that $\exp_q^g(0_q) = q$. Moreover, there exist open subsets $\mathcal{U} \subseteq T_qQ$ and $U \subseteq Q$, with $\mathcal{U}$ starshaped about $0_q \in \mathcal{U}$ and $q \in U$, such that

$$\exp_q^g : \mathcal{U} \to U$$

is a diffeomorphism and

$$\exp_q^g(tv_q) = c_{v_q}(t), \quad T_{0_q}\exp_q^g = id_{T_qQ} : T_{0_q}\mathcal{U} \simeq T_qQ \to T_qQ. \tag{5.2.4}$$

In fact, the curve $t \in [0,1] \to c_{v_q}(t) \in Q$ is a minimizing geodesic from q to $c_{v_q}(1)$. Then the Gauss' lemma (see equation (2.1.8) on Section 2.1) implies that

$$g(\exp_q^g(v_q))((T_{v_q}\exp_q^g)(v_q)_{v_q}^{\mathrm{v}}, (T_{v_q}\exp_q^g)(w_q)_{v_q}^{\mathrm{v}}) = g(q)(v_q, w_q), \tag{5.2.5}$$

for $v_q \in \mathcal{U}$ and $w_q \in T_qQ$, where $(v_q)^{\mathsf{v}}_{v_q}, (w_q)^{\mathsf{v}}_{v_q} \in T_{v_q}(T_qQ)$ are the vertical lifts to TQ at v_q of v_q and w_q, respectively, given by (2.4.5). So, under the linear identification ${}^{\mathsf{v}}_{v_q} : T_qQ \to T_{v_q}(T_qQ)$ between T_qQ and $T_{v_q}(T_qQ)$, equation (5.2.5) gives

$$g(\exp^g_q(v_q))(T_{v_q}\exp^g_q(v_q), T_{v_q}\exp^g_q(w_q)) = g(q)(v_q, w_q). \tag{5.2.6}$$

Thus, if we consider the Riemannian metric $\mathcal{G}_0$ on $\mathcal{U}$ defined by

$$\mathcal{G}_0 = (\exp^g_q)^*(g)$$

then we deduce that

$$\mathcal{G}_0(v_q)(v_q, w_q) = \mathcal{G}_0(0_q)(v_q, w_q).$$

The previous fact motivates the definition below. Given a vector space E equipped with a Riemannian metric $\mathcal{G}$, if $u \in E$ then, as above, we will identify the tangent space T_uE with E.

Definition 5.2.5. We say that the Riemannian manifold $(E, \mathcal{G})$ satisfies the *Gauss condition* if

$$\mathcal{G}(v)(v, w) = \mathcal{G}(0)(v, w), \qquad \forall v, w \in E\ .$$

The previous definition may also be applied to an open subset $\mathcal{U}$ of E which contains the zero vector of E, that is, a Riemannian metric $\mathcal{G}$ on $\mathcal{U}$ satisfies the *Gauss condition* if

$$\mathcal{G}(v)(v, w) = \mathcal{G}(0)(v, w), \qquad \forall v \in \mathcal{U} \text{ and } w \in E. \tag{5.2.7}$$

Let $\mathcal{G}$ be a Riemannian metric on E, $\{e_i\}_{i=1,\dots,n}$ a basis of E, $(x^1, \dots, x^n)$ the global coordinates on E induced by the basis $\{e_i\}_{i=1,\dots,n}$ and $\mathcal{U}$ an open subset of E, with $0 \in \mathcal{U}$. Denote by $\mathcal{G}_{ij}$, with $i, j \in \{1, \dots, n\}$, the components of $\mathcal{G}$ with respect to the global coordinates $(x^1, \dots, x^n)$, that is,

$$\mathcal{G}_{ij} : E \to \mathbb{R}, \;\; v \in E \to \mathcal{G}_{ij}(v) = \mathcal{G}(v)\left(\frac{\partial}{\partial x^i}_{|v}, \frac{\partial}{\partial x^j}_{|v}\right) \in \mathbb{R}.$$

Then, from (5.2.7), it follows that $\mathcal{G}$ satisfies the Gauss condition on $\mathcal{U}$ if and only if

$$x^i(v)(\mathcal{G}_{ij}(v) - \mathcal{G}_{ij}(0)) = 0, \quad \text{for } j \in \{1, \dots, n\} \text{ and } v \in \mathcal{U}.$$

Remark 5.2.6. If $\mathcal{G}$ is the flat metric on E induced by a scalar product on E then it is clear that $\mathcal{G}$ satisfies the Gauss condition (in fact, $\mathcal{G}_{ij}(v) = \mathcal{G}_{ij}(0)$, for every $v \in E$). More generally, let $\overline{\mathcal{G}}$ be an arbitrary Riemannian metric on E and $\exp_0^{\overline{\mathcal{G}}} : T_0E \to E$ the Riemannian exponential at 0. As we know, there exist open subsets $\mathcal{U} \subseteq T_0E$ and $U \subseteq E$, such that $0 \in \mathcal{U} \cap U$ and

$$\exp_0^{\overline{\mathcal{G}}} : \mathcal{U} \to U$$

is a diffeomorphism. Then, proceeding as in the previous discussion to Definition 5.2.5, we deduce that the Riemannian metric on $\mathcal{U}$ given by

$$\mathcal{G} = \left((\exp_0^{\overline{\mathcal{G}}})_{|\mathcal{U}}\right)^* (\overline{\mathcal{G}}_{|U})$$

satisfies the Gauss condition.

Now, let $\mathcal{G}$ be a Riemannian metric on a vector space E and $\exp_0^{\mathcal{G}} : \mathcal{U} \subseteq T_0E \simeq E \to U \subseteq E$ the Riemannian exponential map at the zero vector $0 \in E$ (where we have used the canonical identification between T_0E and E). Then, we denote by $r_0^{\mathcal{G}} : U \subseteq E \to \mathbb{R}$ the standard Riemannian radial function at 0 for the Riemannian manifold $(E, \mathcal{G})$, that is, (see [Car92; O'N83]),

$$r_0^{\mathcal{G}}(v) = \|(\exp_0^{\mathcal{G}})^{-1}(v)\|_{\mathcal{G}(0)}, \quad \text{for } v \in U.$$

Moreover, using that the Riemannian manifold is a vector space, we can also define the *radial distance function* $r^{\mathcal{G}} : E \to \mathbb{R}$ given by

$$r^{\mathcal{G}}(v) = \|v\|_{\mathcal{G}(v)} = \sqrt{\mathcal{G}(v)(v,v)}\ .$$

In general, we have that $r_0^{\mathcal{G}} \neq (r^{\mathcal{G}})_{|U}$. However, if $(E, \mathcal{G})$ satisfies the Gauss condition in $\mathcal{U}$, we will see that $\exp_0^{\mathcal{G}} : \mathcal{U} \subseteq E \to E$ is the canonical inclusion of $\mathcal{U}$ in E (see Theorem 5.2.9 below) and, thus, $r_0^{\mathcal{G}} = (r^{\mathcal{G}})_{|U}$.

First, we will prove the following result:

Lemma 5.2.7. *The radial distance function* $r^{\mathcal{G}} : E \to \mathbb{R}$ *is smooth on* $E \backslash \{0\}$. *Moreover, if* $(E, \mathcal{G})$ *satisfies the Gauss condition, then the gradient vector field* $\operatorname{grad}_{\mathcal{G}} r^{\mathcal{G}}\Big|_{E\setminus\{0\}}$ *of* $r^{\mathcal{G}}$ *on* $E \setminus \{0\}$ *is given by*

$$\operatorname{grad}_{\mathcal{G}} r^{\mathcal{G}}(v) = \frac{v}{\|v\|_{\mathcal{G}(0)}}, \quad \textit{for } v \in E \setminus \{0\}$$

and, in addition, it is unitary relative to the metric $\mathcal{G}$, *that is,*

$$\|\operatorname{grad}_{\mathcal{G}} r^{\mathcal{G}}(v)\|_{\mathcal{G}(v)} = 1, \quad \textit{for } v \in E \setminus \{0\}.$$

Proof. The first part of the lemma is obvious. On the other hand, using the definition of the gradient vector field and the Gauss condition, we have that, for $v \in E \setminus \{0\}$ and $u \in E$,

$$\begin{aligned}
\mathcal{G}(v)(\mathrm{grad}_{\mathcal{G}} r^{\mathcal{G}}(v), u) &= \langle dr^{\mathcal{G}}(v), u\rangle = \langle dr^{\mathcal{G}}(v), \frac{d}{dt}\Big|_{t=0}(v+tu)\rangle \\
&= \frac{d}{dt}\Big|_{t=0} r^{\mathcal{G}}(v+tu) = \frac{d}{dt}\Big|_{t=0}\sqrt{(\mathcal{G}(0)(v+tu, v+tu)} \\
&= \frac{d}{dt}\Big|_{t=0}\sqrt{\|v\|^2_{\mathcal{G}(0)} + 2t\mathcal{G}(0)(u,v) + t^2\|u\|^2_{\mathcal{G}(0)}} \\
&= \frac{\mathcal{G}(0)(v,u)}{\|v\|_{\mathcal{G}(0)}} \\
&= \mathcal{G}(v)\left(\frac{v}{\|v\|_{\mathcal{G}(0)}}, u\right)
\end{aligned}$$

This implies that,

$$\mathrm{grad}_{\mathcal{G}} r^{\mathcal{G}}(v) = \frac{v}{\|v\|_{\mathcal{G}(0)}}.$$

Now, using again the Gauss condition, it follows that

$$\mathcal{G}(v)\left(\frac{v}{\|v\|_{\mathcal{G}(0)}}, \frac{v}{\|v\|_{\mathcal{G}(0)}}\right) = \frac{\mathcal{G}(0)(v,v)}{(\|v\|_{\mathcal{G}(0)})^2} = 1$$

which concludes the proof of the result. □

Another fact which is relevant for our purposes is the following.

Lemma 5.2.8. *If $(E, \mathcal{G})$ satisfies the Gauss condition then the integral curves of the vector field $U = \mathrm{grad}_{\mathcal{G}} r^{\mathcal{G}}$ are geodesic.*

Proof. This follows using that U is unitary and the gradient vector field of a real C^∞-function on $E \setminus \{0\}$.

In fact, if ∇ is the Levi-Civita connection of $\mathcal{G}$ and $X \in \mathfrak{X}(E \setminus \{0\})$ then, since ∇ is metric and torsion free, we have that

$$\mathcal{G}(\nabla_U U, X) = U(\mathcal{G}(U,X)) - \mathcal{G}(U, \nabla_U X) = U(X(r^{\mathcal{G}})) - \mathcal{G}(U, [U,X]) + \mathcal{G}(U, \nabla_X U).$$

On the other hand, using again that ∇ is metric and that U is unitary, we deduce that

$$0 = X(\mathcal{G}(U,U)) = 2\mathcal{G}(U, \nabla_X U).$$

Thus, we obtain that

$$\begin{aligned}\mathcal{G}(\nabla_U U, X) &= U(X(r^{\mathcal{G}})) - [U,X](r^{\mathcal{G}}) \\ &= U(X(r^{\mathcal{G}})) - U(X(r^{\mathcal{G}})) + X(U(r^{\mathcal{G}})) = 0,\end{aligned} \tag{5.2.8}$$

where the last equality follows using that

$$U(r^{\mathcal{G}}) = dr^{\mathcal{G}}(U) = \mathcal{G}(U,U) = 1.$$

Finally, relation (5.2.8) implies that $\nabla_U U = 0$ and, therefore, the integral curves of U are geodesic. □

Now, let us characterize the Riemannian metrics on vector spaces satisfying the Gauss condition on $\mathcal{U}$ as those for which the exponential map is just the inclusion or, equivalently, those for which the geodesics through zero are straight lines.

Theorem 5.2.9. *If $\mathcal{G}$ is a Riemannian metric on a vector space E and*

$$\exp_0^{\mathcal{G}} : \mathcal{U} \subseteq T_0E \simeq E \to E,$$

is the exponential map at the zero vector then the following conditions are equivalent:

i) The map $\exp_0^{\mathcal{G}} : \mathcal{U} \subseteq E \to E$ is the canonical inclusion of $\mathcal{U}$ in E.

ii) For each $u \in \mathcal{U}$, the line

$$t \in [0,1] \to tu \in \mathcal{U}$$

starting at the zero vector is a minimizing geodesic for $(E,\mathcal{G})$ with initial velocity u.

iii) $(E,\mathcal{G})$ satisfies the Gauss condition in $\mathcal{U}$.

Proof. As we know,

$$\exp_0^{\mathcal{G}}(tu) = c_u(t), \quad \text{for } t \in [0,1]$$

where $c_u : [0,1] \to E$ is the minimizing geodesic with initial velocity $u \in \mathcal{U}$. [$i) \Leftrightarrow ii)$] If $\exp_0^{\mathcal{G}} : \mathcal{U} \to E$ is the inclusion of $\mathcal{U}$ in E then $c_u(t) = tu$ which proves $i) \Rightarrow ii)$.

Conversely, if $ii)$ holds then it is clear that

$$\exp_0^{\mathcal{G}}(tu) = tu, \quad \text{for } u \in \mathcal{U},$$

and, thus, $\exp_0^{\mathcal{G}} : \mathcal{U} \subseteq E \to E$ is the canonical inclusion of $\mathcal{U}$ in E.

$[i) \Rightarrow iii)]$ If $u \in \mathcal{U}$ and $v \in E$ then, using the Gauss' Lemma and $i)$, we have that

$$\begin{aligned} \mathcal{G}(u)(u,v) &= \mathcal{G}(\exp_0^{\mathcal{G}}(u))(T_u\exp_0^{\mathcal{G}}(u), T_u\exp_0^{\mathcal{G}}(v)) \\ &= \mathcal{G}(0)(u,v). \end{aligned}$$

So, $(E, \mathcal{G})$ satisfies the Gauss condition in $\mathcal{U}$.

$[iii) \Rightarrow i)]$ Let $r^{\mathcal{G}} : E \to \mathbb{R}$ be the radial distance function and $U = \text{grad}_{\mathcal{G}} r^{\mathcal{G}}$. From Lemma 5.2.7, we have that $U(u) = \frac{u}{\|u\|_{\mathcal{G}(0)}}$, for $u \in E \setminus \{0\}$. This implies that the line $l_{\frac{u}{\|u\|_{\mathcal{G}(0)}}} : (0, \infty) \to E \setminus \{0\}$ given by

$$l_{\frac{u}{\|u\|_{\mathcal{G}(0)}}}(t) = t\frac{u}{\|u\|_{\mathcal{G}(0)}}$$

is an integral curve of U. Then, using Lemma 5.2.8, $l_{\frac{u}{\|u\|_{\mathcal{G}(0)}}}$ is a geodesic. Thus, the homothetic reparametrization of $l_{\frac{u}{\|u\|_{\mathcal{G}(0)}}}$ defined by

$$t \to l_{\frac{u}{\|u\|_{\mathcal{G}(0)}}}(\|u\|_{\mathcal{G}(0)}t) = tu$$

is also a geodesic. It is clear that the curve defined above is continuously extendible to $t = 0$. So, from Lemma 8 of Chapter 5 in [O'N83], it is also extendible as a geodesic. Moreover, its initial velocity is u. Thus

$$\exp_0^{\mathcal{G}}(tu) = tu, \quad \text{for } t \in [0, 1].$$

This proves $i)$. □

We will prove each one of the items in Theorem 5.2.4

Proof of Theorem 5.2.4. $i)$ Take in Proposition 5.2.3

$$\mathcal{M}_q^{nh} = \exp_q^{nh}(\mathcal{U}_0).$$

Then, using (5.2.2), Lemma 5.2.2 and Proposition 5.2.3, we deduce the first part of the theorem.

ii) Let $t \mapsto c_{u_q}(t)$ be a radial kinetic nonholonomic trajectory departing from q, that is, c_{u_q} is a nonholonomic trajectory and

$$c_{u_q}(0) = q, \ \ \dot{c}_{u_q}(0) = u_q \in \mathcal{D}_q.$$

Then, using that $\mathcal{U}_0$ is an open subset of $\mathcal{D}_q$ and $0_q \in \mathcal{U}_0$, there exists $v_q \in \mathcal{U}_0$ and a real number $\lambda > 0$ such that $v_q = \dfrac{u_q}{\lambda}$. Also, by item *i*) of this Theorem, the radial kinetic nonholonomic trajectory

$$c_{v_q} : [0,1] \to Q, \ \ c_{v_q}(t) = \exp_q^{nh}(tv_q)$$

is contained in $\mathcal{M}_q^{nh}$. As

$$\dot{c}_{v_q}(0) = v_q = \frac{u_q}{\lambda} = \frac{\dot{c}_{u_q}(0)}{\lambda},$$

from Lemma 5.2.1, c_{u_q} is just the homothetic reparametrization of c_{v_q} given by

$$c_{u_q}(t) = c_{v_q}(\lambda t), \quad \text{for } 0 < t < \frac{1}{\lambda}.$$

This proves the first part of *ii*) in the theorem.

Now, suppose that g_q^{nh} is a Riemannian metric on $\mathcal{M}_q^{nh}$ and that $\mathcal{G}_0$ is the Riemannian metric on $\mathcal{U}_0$ given by $\mathcal{G}_0 = (\exp_q^{nh})^*(g_q^{nh})$. From Theorem 5.2.9, it follows that the lines through 0_q, which are of the form

$$t \in [0,1] \mapsto tv_q, \quad \text{for } v_q \in \mathcal{U}_0,$$

are minimizing geodesics in the Riemannian manifold $(\mathcal{U}_0, \mathcal{G}_0)$ if and only if $\mathcal{G}_0$ satisfies the Gauss condition.

On the other hand, from the definition of $\mathcal{G}_0$, we have that $\exp_q^{nh} : (\mathcal{U}_0, \mathcal{G}_0) \to (\mathcal{M}_q^{nh}, g_q^{nh})$ is an isometry and, from Lemma 5.2.2, the image by $\exp_q^{nh}$ of the lines through 0_q are just the radial kinetic nonholonomic trajectories departing from q. Thus, we conclude that these trajectories are minimizing geodesics in $(\mathcal{M}_q^{nh}, g_q^{nh})$ if and only if $\mathcal{G}_0$ satisfies the Gauss condition in $\mathcal{U}_0$.

iii) As we know, there exist Riemannian metrics on $\mathcal{U}_0$ which satisfy the Gauss condition (see Remark 5.2.6). So, if $\mathcal{G}_0$ is one of them and we define the Riemannian metric g_q^{nh} on $\mathcal{M}_q^{nh}$ to be given by

$$g_q^{nh} = \left((\exp_q^{nh})^{-1}\right)^* (\mathcal{G}_0),$$

it is clear that, using item ii) in this theorem, the radial kinetic nonholonomic trajectories departing from q are minimizing geodesics in the Riemannian manifold $(\mathcal{M}_q^{nh}, g_q^{nh})$. This proves the first part of item iii).

Now, under the canonical linear identification between $T_q\mathcal{M}_q^{nh}$ and $\mathcal{D}_q$ induced by the linear isomorphism

$$T_{0_q}\exp_q^{nh} : \mathcal{D}_q \to T_q\mathcal{M}_q^{nh},$$

let $\exp_q^{g_q^{nh}} : T_q\mathcal{M}_q^{nh} \simeq \mathcal{D}_q \to \mathcal{M}_q^{nh}$ be the Riemannian exponential associated with g_q^{nh} at the point q. We may assume, without loss of generality, that

$$\exp_q^{g_q^{nh}} : \mathcal{U}_0 \subseteq \mathcal{D}_q \to \mathcal{M}_q^{nh}$$

is a diffeomorphism. Moreover, if $v_q \in \mathcal{U}_0$ then, since the radial kinetic nonholonomic trajectory $c_{v_q} : [0,1] \to \mathcal{M}_q^{nh}$ is a minimizing geodesic in the Riemannian manifold $(\mathcal{M}_q^{nh}, g_q^{nh})$ with initial velocity v_q, we deduce that

$$\exp_q^{g_q^{nh}}(tv_q) = c_{v_q}(t) = \exp_q^{nh}(tv_q), \quad \text{for } t \in [0,1].$$

So, $\exp_q^{g_q^{nh}} = \exp_q^{nh}$. □

Remark 5.2.10. Suppose that the distribution $\mathcal{D} \subseteq TQ$ is integrable and that M_q is the leaf of $\mathcal{D}$ that passes through $q \in Q$. Then, it is well-known that the restriction to M_q of the nonholonomic vector field $\Gamma_{(L_g,\mathcal{D})}$ is tangent to M_q.

Moreover, by Proposition 6.5 in [Lew98], we have that

$$(\nabla_X^{nh} Y)_{|M_q} = \nabla_{X_{|M_q}}^{i_{M_q}^* g} Y_{|M_q}, \quad \forall X, Y \in \Gamma(\mathcal{D}),$$

where $i_{M_q} : M_q \hookrightarrow Q$ is the canonical inclusion and $\nabla^{i_{M_q}^* g}$ is the Levi-Civita connection associated with the Riemannian metric $i_{M_q}^* g$ on M_q.

Therefore, a nonholonomic trajectory $c_{v_q}^{nh} : I \to Q$ satisfying

$$c_{v_q}^{nh}(0) = q, \quad \dot{c}_{v_q}^{nh}(0) = v_q \in \mathcal{D}_q = T_q M_q$$

is contained in M_q, that is, $c_{v_q}^{nh}(I) \subseteq M_q$ and it is a geodesic with respect to the Riemannian metric $i_{M_q}^* g$. Hence, the nonholonomic exponential map at q satisfies

$$\exp_q^{nh} = \exp_q^{i_{M_q}^* g} : \mathcal{U}_0 \subseteq \mathcal{D}_q = T_q M_q \to U \subseteq M_q,$$

where $\mathcal{U}_0$ is a starshaped open subset of $\mathcal{D}_q$ about $0_q \in \mathcal{U}_0$, U is an open subset in M_q, with $q \in U$, and $\exp_q^{i_M^* g}$ is the Riemannian exponential map at q associated with the Riemannian metric $i_{M_q}^* g$. Therefore, we have that $\mathcal{M}_q^{nh}$ is an open subset of M_q and $i_{M_q}^* g$ belongs to the family of Riemannian metrics g_q^{nh} defined in *ii*) of Theorem 5.2.4 since it satisfies Gauss Lemma and, thus,

$$\mathcal{G}_0 = (\exp_q^{nh})^* i_{M_q}^* g = (\exp_q^{i_{M_q}^* g})^* i_{M_q}^* g$$

must satisfy the Gauss condition.

We illustrate the statement of Theorem 5.2.4 with some examples. First of all, we will see that if $i_q : \mathcal{M}_q^{nh} \to Q$ is the canonical inclusion then the Riemannian metric

$$\mathcal{G}_0 = (\exp_q^{nh})^* (i_q^* g)$$

on $\mathcal{U}_0 \subseteq \mathcal{D}_q$ does not satisfy, in general, Gauss condition.

Example 5.2.11. Consider the nonholonomic particle in $Q = \mathbb{R}^3$, that is, g is the standard flat Riemannian metric on $\mathbb{R}^3$ and $\mathcal{D}$ is the constraint distribution determined by

$$\mathcal{D} = \{(x, y, z, \dot{x}, \dot{y}, \dot{z}) \in TQ \mid \dot{z} = y\dot{x}\}.$$

Here, (x, y, z) and $(x, y, z, \dot{x}, \dot{y}, \dot{z})$ are the standard coordinates on $\mathbb{R}^3$ and $T\mathbb{R}^3$, respectively.

For simplicity we will fix $q = 0 = (0, 0, 0)$. It is clear that

$$\mathcal{D}_0 =< U = \left.\frac{\partial}{\partial x}\right|_0, V = \left.\frac{\partial}{\partial y}\right|_0 > .$$

Denote by (u, v) the linear coordinates on $\mathcal{D}_0$ induced by the previous basis.

The nonholonomic exponential map $\exp_0^{nh} : \mathcal{D}_0 \to Q$ is known to be given by (see (4.3.2) and (4.3.3))

$$\exp_0^{nh}(u, v) = \left(\frac{u}{v} \operatorname{arcsinh}(v), v, \frac{u}{v}(\sqrt{v^2 + 1} - 1)\right)$$

if $v \neq 0$ and

$$\exp_0^{nh}(u, 0) = (u, 0, 0), \quad \text{if } v = 0.$$

The tangent map of $\exp_0^{nh}$ at $\bar{U} = U + V \in \mathcal{D}_0$ is represented in coordinates by the matrix

$$T_{(1,1)}\exp_0^{nh} = \begin{pmatrix} \text{arcsinh}(1) & \frac{\sqrt{2}}{2} - \text{arcsinh}(1) \\ 0 & 1 \\ \sqrt{2}-1 & 1-\frac{\sqrt{2}}{2} \end{pmatrix},$$

where $\bar{U}$ is represented in coordinates by $(1,1)$. Therefore, we have that

$$T_{(1,1)}\exp_0^{nh}(\bar{U}) = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 1 \\ \frac{\sqrt{2}}{2} \end{pmatrix}$$
$$T_{(1,1)}\exp_0^{nh}(U) = \begin{pmatrix} \text{arcsinh}(1) \\ 0 \\ \sqrt{2}-1 \end{pmatrix}$$

and so

$$g\left(T_{(1,1)}\exp_0^{nh}(\bar{U}), T_{(1,1)}\exp_0^{nh}(U)\right) = \frac{\sqrt{2}}{2}(\text{arcsinh}(1) + \sqrt{2} - 1) \approx 0.9161$$

But, $g(\bar{U}, U) = 1$, thus, the Riemannian metric $\mathcal{G}_0 = (\exp_0^{nh})^*(i_0^* g)$ on $\mathcal{D}_0$ does not satisfy Gauss condition. $\triangle$

Next, for a fixed point $q \in Q$, we will give examples of Riemannian metrics $\mathcal{G}_0$ satisfying Gauss condition on $\mathcal{D}_q$ and we will obtain the corresponding metrics

$$g_q^{nh} = \left((\exp_q^{nh})^{-1}\right)^* (\mathcal{G}_0)$$

on $\mathcal{M}_q^{nh}$.

Example 5.2.12. Consider again the nonholonomic particle in $Q = \mathbb{R}^3$ and fix the point $q = 0 \in \mathbb{R}^3$.

Let $\mathcal{G}_0$ be the standard flat metric in $\mathcal{D}_0 \simeq \mathbb{R}^2$ so that

$$((\mathcal{G}_0)_{ij}(u,v)) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is clear that $\mathcal{G}_0$ satisfies the Gauss condition in $\mathcal{D}_0$.

Denote by (x, y) the coordinates on $\mathcal{M}_0^{nh}$ induced by the coordinates (u, v) on $\mathcal{D}_0$ and by the nonholonomic exponential map $\exp_0^{nh}$. Since

$$\mathcal{G}_0 = du \otimes du + dv \otimes dv,$$

and

$$(\exp_0^{nh})^{-1}(x, y) = \left(\frac{xy}{\operatorname{arcsinh}(y)}, y\right)$$

we have that

$$g_0^{nh} = Edx \otimes dx + Fdx \otimes dy + Fdy \otimes dx + Gdy \otimes dy,$$

with

$$E = \frac{y^2}{\operatorname{arcsinh}^2(y)}, \quad F = \frac{xy(\operatorname{arcsinh}(y)\sqrt{y^2+1} - y)}{\sqrt{y^2+1}\operatorname{arcsinh}^3(y)}$$
$$G = \frac{-2\operatorname{arcsinh}(y)\sqrt{y^2+1}x^2 y + \operatorname{arcsinh}^2(y)(y^2+1)x^2 + x^2y^2}{\operatorname{arcsinh}^4(y)(y^2+1)} + 1.$$

△

Example 5.2.13. For the same nonholonomic system, one could choose other Riemannian metric satisfying the Gauss condition. Consider the Riemannian metric on $\mathcal{D}_0$ given by

$$\mathcal{G}_0 = (1 - v^2)du \otimes du + uvdu \otimes dv + uvdv \otimes du + (1 - u^2)dv \otimes dv.$$

Note that, this tensor is degenerate on the unitary circle around the origin of $\mathcal{D}_0$ (where the radius is measured with respect to the euclidean metric). To overcome this technicality, we will restrict ourselves to the open ball with unit radius on which the metric $\mathcal{G}_0$ is non-degenerate. This example illustrates that Theorem 5.2.4 could in principle be extended to convex subsets of vector spaces.

The Chrystoffel symbols with respect to this metric are

$$\Gamma^u_{uu} = \frac{2uv^2}{u^2+v^2-1}, \quad \Gamma^u_{uv} = -\frac{(2u^2-1)v}{u^2+v^2-1}, \quad \Gamma^u_{vv} = \frac{2(u^3-u)}{u^2+v^2-1},$$
$$\Gamma^v_{uu} = \frac{2(v^3-v)}{u^2+v^2-1}, \quad \Gamma^v_{uv} = -\frac{(2uv^2-u)}{u^2+v^2-1}, \quad \Gamma^v_{vv} = \frac{2u^2v}{u^2+v^2-1}.$$

Consider now the lines $c_{(u_0,v_0)} : [0,1] \to \mathcal{D}_0$ contained in the unitary open ball in $\mathcal{D}_0$ departing from 0, with the coordinate expression

$$c_{(u_0,v_0)}(t) = (u_0 t, v_0 t).$$

It is easy to check that the curves $c_{(u_0,v_0)}$ satisfy the geodesic equations and therefore the lines through the origin are in fact geodesics. At the same time, it is clear that the exponential map is the identity and we can check that the Riemannian metric $\mathcal{G}_0$ satisfies the Gauss condition: let $X, Y \in \mathcal{D}_0$ with the local expression $(X^u, X^{\mathbf{v}})$ and $(Y^u, Y^{\mathbf{v}})$, respectively.

Then

$$\begin{aligned}\mathcal{G}_0(X)(X,Y) =& (1-(X^{\mathbf{v}})^2)X^uY^u + X^uX^{\mathbf{v}}(X^uY^{\mathbf{v}} + X^{\mathbf{v}}Y^u) \\ &+ (1-(X^u)^2)X^{\mathbf{v}}Y^{\mathbf{v}} \\ =& X^uY^u + X^{\mathbf{v}}Y^{\mathbf{v}} = \mathcal{G}_0(0)(X,Y).\end{aligned}$$

△

Example 5.2.14. Consider the example of the vertical rolling disk with $Q = \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$, parametrized by the coordinates (x, y, θ, φ). This system is described by the Lagrangian function $L : TQ \to \mathbb{R}$ given by

$$L(x, y, \varphi, \theta, \dot{x}, \dot{y}, \dot{\varphi}, \dot{\theta}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2,$$

and subjected to the constraint distribution $\mathcal{D} \subseteq TQ$ determined by the equations

$$\dot{x} = R\cos\varphi\ \dot{\theta}, \quad \dot{y} = R\sin\varphi\ \dot{\theta},$$

where R is the radius of the disk, m is the mass of the disk and I, J are moments of inertia about an axis perpendicular to the plane of the disk and contained in the plane of the disk, respectively.

For simplicity, we will assume from now on that both the mass and the radius are unitary $m = R = 1$. Inspired by an example in [BFM09], we will consider a modified Lagrangian function $L^{mod} : TQ \to \mathbb{R}$ given by

$$L^{mod} = \frac{1}{2}\left(2\dot{\theta}^2 + \dot{x}^2 + \dot{y}^2 + \dot{\varphi}^2\right) - \dot{\theta}\left(\dot{x}\cos\varphi + \dot{y}\sin\varphi\right).$$

The trajectories of the Euler-Lagrange equations for L^{mod} with initial velocity in the distribution $\mathcal{D}$ are exactly the nonholonomic trajectories for $(L, \mathcal{D})$.

Indeed, the vector field $\Gamma_{L^{mod}}$ restricted to points of $\mathcal{D}$ is

$$\begin{aligned}(\Gamma_{L^{mod}})\Big|_{\mathcal{D}} &= \dot\theta\cos\varphi\frac{\partial}{\partial x}+\dot\theta\sin\varphi\frac{\partial}{\partial y}+\dot\theta\frac{\partial}{\partial\theta}+\dot\varphi\frac{\partial}{\partial\varphi}\\ &\quad -\dot\theta\dot\varphi\sin\varphi\frac{\partial}{\partial\dot x}++\dot\theta\dot\varphi\cos\varphi\frac{\partial}{\partial\dot y}\end{aligned}$$

and it is tangent to $\mathcal{D}$ since

$$\Gamma_{L^{mod}}(\dot x-\dot\theta\cos\varphi)=0,\qquad \Gamma_{L^{mod}}(\dot y-\dot\theta\sin\varphi)=0$$

which imply that the trajectories on $\mathcal{D}$ remains on $\mathcal{D}$ and satisfy

$$\dot x=\dot\theta\cos\varphi\,,\qquad \dot y=\dot\theta\sin\varphi\,,\qquad \ddot\theta=0\,,\qquad \ddot\varphi=0,$$

that is, they coincide precisely with the nonholonomic equations.

Consider now trajectories departing from the point q with coordinates $(0,0,0,0)$. We will show that the metric g^{mod} associated to the kinetic Lagrangian L^{mod} is related to a metric on $\mathcal{D}_q$ satisfying the Gauss condition.

Indeed, the Lagrangian L^{mod} is of the type

$$L^{mod}(v)=\frac{1}{2}g^{mod}(v,v),$$

where

$$(g^{mod})_{ij}=\begin{pmatrix}1&0&-\cos\varphi&0\\0&1&-\sin\varphi&0\\-\cos\varphi&-\sin\varphi&2&0\\0&0&0&1\end{pmatrix}.$$

Hence, the trajectories of the Euler-Lagrange equations for L^{mod} are just the geodesics with respect to g^{mod}. Denote by $\exp_q^{mod}:T_qQ\to Q$ the exponential map at q associated to g^{mod} and by $i_q:\mathcal{D}_q\hookrightarrow T_qQ$ the inclusion map. The fact that nonholonomic trajectories of $(L,\mathcal{D})$ coincide with geodesics of g^{mod} starting at $\mathcal{D}$ may be translated into the equation

$$\exp_q^{mod}\circ i_q=\exp_q^{nh},$$

where $\exp_q^{nh}:\mathcal{D}_q\to Q$ is the nonholonomic exponential map. Define now the Riemannian metric on $\mathcal{D}_q$ as

$$\mathcal{G}_0=(\exp_q^{nh})^*g^{mod}.$$

The nonholonomic exponential map may be computed to be

$$\exp_q^{nh}(u,v) = \left(\frac{u}{v}\sin v, \frac{u}{v}(1-\cos v), u, v\right)$$

if $v \neq 0$ and

$$\exp_q^{nh}(u,0) = (u,0,u,0)\,, \quad \text{if } v = 0.$$

Hence, we obtain

$$\mathcal{G}_0 = E du \otimes du + F du \otimes dv + F dv \otimes du + G dv \otimes dv,$$

with

$$E = \frac{2\left(v^2 - v\sin(v) - \cos(v) + 1\right)}{v^2}, \quad F = -\frac{u\left(v^2 - 2\cos(v) - 2v\sin(v) + 2\right)}{v^3}$$
$$G = \frac{u^2v^2 - 2u^2 v\sin(v) - 2u^2\cos(v) + 2u^2 + v^4}{v^4},$$

if $v \neq 0$ and

$$\mathcal{G}_0 = du \otimes du + \left(\frac{u^2}{4} + 1\right) dv \otimes dv, \quad \text{if } v = 0.$$

This metric is easily seen to satisfy the Gauss condition and moreover, the nonholonomic metric g_q^{nh} turns out to be simply the pullback of g^{mod} to the submanifold $\mathcal{M}_q^{nh}$! △

5.3 Trajectories of mechanical nonholonomic systems

In this section, we will state a nonholonomic version of the Maupertuis-Jacobi principle (check Theorem 3.2.4 for the classical unconstrained principle). Then, using Theorem 5.2.4, we will immediately deduce that radial nonholonomic mechanical trajectories with fixed energy $e \in \mathbb{R}$ are, for sufficiently small times, strictly increasing reparametrizations of minimizing Riemannian geodesics on a suitable Riemannian manifold. For the moment, we will restrict ourselves to the necessary results to achieve our immediate purposes. Later, on Chapter 6 (see Section 6.3), we will resume the discussion about mechanical nonholonomic systems, when studying nonholonomic Jacobi fields.

Let g be a Riemannian metric on the n-dimensional manifold Q, $V : Q \to \mathbb{R}$ the potential energy (with V smooth) and let $\mathcal{D}$ be a rank r distribution on Q. Let $L_{(g,V)} : TQ \to \mathbb{R}$ be the mechanical Lagrangian function associated with the Riemannian metric g and potential energy V, that is,

$$L_{(g,V)}(v) = \frac{1}{2}g(v,v) - V \circ \tau_Q(v), \quad v \in TQ.$$

The triple $(Q, L_{(g,V)}, \mathcal{D})$ is called a *nonholonomic mechanical system* and, as it is well-known, it is automatically a regular nonholonomic system. Thus the solutions of the nonholonomic system are the integral curves of the nonholonomic vector field $\Gamma_{(L_{(g,V)},\mathcal{D})} \in \mathfrak{X}(\mathcal{D})$ defined by equations (3.6.6). Hence, given $v_q \in \mathcal{D}$, denote by $c_{v_q} : I \to Q$ the corresponding unique trajectory with initial conditions

$$c_{v_q}(0) = q, \quad \dot{c}_{v_q}(0) = v_q.$$

The energy of the system $(Q, L_{(g,V)}, \mathcal{D})$ is given by the function $E_{L_{(g,V)}} : TQ \to \mathbb{R}$ defined by

$$E_{L_{(g,V)}}(v) = \frac{1}{2}g(v,v) + V \circ \tau_Q(v), \quad v \in TQ.$$

Recall that the energy is a first integral of the vector field $\Gamma_{(L_{(g,V)},\mathcal{D})}$, which implies that the energy is constant along the trajectories c_{v_q}, i.e.,

$$E_{L_{(g,V)}}(\dot{c}_{v_q}(t)) = e, \quad \forall t \in I,$$

where $e \in \mathbb{R}$ is some real number. Note that, $e \geqslant V(c_{v_q}(t))$, for every $t \in I$.

Fixing a real number $e \in \mathbb{R}$, it is possible to classify the mechanical trajectories into two different types:

(i) *Singular trajectories*: the energy of the trajectory c_{v_q} satisfies $e = V(q)$, which automatically implies that the initial velocity is zero $v_q = 0$.

(ii) *Regular trajectories*: the energy of the trajectory c_{v_q} satisfies $e > V(q)$ and the velocity of the trajectory may be written as

$$\|\dot{c}_{v_q}(t)\|^2 = 2(e - V(c_{v_q}(t))), \quad \forall t \in I.$$

So, there exists a real number $\varepsilon > 0$ such that the curve $c_{v_q} : (-\varepsilon, \varepsilon) \to Q$ is a regular trajectory.

Now, if for a fixed $e \in \mathbb{R}$, the curve c_{v_q} is a regular trajectory, that, is $e > V(q)$, then it is clear that the initial velocity is in the sphere centred at the zero vector 0_q and with radius $\sqrt{2(e - V(q))}$, which we will denote by

$$v_q \in S_g\left(\sqrt{2(e - V(q))}\right),$$

where the subscript g indicates that the norm is measured relative to the Riemannian metric g.

Remark 5.3.1. The set $\{q \in Q \mid e \geq V(q)\}$ is usually called the Hill region and the set $\{q \in Q \mid e = V(q)\}$ is called the Hill boundary or also sometimes called the zero velocity surface.

Now, take $e \in \mathbb{R}$ such that the set

$$U_e = \{q \in Q \mid e > V(q)\}$$

is a non-empty subset of Q. In fact, U_e is an open subset of Q and if it is non-empty, it inherits the smooth manifold structure of Q. We can consider on it the *Jacobi metric*

$$g_e = (e - V)g \tag{5.3.1}$$

and the kinetic nonholonomic system $(U_e, g_e, \mathcal{D}_e)$, where the distribution $\mathcal{D}_e$ is nothing but the fibers of $\mathcal{D}$ at the points in U_e. In other words, $\mathcal{D}_e = (\tau_{\mathcal{D}})^{-1}(U_e)$, where $\tau_{\mathcal{D}} : \mathcal{D} \to Q$ is the bundle projection.

Given a vector $v_q \in \mathcal{D}_e$, we will denote by $c^e_{v_q} : I \to U_e$ the nonholonomic trajectory of $(U_e, g_e, \mathcal{D}_e)$, with initial velocity v_q, that is

$$\nabla^{nh,e}_{\dot{c}^e_{v_q}(t)} \dot{c}^e_{v_q}(t) = 0\ , \qquad \dot{c}^e_{v_q}(0) = v_q \in \mathcal{D}_{c^e_{v_q}(0)} \tag{5.3.2}$$

where $\nabla^{nh,e}_X Y := P(\nabla^{g_e}_X Y) + \nabla^{g_e}_X [P'(Y)]$, $X, Y \in \Gamma(\mathcal{D}_e)$. Observe that since g_e and g are in the same conformal class of metrics the orthogonal projectors P and P' are the same for both metrics.

Therefore, there exists a SODE $\Gamma_{(L_{g_e}, \mathcal{D}_e)} \in \mathfrak{X}(\mathcal{D}_e)$ whose integral curves with initial velocity v_q are precisely the tangent lift of the trajectories $c^e_{v_q} : I \to U_e$.

Moreover, the energy of this system is simply given by the Lagrangian itself, that is, $E_{L_{g_e}} : TU_e \to \mathbb{R}$ coincides with the Lagrangian function $L_{g_e} : TU_e \to \mathbb{R}$ given by

$$L_{g_e}(u) = \frac{1}{2} g_e(u, u), \quad u \in TU_e.$$

Thus, $L_{g_e}|_{\mathcal{D}_e}$ is a first integral of $\Gamma_{(L_{g_e},\mathcal{D}_e)}$.

Moreover, it is not difficult to prove that if the trajectory $c^e_{v_q}$ has energy equal to 1, then the initial velocity v_q satisfies

$$v_q \in S_g\left(\sqrt{\frac{2}{e - V(q)}}\right),$$

using the same notation as before.

Now, let us introduce two projections and a suitable diffeomorphism between the two spheres mentioned above.

Let $\mathcal{P}_q : T_qQ \setminus \{0_q\} \to S_g\left(\sqrt{2(e - V(q))}\right)$ be the projection given by

$$\mathcal{P}_q(v_q) = \sqrt{2(e - V(q))}\frac{v_q}{\|v_q\|_g}$$

and $\mathcal{Q}_q : T_qQ \setminus \{0_q\} \to S_g\left(\sqrt{\frac{2}{e-V(q)}}\right)$ be the projection given by

$$\mathcal{Q}_q(v_q) = \sqrt{\frac{2}{e - V(q)}}\frac{v_q}{\|v_q\|_g}.$$

Consider the map $\Psi_q : S_g\left(\sqrt{2(e - V(q))}\right) \to S_g\left(\sqrt{\frac{2}{e-V(q)}}\right)$ that makes the diagram 5.1 below commute. Observe that Ψ_q is a diffeomorphism with

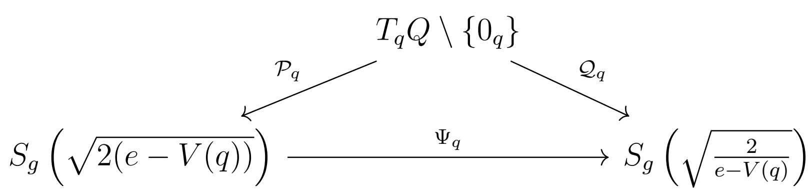


Figure 5.1: Definition of the diffeomorphism Ψ_q between spheres.

explicit expression

$$\Psi_q(v_q) = \frac{1}{(e - V(q))}v_q.$$

5.3.1 Nonholonomic Maupertuis-Jacobi principle

Now, we present the nonholonomic version of Maupertuis-Jacobi principle below relating nonholonomic mechanical trajectories with nonholonomic trajectories of the kinetic nonholonomic problem associated with the Jacobi metric.

Theorem 5.3.2 (Nonholonomic Maupertuis-Jacobi theorem). *Let $(Q, L_{(g,V)}, \mathcal{D})$ be a mechanical nonholonomic system, $q \in Q$ a fixed point of the manifold and let $e \in \mathbb{R}$ such that $e > V(q)$. For a non-zero $v_q \in T_qU_e$ denote by*

$$c_{\mathcal{P}_q(v_q)} : J \longrightarrow U_e \quad \textit{and} \quad c_{\mathcal{Q}_q(v_q)} : I \longrightarrow U_e \textit{ with } 0 \in I, J$$

the nonholonomic trajectories for the systems $(U_e, L_{(g,V)}|_{TU_e}, \mathcal{D}_e)$ and $(U_e, L_{g_e}, \mathcal{D}_e)$ with initial velocities $P_q(v_q)$ and $\mathcal{Q}_q(v_q)$, respectively. Then, we have that

$$c_{P_q(v_q)}(s) = c_{\mathcal{Q}_q(v_q)}(h(s)),$$

where $h : J \to I$ is a strictly increasing reparametrization satisfying

$$\frac{dh}{ds} = e - V \circ c_{P_q(v_q)}, \quad h(0) = 0.$$

This theorem will be proven in an intrinsic geometric setting called the contact bundle formulation of the nonholonomic Maupertuis-Jacobi principle. This section is precisely devoted to develop this formalism and proving the above results. Indeed, we develop the machinery we need to prove the nonholonomic Maupertuis-Jacobi Theorem 5.3.2. In order to do that, we must first discuss the symplectic bundle formulation of nonholonomic mechanical systems.

For the rest of the section, let $(Q, L_{(g,V)}, \mathcal{D})$ be a mechanical nonholonomic system with rank $\mathcal{D} = k$ and let $\Gamma_{(L_{(g,V)},\mathcal{D})}$ be the corresponding nonholonomic vector field.

The Lagrangian side

We will review the main ingredients of the construction given by [BS93] (see also [Cor02; Cor+09b]). In their paper, they define the set

$$\mathcal{T}^{\mathcal{D}}\mathcal{D} = \bigcup_{\substack{v_q \in \mathcal{D}_q \\ q \in Q}} \{X \in T_{v_q}\mathcal{D} \mid (T_{v_q}\tau_Q)(X) \in \mathcal{D}_q\}$$

which is a symplectic vector bundle of rank $2k$ over $\mathcal{D}$, that is,

$$(\mathcal{T}^{\mathcal{D}}_{v_q}\mathcal{D}, \omega_{L_{(g,V)}}(v_q)\Big|_{\mathcal{T}^{\mathcal{D}}_{v_q}\mathcal{D}})$$

is a symplectic vector space of dimension $2k$, for all $v_q \in \mathcal{D}_q$, where $\omega_{L_{(g,V)}}$ is the Poincaré-Cartan 2-form associated with the mechanical Lagrangian $L_{(g,V)}$.

Let $E_{L_{(g,V)}}$ be the corresponding Lagrangian energy. Then we have that

$$dE_{L_{(g,V)}}(v_q)\Big|_{\mathcal{T}^{\mathcal{D}}_{v_q}\mathcal{D}} \in (\mathcal{T}^{\mathcal{D}}_{v_q}\mathcal{D})^*, \text{ for all } v_q \in \mathcal{D}_q.$$

Moreover, we have that the nonholonomic vector field $\Gamma_{(L_{(g,V)},\mathcal{D})}$ is geometrically characterized by the equations

$$\begin{aligned} &\left(i_{\Gamma_{(L_{(g,V)},\mathcal{D})}}\omega_{L_{(g,V)}}|_{\mathcal{D}}\right)|_{\mathcal{T}^{\mathcal{D}}\mathcal{D}} = \left(dE_{L_{(g,V)}}|_{\mathcal{D}}\right)|_{\mathcal{T}^{\mathcal{D}}\mathcal{D}} \\ &\Gamma_{(L_{(g,V)},\mathcal{D})} \in \Gamma(\mathcal{T}^{\mathcal{D}}\mathcal{D}). \end{aligned} \tag{5.3.3}$$

As an immediate consequence, we deduce the preservation of energy for the nonholonomic trajectories:

$$\Gamma_{(L_{(g,V)},\mathcal{D})}(E_{L_{(g,V)}}|_{\mathcal{D}}) = 0. \tag{5.3.4}$$

The Hamiltonian side

Given a Riemannian metric g and a potential energy function V on the manifold Q, we may consider the Hamiltonian function $H_{(g,V)} : T^*Q \to \mathbb{R}$ given by

$$H_{(g,V)}(\alpha_q) = \frac{1}{2}g^\sharp_q(\alpha_q, \alpha_q) + V(q), \quad \alpha_q \in T^*_qQ,$$

where we are denoting by $g^\sharp$ the *co-metric* associated to the Riemannian metric g (see Section 3.3.1).

If $\mathcal{D}^\perp$ is the orthogonal complement of $\mathcal{D}$ with respect to the metric g and

$$(\mathcal{D}^\perp)^o = \bigcup_{q\in Q}\{\alpha_q \in T^*_qQ \mid \langle\alpha_q, v_q\rangle = 0, \ \forall v_q \in \mathcal{D}^\perp_q\}$$

then we have that

$$\mathbb{F}L_{(g,V)}(\mathcal{D}) = (\mathcal{D}^\perp)^o.$$

It is clear that $(i^*_{\mathcal{D}})|_{(\mathcal{D}^\perp)^o} : (\mathcal{D}^\perp)^o \to \mathcal{D}^*$ is an isomorphism of vector bundles where $i_{\mathcal{D}} : \mathcal{D} \hookrightarrow TQ$ is the canonical inclusion. From now on, we will use the previous canonical identification. We have that $T\mathbb{F}L_{(g,V)} = T\flat_g$ is a vector bundle isomorphism over $\mathbb{F}L_{(g,V)} = \flat_g$. Hence, if $\pi_Q : T^*Q \to Q$ is the cotangent bundle projection, considering the following vector bundle over $\mathcal{D}^*$

$$\mathcal{T}^{\mathcal{D}^*}\mathcal{D}^* = \bigcup_{\substack{\alpha_q \in \mathcal{D}^*_q \\ q \in Q}} \{Y \in T_{\alpha_q}\mathcal{D}^* \mid (T_{\alpha_q}\pi_Q)(Y) \in \mathcal{D}_q\},$$

we have that

$$T\flat_g(\mathcal{T}^{\mathcal{D}}\mathcal{D}) = \mathcal{T}^{\mathcal{D}^*}\mathcal{D}^*.$$

So, the previous discussion and the first equality in (3.3.2) imply that

$$(\mathcal{T}^{\mathcal{D}^*}\mathcal{D}^*, \omega_Q|_{\mathcal{T}^{\mathcal{D}^*}\mathcal{D}^* \times \mathcal{T}^{\mathcal{D}^*}\mathcal{D}^*})$$

is also a symplectic vector bundle over $\mathcal{D}^*$.

As a consequence, there exists a unique section $X_{(H_{(g,V)},\mathcal{D})} \in \Gamma(\mathcal{T}^{\mathcal{D}^*}\mathcal{D}^*)$ satisfying

$$\left(i_{X_{(H_{(g,V)},\mathcal{D})}}\omega_Q|_{\mathcal{D}^*}\right)|_{\mathcal{T}^{\mathcal{D}^*}\mathcal{D}^*} = \left(dH_{(g,V)}|_{\mathcal{D}^*}\right)|_{\mathcal{T}^{\mathcal{D}^*}\mathcal{D}^*}. \tag{5.3.5}$$

Note that, $X_{(H_{(g,V)},\mathcal{D})} \in \mathfrak{X}(\mathcal{D}^*)$. Moreover, since

$$X_{(H_{(g,V)},\mathcal{D})} \circ (\flat_g)|_{\mathcal{D}} = (T\flat_g)|_{\mathcal{T}^{\mathcal{D}}\mathcal{D}} \circ \Gamma_{(L_{(g,V)},\mathcal{D})} \tag{5.3.6}$$

it follows that if $\sigma : I \to \mathcal{D}^*$ is an integral curve of $X_{(H_{(g,V)},\mathcal{D})}$ then

$$\pi_Q \circ \sigma : I \to Q$$

is a trajectory of the nonholonomic mechanical system $(Q, L_{(g,V)}, \mathcal{D})$.

The kinetic nonholonomic system with Jacobi metric

Let $H_{(g,V)} : T^*Q \to \mathbb{R}$ be the Hamiltonian function associated to the mechanical nonholonomic system $(Q, L_{(g,V)}, \mathcal{D})$ and consider the corresponding Hamiltonian vector field $X_{(H_{(g,V)},\mathcal{D})} \in \mathfrak{X}(\mathcal{D}^*)$.

Suppose that $e \in \mathbb{R}$ is such that $U_e = \{q \in Q \mid e > V(q)\}$ is non-empty. Again consider the Jacobi metric g_e on U_e defined in (5.3.1) as well as the distribution $\mathcal{D}_e$ and its dual distribution

$$\mathcal{D}_e^* = \bigcup_{q \in U_e} \mathcal{D}_q^* \subseteq T^*U_e.$$

In the Hamiltonian side of the nonholonomic kinetic system $(U_e, g_e, \mathcal{D}_e)$, we will denote by $X_{(H_{g_e}, \mathcal{D}_e)} \in \mathfrak{X}(\mathcal{D}_e^*)$ the corresponding Hamiltonian vector field.

As we know

$$(\mathcal{T}^{\mathcal{D}_e^*}\mathcal{D}_e^*, \omega_Q|_{\mathcal{T}^{\mathcal{D}_e^*}\mathcal{D}_e^* \times \mathcal{T}^{\mathcal{D}_e^*}\mathcal{D}_e^*})$$

is a symplectic vector bundle over $\mathcal{D}_e^*$ and also

$$\left(i_{X_{(H_{g_e}, \mathcal{D}_e)}} \omega_Q|_{\mathcal{D}_e^*}\right)|_{\mathcal{T}^{\mathcal{D}_e^*}\mathcal{D}_e^*} = \left(dH_{g_e}|_{\mathcal{D}_e^*}\right)|_{\mathcal{T}^{\mathcal{D}_e^*}\mathcal{D}_e^*}, \tag{5.3.7}$$

where $H_{g_e} : T^*U_e \to \mathbb{R}$ is the Hamiltonian function in the Hamiltonian side of the kinetic nonholonomic system $(U_e, g_e, \mathcal{D}_e)$. It is important to note that the Hamiltonian function H_{g_e} is defined by

$$H_{g_e}(\alpha_q) = \frac{1}{2} g_e^\sharp(\alpha_q, \alpha_q),$$

where $g_e^\sharp$ is the *Jacobi co-metric* which is given by

$$g_e^\sharp = \frac{1}{e - V} g^\sharp. \tag{5.3.8}$$

For simplicity, when we have a co-metric $g^\sharp$ associated to a Riemannian metric g, we will denote by $\|\alpha_q\|_g$, with $\alpha_q \in T_q^*Q$, the norm on the fibers of T^*Q induced by $g^\sharp$, i.e.,

$$\|\alpha_q\|_g^2 = g^\sharp(\alpha_q, \alpha_q).$$

Let us introduce the subset S_e^* of $\mathcal{D}_e^*$ given by

$$S_e^* = \bigcup_{q \in U_e} \{\alpha_q \in \mathcal{D}_q^* \mid \|\alpha_q\|_g^2 = 2(e - V(q))\}.$$

Then we may prove the following result:

Theorem 5.3.3 (Contact bundle formulation of the nonholonomic Maupertuis-Jacobi principle)**.** *Using the notation we have introduced before, we have the following:*

1. *The subset S^*_e satisfies*

$$S^*_e = \left(H_{(g,V)}|_{\mathcal{D}^*_e}\right)^{-1}(e) = \left(H_{g_e}|_{\mathcal{D}^*_e}\right)^{-1}(1)$$

and, in addition, if $\alpha_q \in S^*_e$ then

$$\left(dH_{(g,V)}(\alpha_q)\right)|_{\mathcal{T}^{\mathcal{D}^*_e}\mathcal{D}^*_e} = \left(dH_{g_e}(\alpha_q)\right)|_{\mathcal{T}^{\mathcal{D}^*_e}\mathcal{D}^*_e} \neq 0,$$

so S^*_e is a submanifold of codimension 1 in $\mathcal{D}^*_e$. In fact,

$$\begin{aligned} T_{\alpha_q}S^*_e &= \{X \in T_{\alpha_q}\mathcal{D}^*_e \mid \langle dH_{(g,V)}(\alpha_q), X\rangle = 0\} \\ &= \{X \in T_{\alpha_q}\mathcal{D}^*_e \mid \langle dH_{g_e}(\alpha_q), X\rangle = 0\} \end{aligned}$$

and S^*_e is a bundle over U_e with fiber at $q \in U_e$ the sphere centred at $0_q \in \mathcal{D}^*_e$ and radius $\sqrt{2(e - V(q))}$, with respect to the Riemannian metric g.

2. *If $\mathcal{C}_e$ is defined by*

$$\mathcal{C}_e = \bigcup_{\substack{\alpha_q \in \mathcal{D}^*_q \\ q \in U_e}} \left(T_{\alpha_q}S^*_e \cap \mathcal{T}^{\mathcal{D}^*_e}_{\alpha_q}\mathcal{D}^*_e\right)$$

then $\mathcal{C}_e$ is a vector bundle over S^*_e which admits a contact bundle structure and the Reeb section R_e is just $X_{(H_{g_e},\mathcal{D}_e)}|_{\mathcal{D}^*_e}$.

3. *We have that*

$$(e - V)|_{U_e} R_e = X_{(H_{(g,V)},\mathcal{D})}|_{S^*_e}.$$

4. *If $v_q \in \mathcal{D}_q$ is a non-zero vector with $q \in U_e$ and $c_{\mathcal{P}_q(v_q)} : J \to U_e$, $c_{\mathcal{Q}_q(v_q)} : I \to U_e$ are the nonholonomic trajectories of the systems $(U_e, L_{(g,V)}|_{U_e}, \mathcal{D}_e)$, $(U_e, g_e, \mathcal{D}_e)$ with initial velocities $\mathcal{P}_q(v_q)$ and $\mathcal{Q}_q(v_q)$, respectively, then*

$$c_{\mathcal{P}_q(v_q)}(s) = c_{\mathcal{Q}_q(v_q)}(h(s)),$$

where $h : J \to I$ is a strictly increasing reparametrization satisfying

$$\frac{dh}{ds} = e - V \circ c_{\mathcal{P}_q(v_q)}, \quad h(0) = 0.$$

Remark 5.3.4. In the above theorem we used some notations introduced in the previous sections, namely the projections

$$\mathcal{P}_q : \mathcal{D}_q \setminus \{0_q\} \to \left(E_{L_{(g,V)}}\right)^{-1}(e) \cap \mathcal{D}_q$$

given by

$$\mathcal{P}_q(v_q) = \sqrt{2(e - V(q))}\frac{v_q}{\|v_q\|_g}$$

and $\mathcal{Q}_q : T_qQ \setminus \{0_q\} \to \left(E_{L_{g_e}}\right)^{-1}(1) \cap \mathcal{D}_q$ given by

$$\mathcal{Q}_q(v_q) = \sqrt{\frac{2}{e - V(q)}}\frac{v_q}{\|v_q\|_g}.$$

Proof. Let us prove each item in the Theorem by order of appearance:

1. We have that

$$\|\alpha_q\|_g^2 = 2(e - V(q))$$

   is equivalent to

$$\frac{\|\alpha_q\|_g^2}{2(e - V(q))} = 1$$

   and so, using the definition of the Jacobi co-metric $g_e^\sharp$ in (5.3.8) we have that

$$\frac{1}{2}\|\alpha_q\|_{g_e}^2 = 1,$$

   which proves that $\alpha_q \in \left(H_{(g,V)}|_{\mathcal{D}_e^*}\right)^{-1}(e)$ if and only if $\alpha_q \in \left(H_{g_e}|_{\mathcal{D}_e^*}\right)^{-1}(1)$.

   Now, let Δ^* be the Euler vector field of $\mathcal{D}^*$ defined by

$$\Delta^*(\alpha_q) = (\alpha_q)_{\alpha_q}^V = \frac{d}{dt}\Big|_{t=0}((1+t)\alpha_q) \in \mathcal{T}_{\alpha_q}^{\mathcal{D}_e^*}\mathcal{D}_e^*.$$

   Then, if $\alpha_q \in S_e^*$ we have that

$$\left\langle dH_{(g,V)}(\alpha_q), \Delta^*(\alpha_q)\right\rangle = \|\alpha_q\|_g^2 = 2(e - V(q)) > 0$$

   as well as

$$\langle dH_{g_e}(\alpha_q), \Delta^*(\alpha_q)\rangle = \|\alpha_q\|_{g_e}^2 = 2 > 0.$$

Hence, S_e^* is a submanifold of $\mathcal{D}_e^*$ of codimension 1 and

$$\begin{aligned} T_{\alpha_q} S_e^* &= \{X \in T_{\alpha_q}\mathcal{D}_e^* \mid \langle dH_{(g,V)}(\alpha_q), X\rangle = 0\} \\ &= \{X \in T_{\alpha_q}\mathcal{D}_e^* \mid \langle dH_{g_e}(\alpha_q), X\rangle = 0\}. \end{aligned}$$

Thus,

$$T_{\alpha_q}\mathcal{D}_e^* = T_{\alpha_q} S_e^* \oplus \langle \Delta^*(\alpha_q)\rangle.$$

Therefore, using that Δ^* is vertical with respect to the projection $\tau_e^* : \mathcal{D}_e^* \to U_e$, it follows that the restriction of τ_e^* to S_e^* is also a bundle with projection $\tau_e^*|_{S_e^*} : S_e^* \to U_e$. In addition, it is easy to prove that the fiber of $\tau_e^*|_{S_e^*}$ at $q \in U_e$ is just the sphere centred at $0_q \in \mathcal{D}_e^*$ and radius $\sqrt{2(e - V(q))}$, with respect to the Riemannian metric g.

2. If $\alpha_q \in S_e^*$ then, from the previous item, we deduce that the set

$$T_{\alpha_q}^{\mathcal{D}_e^*}\mathcal{D}_e^* \cap T_{\alpha_q} S_e^*$$

is a vector subspace of codimension 1 of $\mathcal{T}_{\alpha_q}^{\mathcal{D}_e^*}\mathcal{D}_e^*$. Therefore,

$$\mathcal{C}_e = \bigcup_{\substack{\alpha_q \in \mathcal{D}_q^* \\ q \in U_e}} \left(T_{\alpha_q} S_e^* \cap \mathcal{T}_{\alpha_q}^{\mathcal{D}_e^*}\mathcal{D}_e^* \right)$$

is a vector bundle over S_e^* with rank $2r-1$.

Now, we consider the sections $(\theta_Q)_e$ and $(\omega_Q)_e$ of the vector bundles $\mathcal{C}_e^* \to S_e^*$ and $\Lambda^2(\mathcal{C}_e^*) \to S_e^*$, respectively, given by

$$(\theta_Q)_e(\alpha_q) = \frac{1}{2}\theta_Q(\alpha_q)|_{(\mathcal{C}_e)_{\alpha_q}}$$

and

$$(\omega_Q)_e(\alpha_q) = \frac{1}{2}\omega_Q(\alpha_q)|_{(\mathcal{C}_e)_{\alpha_q} \times (\mathcal{C}_e)_{\alpha_q}}$$

for $\alpha_q \in S_e^*$.

We will see that $((\theta_Q)_e, (\omega_Q)_e)$ is a contact bundle structure on the vector bundle $C_e \to S_e^*$, that is,

$$(\theta_Q)_e \wedge (\omega_Q)_e^{r-1} \in \Gamma(\Lambda^{2r}(C_e^*))$$

is non-vanishing at every point of S_e^*. In fact, using that

$$X_{(g_e,\mathcal{D}_e)}(\alpha_q)\left(H_{g_e}|_{\mathcal{D}_e^*}\right)=0$$

it follows that $X_{(g_e,\mathcal{D}_e)}(\alpha_q)\in(\mathcal{C}_e)_{\alpha_q}$. Thus, we deduce that $X_{(g_e,\mathcal{D}_e)}|_{S_e^*}\in\Gamma(\mathcal{C}_e)$. In addition,

$$\begin{aligned}\langle(\theta_Q)_e(\alpha_q),X_{(g_e,\mathcal{D}_e)}(\alpha_q)\rangle&=\frac{1}{2}\langle\theta_Q(\alpha_q),X_{(g_e,\mathcal{D}_e)}(\alpha_q)\rangle\\&=\frac{1}{2}\langle\alpha_q,T_{\alpha_q}\tau_e^*\left(X_{(g_e,\mathcal{D}_e)}(\alpha_q)\right)\rangle,\end{aligned}$$

where we used the definition of the canonical 1-form of the cotangent bundle and $\tau_e^*:\mathcal{D}_e^*\to U_e$ is the bundle projection. On the other hand, from (5.3.6), we have that

$$X_{(g_e,\mathcal{D}_e)}\circ(\flat_{g_e})|_{\mathcal{D}_e}=(T\flat_{g_e})|_{\mathcal{T}^{\mathcal{D}_e}\mathcal{D}_e}\circ\Gamma_{(g_e,\mathcal{D}_e)}.$$

Now, denote by $\sharp_{g_e}:\mathcal{D}_e^*\to\mathcal{D}_e$ the inverse map of the flat isomorphism $\flat_{g_e}:\mathcal{D}_e\to\mathcal{D}_e^*$. Then, using that

$$T_{\alpha_q}\tau_e^*\circ(T\flat_g)|_{\mathcal{T}^{\mathcal{D}}\mathcal{D}}=T_{\sharp_{g_e}(\alpha_q)}\tau_e,$$

where $\tau_e:\mathcal{D}_e\to U_e$ is the canonical bundle projection, we deduce that

$$\langle(\theta_Q)_e(\alpha_q),X_{(g_e,\mathcal{D}_e)}(\alpha_q)\rangle=\frac{1}{2}\langle\alpha_q,T_{\sharp_{g_e}(\alpha_q)}\tau_{\mathcal{D}_e}\left(\Gamma_{(g_e,\mathcal{D}_e)}\circ\sharp_{g_e}(\alpha_q)\right)\rangle.$$

But, since $\Gamma_{(g_e,\mathcal{D}_e)}$ is a SODE the previous relation reduces to

$$\begin{aligned}\langle(\theta_Q)_e(\alpha_q),X_{(g_e,\mathcal{D}_e)}(\alpha_q)\rangle&=\frac{1}{2}\langle\alpha_q,\sharp_{g_e}(\alpha_q)\rangle\\&=\frac{1}{2}\|\alpha_q\|_{g_e}^2=1.\end{aligned}$$

Moreover, we have that

$$\left[i_{X_{(g_e,\mathcal{D}_e)}(\alpha_q)}(\omega_Q)_e(\alpha_q)\right]\Big|_{(\mathcal{C}_e)_{\alpha_q}}=dH_{g_e}(\alpha_q)|_{(\mathcal{C}_e)_{\alpha_q}}=0.$$

This implies that $((\theta_Q)_e,(\omega_Q)_e)$ is a contact bundle structure on the vector bundle $\mathcal{C}_e$ and that $X_{(g_e,\mathcal{D}_e)}|_{S_e^*}\in\Gamma(C_e)$ is the Reeb section of this contact structure, that is,

$$i_{X_{(g_e,\mathcal{D}_e)}|_{S_e^*}}(\theta_Q)_e=1,\quad i_{X_{(g_e,\mathcal{D}_e)}|_{S_e^*}}(\omega_Q)_e=0.$$

3. Using that
$$\left(X_{(H_{(g,V)},\mathcal{D})}|_{S_e^*}\right)\left(H_{(g,V)}|_{\mathcal{D}_e}\right) = 0$$
it follows that $X_{(H_{(g,V)},\mathcal{D})}|_{S_e^*} \in \Gamma(\mathcal{C}_e)$. In addition, proceeding as in the previous item, one may prove that if $\alpha_q \in S_e^*$ then
$$\langle (\theta_Q)_e(\alpha_q), X_{(H_{(g,V)},\mathcal{D})}(\alpha_q)\rangle = \frac{1}{2}\|\alpha_q\|_g^2 = e - V(q)$$
and
$$\left[i_{X_{(H_{(g,V)},\mathcal{D})}}(\alpha_q)(\omega_Q)_e(\alpha_q)\right]\Big|_{(\mathcal{C}_e)_{\alpha_q}} = dH_{(g,V)}(\alpha_q)|_{(\mathcal{C}_e)_{\alpha_q}} = 0.$$
Therefore,
$$(e - V(q))|_{U_e} X_{(H_{g_e},\mathcal{D}_e)}|_{S_e^*} = X_{(H_{(g,V)},\mathcal{D})}|_{S_e^*}. \tag{5.3.9}$$

4. It is easy to prove that the following diagram commutes:

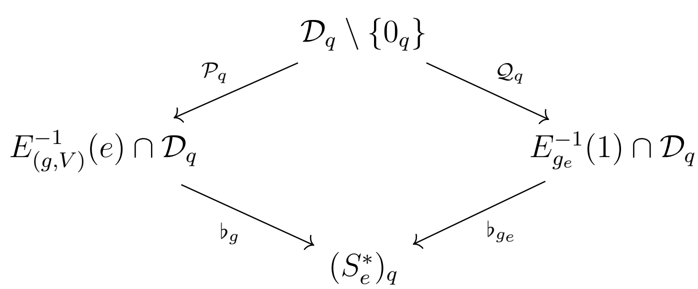


Figure 5.2: Commutative diagram.

Thus, if $v_q \in \mathcal{D}_q \setminus \{0_q\}$ then
$$\flat_g(\mathcal{P}_q(v_q)) = \flat_{g_e}(\mathcal{Q}_q(v_q)) = \alpha_q \in S_e^*. \tag{5.3.10}$$
Now, we consider the integral curves $\sigma_{\alpha_q} : J \to S_e^*$ and $\sigma^e_{\alpha_q} : I \to S_e^*$ (with $0 \in I, J$) of the vector fields $X_{(g,V,\mathcal{D})}|_{S_e^*}$ and $X_{(g_e,\mathcal{D}_e)}|_{S_e^*}$, respectively, satisfying the initial conditions
$$\sigma_{\alpha_q}(0) = \sigma^e_{\alpha_q}(0) = \alpha_q.$$

Then, using Equation (5.3.9) in the previous item , it follows that there exists a strictly increasing reparametrization $h : J \to I$ such that

$$\frac{dh}{ds} = e - V \circ \pi_{\mathcal{D}^*} \circ \sigma_{\alpha_q}, \quad h(0) = 0$$

and

$$\sigma_{\alpha_q}(s) = \sigma^e_{\alpha_q}(h(s)), \quad \text{for } s \in J,$$

with $\pi_{\mathcal{D}^*} : \mathcal{D}^* \to Q$ the canonical projection. But, recall that, if $v_q = \sharp_q(\alpha_q)$ then using (5.3.6), Figure 5.2 and (5.3.10), we deduce that

$$\pi_{\mathcal{D}^*} \circ \sigma_{\alpha_q} = c_{\mathcal{P}_q(v_q)} \quad \text{and} \quad \pi_{\mathcal{D}^*} \circ \sigma^e_{\alpha_q} = c_{\mathcal{Q}_q(v_q)},$$

which implies the result.

□

Remark 5.3.5. A coordinate derivation of nonholonomic Maupertuis-Jacobi principle (see also [Koi92])

Having chosen a system of coordinates (q^i), $1 \leq i \leq n = \dim Q$ then we induce a system of coordinates $(q^i, \dot{q}^i)$ on TQ. In these coordinates, the Lagrangian $L_{(g,V)} : TQ \to \mathbb{R}$ is written as

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

where $g_{ij} = g(\partial/\partial q^i, \partial/\partial q^j)$. The linear velocity constraints are determined by the distribution $\mathcal{D}$ where rank$\mathcal{D} = m \leq n$ and it is locally determined by its annihilator:

$$\mathcal{D}^o = \text{span}\{\mu^\alpha = \mu^\alpha_i(q)\, dq^i; m + 1 \leq \alpha \leq n\}$$

However in the case of nonholonomic mechanics it can be better to adapt the coordinates on the tangent bundle to the linear velocity constraints and to the Riemannian metric. To this end, consider a local basis $\{X_a, Y_\alpha\}$, $1 \leq a \leq m$ and $m + 1 \leq \alpha \leq n$ of vector fields such that locally

$$\mathcal{D}_q = \text{span}\{X_a(q)\} \quad \text{and} \quad \mathcal{D}^{\perp,g}_q = \text{span}\{Y_\alpha(q)\},$$

where $\mathcal{D}^{\perp,g}_q$ is the Riemannian-orthogonal to $\mathcal{D}$, i.e.

$$g(X_a, Y_\alpha) = 0, \quad 1 \leq a \leq m \quad \text{and} \quad m + 1 \leq \alpha \leq n.$$

Denote by $g_{ab} = g(X_a, X_b)$ and consider the Lie bracket:

$$[X_a, X_b] = \mathcal{C}^c_{ab} X_c + \mathcal{C}^\alpha_{ab} Y_\alpha$$

Observe that the non-vanishing of some of functions $\mathcal{C}^\alpha_{ab}$ implies the non-integrability of the distribution $\mathcal{D}$.

Obviously we have that $T_q Q = \mathcal{D}_q \oplus \mathcal{D}_q^{\perp, g}$. Therefore, the adapted basis $\{X_a, Y_\alpha\}$ induces a new set of coordinates on the tangent bundle (q^i, y^a, y^α). Observe that the elements $v_q \in \mathcal{D}_q$ are distinguished by the condition $y^\alpha = 0$. That is, the nonholonomic constraints are now $y^\alpha = 0$ and $\mathcal{D}$ is completely described by coordinates (q^i, y^a).

Denote by $\{X^a, Y^\alpha\}$ the dual basis corresponding to $\{X_a, Y_\alpha\}$ inducing coordinates (q^i, p_a, p_α) on T^*Q and (q^i, p_a) on D^*. The Hamiltonian is now

$$H_{(g,V)}|_{\mathcal{D}^*}(q^i, p_a) = \frac{1}{2} g^{ab}(q) p_a p_b + V(q)\,.$$

The equations of motion of a nonholonomic system are written in the system of adapted coordinates (q^i, p_a) as follows (see, for instance, [Cor+09a; Cel+19]):

$$\begin{aligned}
\dot{q}^i &= X^i_b \frac{\partial H_{(g,V)}|_{\mathcal{D}^*}}{\partial p_b} = X^i_b g^{ab} p_a \\
\dot{p}_a &= -\mathcal{C}^c_{ab} p_c \frac{\partial H_{(g,V)}|_{\mathcal{D}^*}}{\partial p_b} - X^i_a \frac{\partial H_{(g,V)}|_{\mathcal{D}^*}}{\partial q^i} \\
&= -\mathcal{C}^c_{ab} g^{bd} p_c p_d - X^i_a \left(\frac{1}{2} \frac{\partial g^{cb}}{\partial q^i} p_c p_b + \frac{\partial V}{\partial q^i} \right),
\end{aligned} \tag{5.3.11}$$

where $X_a = X^i_a \frac{\partial}{\partial q^i}$. The dynamics is precisely the given by the vector field $X_{(H_{(g,V)}, \mathcal{D})}$ intrinsically defined in Equation (5.3.5).

From the other hand, if we consider the Hamiltonian $H_{g_e}|_{\mathcal{D}^*_e} : \mathcal{D}^*_e \to \mathbb{R}$:

$$H_{g_e}|_{\mathcal{D}^*_e}(q^i, p_a) = \frac{1}{2(e - V(q))} g^{ab} p_a p_b\,.$$

Then the corresponding nonholonomic equations are:

$$\begin{aligned}
\dot{q}^i =& \frac{1}{e - V(q)} X^i_b g^{ab} p_a \\
\dot{p}_a =& -\frac{1}{e - V(q)} \mathcal{C}^c_{ab} g^{bd} p_c p_d \\
& - X^i_a(q) \left(\frac{1}{2(e - V(q))} \frac{\partial g^{cb}}{\partial q^i} p_c p_b + \frac{1}{2(e - V(q))^2} \frac{\partial V}{\partial q^i} g^{cb} p_c p_b \right)
\end{aligned} \tag{5.3.12}$$

These equations are precisely the ones defined by the integral curves of the vector field $X_{(H_{ge},\mathcal{D}_e)}$ given in Equation (5.3.7).

Therefore

$$X_{(H_{ge},\mathcal{D}_e)} - \frac{1}{e-V(q)} X_{(H_{(g,V)},\mathcal{D})}|_{D_e^*} = X_a^i(q)\left(\frac{1}{2(e-V(q))}\frac{\partial V}{\partial q^i} g^{cb} p_c p_b - \frac{\partial V}{\partial q^i}\right)\frac{\partial}{\partial p_a}$$

Along the set $S_e^* = \left(H_{(g,V)}|_{\mathcal{D}_e^*}\right)^{-1}(e)$ we have that $\frac{1}{2} g^{cb} p_c p_b = e - V(q)$ and in consequence,

$$\mathcal{R}_e = X_{(H_{ge},\mathcal{D}_e)}|_{S_e^*} = \frac{1}{e-V(q)} X_{(H_{(g,V)},\mathcal{D})}|_{S_e^*}$$

as appears in Theorem 5.3.3.

5.3.2 The geodesic character of radial mechanical trajectories

We are now in position to formulate the main result of this section.

Theorem 5.3.6. *Let $(Q,(g,V),\mathcal{D})$ be a mechanical nonholonomic system, $q \in Q$ a fixed point of the manifold Q and let $e \in \mathbb{R}$ such that $e > V(q)$. Then:*

i) There exists $\varepsilon > 0$ and a submanifold $\mathcal{M}_q^{nh,e} \subset Q$ with $q \in \mathcal{M}_q^{nh,e}$ and a diffeomorphism

$$\exp_q^{nh,e} : B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e-V(q)}}\right) \subseteq \mathcal{D}_q \to \mathcal{M}_q^{nh,e},$$

where the domain denotes the open ball in $\mathcal{D}_q$ around 0_q with radius $\sqrt{\frac{2\varepsilon}{e-V(q)}}$, with respect to the Riemannian metric g.

Moreover we have that $\exp_q^{nh,e}(0_q) = q$ and:

(a) The tangent map of $\exp_q^{nh,e}$ at 0_q, under the canonical linear identification between $\mathcal{D}_q$ and $T_{0_q}\left(B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e-V(q)}}\right)\right)$,

$$T_{0_q}\exp_q^{nh,e} : \mathcal{D}_q \longrightarrow T_qQ,$$

is just the canonical inclusion of $\mathcal{D}_q$ in T_qQ.

(b) For every non-zero vector $v_q \in B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e-V(q)}}\right)$ *the nonholonomic mechanical trajectory* $c_{\mathcal{P}_q(v_q)} : [0,\lambda] \to Q$ *satisfies*

$$c_{\mathcal{P}_q(v_q)}(s) = \exp_q^{nh,e}(h(s)\mathcal{Q}_q(v_q)), \tag{5.3.13}$$

where $h : [0,\lambda] \to [0,\delta]$ *is a strictly increasing reparametrization satisfying*

$$\frac{dh}{ds} = e - V \circ c_{\mathcal{P}_q(v_q)}, \quad h(0) = 0$$

and λ *is sufficiently small in such a way that*

$$h(s)\mathcal{Q}_q(v_q) \in B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e - V(q)}}\right), \quad \forall s \in [0,\lambda].$$

ii) All the nonholonomic trajectories with starting point q *and energy* e *are, for sufficiently small times, of the form (5.3.13). In addition, if* $g_q^{nh,e}$ *is a Riemannian metric on* $\mathcal{M}_q^{nh,e}$ *such that* $\mathcal{G}_0^e = (\exp_q^{nh,e})^* g_q^{nh,e}$ *satisfies the Gauss condition, then the curves*

$$t \in [0,1] \mapsto \exp_q^{nh,e}(tv_q) \in \mathcal{M}_q^{nh,e},$$

with $v_q \in B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e-V(q)}}\right)$ *are geodesics for* $g_q^{nh,e}$ *and, therefore, the nonholonomic trajectories*

$$s \in [0,\lambda] \mapsto c_{P_q(v_q)}(s) \in \mathcal{M}_q^{nh,e}$$

are reparametrizations of minimizing geodesics for the metric $g_q^{nh,e}$*. In particular, these nonholonomic trajectories minimize length in* $\mathcal{M}_q^{nh,e}$*.*

iii) The Riemannian metrics $g_q^{nh,e}$ *on* $\mathcal{M}_q^{nh,e}$ *always exist.*

Remark 5.3.7. We have that the map

$$\exp_q^{nh,e} : B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e - V(q)}}\right) \subseteq \mathcal{D}_q \longrightarrow U_e \subseteq Q$$

is given by

$$\exp_q^{nh,e}(v_q) = \tau_Q\left(\phi_1^{\Gamma_{(g_e,\mathcal{D}_e)}}(v_q)\right)$$

for $v_q \in B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e-V(q)}}\right)$ and where $\phi_t^{\Gamma_{(g_e,\mathcal{D}_e)}}$ is the flow of the SODE $\Gamma_{(g_e,\mathcal{D}_e)}$ along $\mathcal{D}_e$. In other words, $\exp_q^{nh,e}$ is the nonholonomic exponential map at q associated with the kinetic non-holonomic system $(U_e, g_e, \mathcal{D}_e)$.

Proof of Theorem 5.3.6. Now we have all the ingredients to prove of Theorem 5.3.6 since it is a direct consequence combining first the nonholonomic Maupertuis-Jacobi principle stated in Theorem 5.3.3 and then Theorem 5.2.4. We just add a few reasons why we take the open ball $B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e-V(q)}}\right)$, with ε a sufficiently small positive number, as the domain of the map $\exp_q^{nh,e}$:

- It is clear that $B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e-V(q)}}\right)$ is a star-shaped open subset of $\mathcal{D}_q$ about $0_q \in \mathcal{D}_q$;
- If $v_q \in \mathcal{D}_q \setminus \{0_q\}$, then $\mathcal{Q}_q(v_q) \in B_g\left(0_q; \sqrt{\frac{2}{e-V(q)}}\right)$. So, if we fix $\varepsilon > 0$ small enough, it is possible to choose a sufficiently small positive number λ such that

$$h(s)\mathcal{Q}_q(v_q) \in B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e-V(q)}}\right), \; \forall s \in [0, \lambda]$$

  (note that $h(0) = 0$);
- Using the previous facts, we can directly apply Theorem 5.2.4 to the map $\exp_q^{nh,e} : B_g\left(0_q; \sqrt{\frac{2\varepsilon}{e-V(q)}}\right) \subseteq \mathcal{D}_q \to Q$.

$\square$

Example 5.3.8. Let us first consider a mechanical nonholonomic system describing a particle with unitary mass in euclidean three dimensional space $Q = \mathbb{R}^3$ equipped with the euclidean metric g, subjected to a potential force $V : Q \to \mathbb{R}$ given by

$$V(x, y, z) = z,$$

and to the nonholonomic constraint determined by

$$\mathcal{D} = \{(q, \dot{q}) \in TQ \mid \dot{z} = y\dot{x}\}.$$

Let $e \in R$ be a fixed energy value and consider the set

$$U_e = \{(x, y, z) \in Q \mid z < e\}$$

where the Jacobi metric

$$g_e = (e - z)g$$

is defined. The kinetic nonholonomic system $(g_e, \mathcal{D}_e)$ associated to the mechanical nonholonomic system $(g, V, \mathcal{D})$ is associated to the kinetic Lagrangian $L_{g_e} : TU_e \to \mathbb{R}$ given by

$$L_{g_e}(q, \dot{q}) = \frac{e-z}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right).$$

To observe explicitly the results of Theorem 5.3.3, it is easier to work on the Hamiltonian side and using a basis adapted to $\mathcal{D}$, as in Remark 5.3.5. In that sense, we will use the basis given by

$$X_1 = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}, \quad X_2 = \frac{\partial}{\partial y}$$

spanning $\mathcal{D}$ and the vector

$$Y_1 = -y\frac{\partial}{\partial x} + \frac{\partial}{\partial z}$$

spanning the orthogonal complement $\mathcal{D}^{\perp}$. Hence, we obtain the following non-vanishing components of the Riemannian metric

$$g_{11} = 1 + y^2, \quad g_{22} = 1.$$

Finally, the non-vanishing structure functions (C_{ab}^c) relative to this basis are

$$C_{12}^1 = -\frac{y}{y^2+1} = -C_{23}^3, \quad C_{23}^1 = -\frac{1}{y^2+1} = C_{12}^3.$$

The Hamiltonian function is written with respect to this basis as

$$H_{(g,V)}|_{\mathcal{D}^*}(q^i, p_a) = \frac{1}{2}\left(\frac{p_1^2}{y^2+1} + p_2^2\right) + z$$

and the corresponding Hamiltonian equations in this adapted coordinates are

$$\begin{cases} \dot{x} = \frac{p_1}{y^2+1} \\ \dot{y} = p_2 \\ \dot{z} = \frac{yp_1}{y^2+1} \end{cases} \qquad \begin{cases} \dot{p}_1 = \frac{yp_1p_2}{y^2+1} - y \\ \dot{p}_2 = 0 \end{cases}$$

On the other hand, the kinetic Hamiltonian function $H_{g_e} : T^*U_e \to \mathbb{R}$ is given on these coordinates by

$$H_{g_e}(q^i, p_a) = \frac{1}{2(e-z)}\left(\frac{p_1^2}{y^2+1} + p_2^2\right)$$

implying the following Hamiltonian equations

$$\begin{cases} \dot{x} = \frac{1}{e-z}\frac{p_1}{y^2+1} \\ \dot{y} = \frac{1}{e-z}p_2 \\ \dot{z} = \frac{1}{e-z}\frac{yp_1}{y^2+1} \end{cases} \qquad \begin{cases} \dot{p}_1 = \frac{1}{e-z}\frac{yp_1p_2}{y^2+1} - \frac{y}{2(e-z)^2}\left(\frac{p_1^2}{y^2+1} + p_2^2\right) \\ \dot{p}_2 = 0. \end{cases}$$

Then it is clear that if we restrict to the set $S_e^* = \left(H_{(g,V)}|_{\mathcal{D}_e^*}\right)^{-1}(e)$, we have that

$$\frac{1}{2(e-z)}\left(\frac{p_1^2}{y^2+1} + p_2^2\right) = e - z$$

on S_e^*, showing that

$$X_{(H_{g_e},\mathcal{D}_e)}|S_e^* = \frac{1}{e-z}X_{(H_{(g,V)},\mathcal{D})}|S_e^*.$$

It is now clear that the integral curves of $X_{(H_{(g,V)},\mathcal{D})}$ must be a reparametrization of the integral curves of the Hamiltonian vector field $X_{(H_{g_e},\mathcal{D}_e)}$ on S_e^*. △

Example 5.3.9. The vertical rolling disk with harmonic potential in the steering angle. Consider the mechanical Lagrangian function $L_{(g,V)} : TQ \to \mathbb{R}$ in the manifold $Q = \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$ given by

$$L_{(g,V)}(q,\dot{q}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{\theta}^2 + \dot{\varphi}^2) - \frac{\varphi^2}{2},$$

subject to the constraint

$$\mathcal{D} = \{(q,\dot{q}) \in TQ \mid \dot{x} = \dot{\theta}\cos\varphi,\ \dot{y} = \dot{\theta}\sin\varphi\}.$$

It is not difficult to show that the general solution is

$$\begin{cases} x(t) & = \int_0^t \cos(\varphi(s))\ ds + x_0 \\ y(t) & = \int_0^t \sin(\varphi(s))\ ds + y_0 \\ \theta(t) & = \Omega t + \theta_0 \\ \varphi(t) & = \varphi_0\cos(t) + \omega\sin(t), \end{cases}$$

with $q_0 = (x_0, y_0, \theta_0, \varphi_0) \in Q$, $(\Omega, \omega) \in \mathbb{R}^2$ a coordinate chart on $\mathcal{D}_{q_0}$, representing the initial angular velocities. Then the nonholonomic exponential map $\exp_{q_0}^{nh} : \mathcal{D}_{q_0} \to Q$

$$\exp_{q_0}^{nh}(v_{q_0}) = (\tau_Q \circ \phi_1^{\Gamma_{(L_{(g,V)},\mathcal{D})}})(v_q),$$

where $\phi_t^{\Gamma_{(L_{(g,V)},\mathcal{D})}}$ is the flow of the nonholonomic mechanical system $(L_{(g,V)}, \mathcal{D})$, is a local diffeomorphism onto its image and so its inverse map is $R^{nh}_{q_0} : \mathcal{M}^{nh}_{q_0} \to \mathcal{D}_{q_0}$ given by

$$R^{nh}_{q_0}(\theta, \varphi) = \left(\theta - \theta_0, \frac{\varphi - \varphi_0 \cos(1)}{\sin(1)}\right).$$

The corresponding kinetic nonholonomic system is determined by the Lagrangian function $L_{g_e} : TU_e \to \mathbb{R}$ given by

$$L_{g_e}(q, \dot{q}) = \frac{e - \frac{\varphi^2}{2}}{2}(\dot{x}^2 + \dot{y}^2 + \dot{\theta}^2 + \dot{\varphi}^2).$$

After some computations, we may eliminate the Lagrange multipliers appearing in Lagrange-d'Alembert equations and find that the trajectories of the nonholonomic system $(L_{g_e}, \mathcal{D}_e)$ must satisfy

$$\begin{cases} \dot{x} &= \dot{\theta}\cos\varphi \\ \dot{y} &= \dot{\theta}\sin\varphi \\ \ddot{\theta} &= \frac{2\varphi\dot{\varphi}\dot{\theta}}{e - \frac{\varphi^2}{2}} \\ \ddot{\varphi} &= \frac{\varphi\dot{\varphi}^2 - \varphi\dot{\theta}^2}{e - \frac{\varphi^2}{2}}. \end{cases}$$

Then the trajectories of this system form the exponential map $\exp^{(g_e, \mathcal{D}_e)}_{q_0} : \mathcal{D}_{q_0} \to Q$.

Moreover, using Theorem 5.3.2, we know there is a strictly increasing function $h : J \to I$ satisfying

$$\frac{dh}{ds} = e - V \circ c_{P_q(v_q)}, \quad h(0) = 0.$$

Solving the differential equation above, we obtain that

$$h(s) = es - \frac{1}{2}\left(\frac{(\varphi_0^2 - \omega^2)\cos s \sin s}{2} + \frac{(\varphi_0^2 + \omega^2)s}{2} + \varphi_0\omega\sin^2 s\right).$$

△

Example 5.3.10. The vertical rolling disk with linear potential in the steering angle. Consider the mechanical Lagrangian function $L_{(g,V)} : TQ \to \mathbb{R}$ in the manifold $Q = \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$ given by

$$L_{(g,V)}(q, \dot{q}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{\theta}^2 + \dot{\varphi}^2) - \varphi,$$

subject to the constraint

$$\mathcal{D} = \{(q, \dot{q}) \in TQ \mid \dot{x} = \dot{\theta}\cos\varphi,\ \dot{y} = \dot{\theta}\sin\varphi\}.$$

It is not difficult to show that the general solution is

$$\begin{cases} x(t) & = \int_0^t \cos(\varphi(s))\ ds + x_0 \\ y(t) & = \int_0^t \sin(\varphi(s))\ ds + y_0 \\ \theta(t) & = \Omega t + \theta_0 \\ \varphi(t) & = \omega t + \varphi_0 - \frac{t^2}{2}, \end{cases}$$

with $q_0 = (x_0, y_0, \theta_0, \varphi_0) \in Q$ and $(\Omega, \omega) \in \mathbb{R}^2$ a coordinate chart on $\mathcal{D}_{q_0}$, representing the initial angular velocities. Then the nonholonomic exponential map $\exp^{nh}_{q_0} : \mathcal{D}_{q_0} \to Q$ given by

$$\exp^{nh}_{q_0}(v_{q_0}) = (\tau_Q \circ \phi_1^{\Gamma_{(L_{(g,V)},\mathcal{D})}})(v_q),$$

where $\phi_t^{\Gamma_{(L_{(g,V)},\mathcal{D})}}$ is the flow of the nonholonomic mechanical system $(L_{(g,V)}, \mathcal{D})$, is a local diffeomorphism onto its image and so its inverse map is $R^{nh}_{q_0} : \mathcal{M}^{nh}_{q_0} \to \mathcal{D}_{q_0}$ given by

$$R^{nh}_{q_0}(\theta, \varphi) = \left(\theta - \theta_0, \varphi - \varphi_0 + \frac{1}{2}\right).$$

The corresponding kinetic nonholonomic system is determined by the Lagrangian function $L_{g_e} : TU_e \to \mathbb{R}$ given by

$$L_{g_e}(q, \dot{q}) = \frac{e - \varphi}{2}(\dot{x}^2 + \dot{y}^2 + \dot{\theta}^2 + \dot{\varphi}^2).$$

After some computations, we may eliminate the Lagrange multipliers appearing in the corresponding Lagrange-d'Alembert equations and find that the trajectories of the nonholonomic system $(L_{g_e}, \mathcal{D}_e)$ must satisfy

$$\begin{cases} \dot{x} & = \dot{\theta}\cos\varphi \\ \dot{y} & = \dot{\theta}\sin\varphi \\ \ddot{\theta} & = \frac{\dot{\varphi}(\dot{\theta} + \sin(\varphi)\dot{y} + \cos(\varphi)\dot{x})}{2e - 2\varphi} \\ \ddot{\varphi} & = \frac{\dot{\varphi}^2 - 2\dot{\theta}^2}{2e - 2\varphi}. \end{cases}$$

Then the trajectories of this system form the exponential map $\exp_{q_0}^{(g_e,\mathcal{D}_e)} : \mathcal{D}_{q_0} \to Q$.

Moreover, using Theorem 5.3.2, we know there is a strictly increasing function $h : J \to I$ satisfying

$$\frac{dh}{ds} = e - V \circ c_{P_q(v_q)}, \quad h(0) = 0.$$

Solving the differential equations, we obtain that

$$h(s) = es + \frac{s^3}{6} - \frac{\omega s^2}{2} - \varphi_0 s.$$

Moreover, by the definition of nonholonomic exponential map we have that

$$\exp_q^{nh,e}(v_q) = c_{v_q}^e(1),$$

where $c_{v_q}^e$ is the trajectory of the kinetic nonholonomic system $(L_{g_e}, \mathcal{D}_e)$. In addition, note that every non-zero vector in $\mathcal{D}$ might be uniquely written in the form

$$v_q = \lambda(v_q)\mathcal{Q}_q(v_q), \quad \lambda(v_q) = \sqrt{\frac{e - V(q)}{2}}\|v_q\|_g.$$

Hence, by the homothetic property of kinetic nonholonomic trajectories we deduce

$$\exp_q^{nh,e}(v_q) = c_{\mathcal{Q}_q(v_q)}^e(\lambda(v_q)).$$

Alternatively, using again Theorem 5.3.2 we may also write

$$\exp_q^{nh,e}(v_q) = c_{\mathcal{P}_q(v_q)}(h^{-1}(\lambda(v_q))).$$

Let (Ω, ω) be coordinates on $\mathcal{D}$ associated to the basis

$$\{\cos\varphi\frac{\partial}{\partial x} + \sin\varphi\frac{\partial}{\partial y} + \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \varphi}\}$$

and (θ, φ) coordinates on $\mathcal{M}_{q_0}^{nh,e} = \mathcal{M}_{q_0}^{nh}$ under which

$$c_{v_q}(t) = \left(\Omega t + \theta_0, \omega t + \varphi_0 - \frac{t^2}{2}\right).$$

Now, $E_{L_{(g,V)}}(\Omega, \omega) = e$ if and only if the initial velocity Ω is equal to

$$\Omega^{\pm} := \pm\sqrt{(e - \varphi_0) - \frac{\omega}{2}},$$

so that

$$c_{P(v_q)}(t) = \left(\Omega^{\pm} t + \theta_0, \omega t + \varphi_0 - \frac{t^2}{2}\right).$$

If $k = h^{-1} \circ \lambda$ then we have that

$$\exp_q^{nh,e}(\Omega^{\pm}, \omega) = \left(\Omega^{\pm} k(\Omega^{\pm}, \omega) + \theta_0, \omega k(\Omega^{\pm}, \omega) + \varphi_0 - \frac{k^2(\Omega^{\pm}, \omega)}{2}\right).$$

Considering the flat metric in $\mathcal{D}$ as a Gauss metric, i.e., the metric

$$\mathcal{G}_0 = d\Omega \otimes d\Omega + d\omega \otimes d\omega$$

then the reparametrization by the function h of the unit energy geodesics (i.e., with velocity 1) with respect to the metric

$$g_q^{nh,e} = \left((\exp_q^{nh,e})^{-1}\right)^* \mathcal{G}_0$$

are just the mechanical nonholonomic trajectories with energy e with initial point q. Therefore, the nonholonomic trajectories are reparametrizations of minimizing geodesics for the Riemannian metric $g_q^{nh,e}$. In particular, they minimize the Riemannian length associated with this metric. △

Chapter 6

Nonholonomic Jacobi fields

In this chapter, we will introduce the notion of a nonholonomic Jacobi field and we will discuss some general classes of examples. Then we will use a technique of lifting the nonholonomic systems to the tangent bundle to find a nonholonomic Jacobi equation satisfied by nonholonomic Jacobi fields, involving the torsion and the curvature of the nonholonomic connection, which is the nonholonomic analogue of the well known Jacobi equation in Riemannian geometry. The chapter is subdivided in two sections: the first dedicated to kinetic nonholonomic systems and the second to mechanical nonholonomic systems.

As it happens with Riemannian geodesics, nonholonomic Jacobi fields measure how much nonholonomic trajectories spread apart or join together. However, as we will see in the following, the *tidal forces* associated to the nonholonomic connection, are much more complex than in Riemannian geometry, in which only curvature is responsible for this effect.

6.1 Definition

First of all, we will introduce the notion of a Jacobi field for a general nonholonomic system as an extension of the definition of a Jacobi field (over a geodesic) for a Riemannian metric.

Definition 6.1.1. Let $(L, \mathcal{D})$ be a nonholonomic system with configuration manifold Q. A vector field $W : I \to TQ$ along a curve $c : I \to Q$ is said to be a *nonholonomic Jacobi field* for the system $(L, \mathcal{D})$ if it is the infinitesimal variation vector field of a family of nonholonomic trajectories of $(L, \mathcal{D})$.

So, according to the definition,

$$W(t) = \frac{\partial}{\partial s}\bigg|_{s=0} (\tau_Q \circ \Phi)(s,t),$$

where

$$\begin{aligned} \Phi : (-\varepsilon, \varepsilon) \times I &\to \mathcal{D} \\ (s,t) &\mapsto \Phi_s(t) \end{aligned}$$

is a smooth map and, for each $s \in (-\varepsilon, \varepsilon)$, $\Phi_s : I \to \mathcal{D}$ is the tangent lift $\dot{c}_s : I \to \mathcal{D}$ of a trajectory $c_s : I \to Q$ of $(L, \mathcal{D})$, with $c_0 = c$.

We remark that, in general, W is not a section of $\mathcal{D}$ (see example 6.2.6 below).

6.2 Nonholonomic Jacobi fields for kinetic systems

In what follows, consider the nonholonomic kinetic systems $(L_g, \mathcal{D})$ with g a Riemannian metric, so that the nonholonomic system is regular. Thus, we may consider the nonholonomic SODE $\Gamma_{(L_g,\mathcal{D})}$ and

$$\Phi_s(t) = \phi_t^{\Gamma_{(L_g,\mathcal{D})}}(v(s)),$$

where $\phi_t^{\Gamma_{(L_g,\mathcal{D})}}$ is the local flow of $\Gamma_{(L_g,\mathcal{D})}$ and $v : (-\varepsilon, \varepsilon) \to \mathcal{D}$ is a smooth curve on $\mathcal{D}$. Therefore, a Jacobi field could be written as

$$W(t) = \frac{\partial}{\partial s}\bigg|_{s=0} \left(\tau_Q \circ \phi_t^{\Gamma_{(L_g,\mathcal{D})}}(v(s))\right).$$

In particular, if the variation has fixed starting point q, then for each s, the curve $c_s(t) = \tau_Q \circ \phi_t^{\Gamma_{(L_g,\mathcal{D})}}(v(s))$ is a radial nonholonomic trajectory departing from q. Hence, it lives on the exact discrete constraint submanifold $\mathcal{M}_q^{nh}$ (see Theorem 5.2.4). Therefore, following Theorem 5.2.4, we deduce that the nonholonomic Jacobi field W associated with this variation must be also a (Riemannian) Jacobi field associated with some Riemannian metric g_q^{nh} satisfying the Gauss condition (see Theorem 5.2.4).

On the other hand, note that the tangent bundle of $\mathcal{M}_q^{nh}$ is generated exactly by nonholonomic Jacobi fields associated with variations by radial nonholonomic trajectories departing from q.

Remark 6.2.1. If the system is unconstrained, that is $\mathcal{D} = TQ$, then it is clear that $W : I \to TQ$ is a Jacobi field for the system (L_g, TQ) if and only if W is a Jacobi field for the Riemannian metric g on Q (see Section 2.1.4 and also, for instance, [O'N83]).

Next, we will present a method that allows us to obtain, under certain conditions, nonholonomic Jacobi fields.

Theorem 6.2.2. *Let $(L_g, \mathcal{D})$ be a purely kinematic nonholonomic system on the manifold Q associated with the Riemannian metric g, $c_v : I \to Q$ a nonholonomic trajectory and let $W \in \mathfrak{X}(Q)$ be a vector field satisfying the following three conditions:*

(i) $[W, \Gamma(\mathcal{D})] \subseteq \Gamma(\mathcal{D})$ *;*

(ii) $\mathcal{L}_W g|_{\Gamma(\mathcal{D})\times\Gamma(\mathcal{D})} = 0$*;*

(iii) $\mathcal{L}_W g|_{[\Gamma(\mathcal{D}),\Gamma(\mathcal{D})]\times\Gamma(\mathcal{D})} = 0$*.*

Then $W \circ c_v : I \longrightarrow TQ$ is a Jacobi field along the nonholonomic trajectory c_v.

Proof. Let us first show that the vector field $W^{\mathbf{c}}|_{\mathcal{D}} \in \mathfrak{X}(\mathcal{D})$, which is clearly equivalent to having its flow $T\phi_t^W$ contained in $\mathcal{D}$. Given $\alpha \in \Gamma(\mathcal{D}^o)$, its associated fiberwise linear function $\widehat{\alpha} \in C^\infty(TQ)$ vanishes on $\mathcal{D}$. In fact,

$$\mathcal{D} = \{v \in TQ \mid \widehat{\alpha}(v) = 0, \ \forall \ \alpha \in \Gamma(\mathcal{D}^o)\}.$$

Therefore, it is enough to show that $W^{\mathbf{c}}(\widehat{\alpha})|_{\mathcal{D}} = 0$, for $\alpha \in \Gamma(\mathcal{D}^o)$. Let $X \in \Gamma(\mathcal{D})$ then

$$W^{\mathbf{c}}(\widehat{\alpha}) \circ X = \widehat{\mathcal{L}_W \alpha} \circ X,$$

using the definition of complete lift (see equation (2.4.2)). Applying now the characterization of the Lie derivative of a one-form we deduce

$$W^{\mathbf{c}}(\widehat{\alpha}) \circ X = W(\alpha(X)) - \alpha([W, X]).$$

The first term vanishes identically since α is a section of the annihilator of $\mathcal{D}$ and X is a section of $\mathcal{D}$ while the second one vanishes since $[W, X]$ is a section of $\mathcal{D}$, by the first hypothesis in the statement of the theorem. Hence, since X was arbitrary, we deduce that $W^{\mathbf{c}}|_{\mathcal{D}} \in \mathfrak{X}(\mathcal{D})$.

Now, assume that the vector fields $W^{\mathbf{c}}|_{\mathcal{D}}$ and $\Gamma_{(L_g,\mathcal{D})}$ commute. Then their flows $T\phi_s^W$ and $\phi_t^{\Gamma_{(L_g,\mathcal{D})}}$, respectively, also commute. Take $v \in \mathcal{D}$ and project the composition of the flows to Q using the bundle projection $\tau_{\mathcal{D}} : \mathcal{D} \to Q$. Then

$$\left(\tau_{\mathcal{D}} \circ T\phi_s^W \circ \phi_t^{\Gamma_{(L_g,\mathcal{D})}}\right)(v) = \left(\tau_{\mathcal{D}} \circ \phi_t^{\Gamma_{(L_g,\mathcal{D})}} \circ T\phi_s^W\right)(v).$$

Since the tangent lift of the flow of W is a vector bundle isomorphism over ϕ_s^W and since the projection $\tau_{\mathcal{D}}(\phi_t^{\Gamma_{L_g,\mathcal{D}}}(v))$ of $\phi_t^{\Gamma_{L_g,\mathcal{D}}}(v)$ is just the trajectory of $\Gamma_{(L_g,\mathcal{D})}$ with initial velocity $v \in \mathcal{D}$, which we denote in general by c_v, we find

$$\left(\phi_s^W \circ \tau_{\mathcal{D}} \circ \phi_t^{\Gamma_{(L_g,\mathcal{D})}}\right)(v) = c_{T\phi_s^W(v)}(t).$$

And applying similar considerations again, the last line reduces to

$$\phi_s^W \circ c_v(t) = c_{T\phi_s^W(v)}(t).$$

This computation proves that the 2-parameter family

$$\Phi : (t,s) \mapsto \phi_s^W \circ c_v(t) \tag{6.2.1}$$

is actually a variation by trajectories of $\Gamma_{(L_g,\mathcal{D})}$. Moreover, its infinitesimal variation vector field is given by

$$\left.\frac{d\Phi}{ds}\right|_{s=0}(t) = \left.\frac{d}{ds}\right|_{s=0} \phi_s^W \circ c_v(t) = W(c_v(t)).$$

Therefore, $W \circ c_v : I \to TQ$ is a Jacobi field along c_v.

So, all we need to show is that $W^{\mathbf{c}}|_{\mathcal{D}}$ and $\Gamma_{(L_g,\mathcal{D})}$ commute. We will prove this result in the next proposition. $\square$

Proposition 6.2.3. *If $(L_g, \mathcal{D})$ is a nonholonomic system on Q and W is a vector field on Q in the same conditions as in Theorem 6.2.2, then we have that*

$$[W^{\mathbf{c}}|_{\mathcal{D}}, \Gamma_{(L_g,\mathcal{D})}] = 0.$$

Proof. We will prove the proposition by computing the action of $[W^{\mathbf{c}}|_{\mathcal{D}}, \Gamma_{L_g,\mathcal{D}}]$ on basic and fiberwise linear functions in $\mathcal{C}^\infty(\mathcal{D})$, which are generated by functions of the type $f \circ \tau_{\mathcal{D}}$ and $\widehat{\alpha}$, with $f \in \mathcal{C}^\infty(Q)$ and $\alpha \in \Gamma(\mathcal{D}^*)$. We have that

$$[W^{\mathbf{c}}|_{\mathcal{D}}, \Gamma_{(L_g,\mathcal{D})}](f \circ \tau_Q) = W^{\mathbf{c}}|_{\mathcal{D}}(\Gamma_{(L_g,\mathcal{D})}(f \circ \tau_Q)) - \Gamma_{(L_g,\mathcal{D})}(W^{\mathbf{c}}|_{\mathcal{D}}(f \circ \tau_Q)).$$

Using again equation (2.4.2) and Theorem 5.1.10, the last line becomes

$$\begin{aligned}[W^{\mathbf{c}}|_{\mathcal{D}}, \Gamma_{(L_g,\mathcal{D})}](f\circ\tau_{\mathcal{D}}) &= W^{\mathbf{c}}|_{\mathcal{D}}(\widehat{df}|_{\mathcal{D}}) - \Gamma_{(L_g,\mathcal{D})}(W(f)\circ\tau_{\mathcal{D}})\\ &= \widehat{\mathcal{L}_W df}|_{\mathcal{D}} - d\widehat{(W(f))}|_{\mathcal{D}} = 0.\end{aligned}$$

On the other hand, the action over functions $\widehat{\alpha}$ with $\alpha\in\Gamma(D^*)$ is given by

$$[W^{\mathbf{c}}|_{\mathcal{D}}, \Gamma_{(L_g,\mathcal{D})}](\widehat{\alpha}) = W^{\mathbf{c}}|_{\mathcal{D}}(\Gamma_{(L_g,\mathcal{D})}(\widehat{\alpha})) - \Gamma_{(L_g,\mathcal{D})}(W^{\mathbf{c}}|_{\mathcal{D}}(\widehat{\alpha})).$$

It is a simple computation to show that on $\mathcal{D}$

$$\widehat{\alpha} = \widehat{P^*\alpha}|_{\mathcal{D}}, \quad (\nabla^{nh}\alpha)^{\mathfrak{q}} = (\nabla^{nh}P^*\alpha)^{\mathfrak{q}}|_{\mathcal{D}} = (\nabla^g P^*\alpha)^{\mathfrak{q}}|_{\mathcal{D}}, \tag{6.2.2}$$

where $P: TQ\to\mathcal{D}$ is the orthogonal projector, ∇^{nh} is the nonholonomic connection and ∇^g is the Levi-Civita connection with respect to g. In the expression above, we extended the notation for fiberwise quadratic functions we introduced before (see equation (2.4.3) in Section 2.4.1). Indeed, given any vector bundle $V\to Q$ and a section T of $V^*\otimes V^*\to Q$, then $T^{\mathfrak{q}}$ is the fiberwise quadratic function on V induced by T.

Hence, from (2.4.2) and Theorem 5.1.10 we have that

$$[W^{\mathbf{c}}|_{\mathcal{D}}, \Gamma_{(L_g,\mathcal{D})}](\widehat{\alpha}) = (\mathcal{L}_W(\nabla^g P^*\alpha))^{\mathfrak{q}}|_{\mathcal{D}} - (\nabla^{nh}\mathcal{L}_W(P^*\alpha)|_{\mathcal{D}})^{\mathfrak{q}}, \tag{6.2.3}$$

where we have also used equation (2.4.4) on the first term of the right-hand side. Both terms appearing above are fiberwise quadratic functions associated to $(0,2)$-tensors. Given $X\in\Gamma(\mathcal{D})$, the first term reduces to

$$\mathcal{L}_W(\nabla^g P^*\alpha)(X,X) = W(\nabla^g_X P^*\alpha(X)) - \nabla^g_{[W,X]}P^*\alpha(X) - \nabla^g_X P^*\alpha([W,X]).$$

Note that there exists a section $Y\in\Gamma(\mathcal{D})$ such that $P^*\alpha = \flat_g(Y)$. So we can rewrite the expression above in terms of the vector field Y. Moreover, using the identity

$$\flat_g(\nabla^g_X Y) = \nabla^g_X \flat_g(Y), \quad X, Y\in\mathfrak{X}(Q), \tag{6.2.4}$$

we get

$$\mathcal{L}_W(\nabla^g P^*\alpha)(X,X) = W(\flat_g(\nabla^g_X Y)(X)) - \flat_g(\nabla^g_{[W,X]}Y)(X) - \flat_g(\nabla^g_X Y)([W,X]).$$

Now we use Lemma 2.1.3 to reduce the previous to

$$\begin{aligned}\mathcal{L}_W(\nabla^g P^*\alpha)(X,X) &= \frac{1}{2}W(\mathcal{L}_Y g(X,X)) - \frac{1}{2}(\mathcal{L}_Y g([W,X],X) \\ &- d(\flat_g(Y))([W,X],X)) - \frac{1}{2}(\mathcal{L}_Y g(X,[W,X]) \\ &- d(\flat_g(Y))(X,[W,X])) \\ &= \frac{1}{2}W(\mathcal{L}_Y g(X,X)) - \mathcal{L}_Y g([W,X],X) \\ &= \frac{1}{2}\mathcal{L}_W(\mathcal{L}_Y g)(X,X).\end{aligned}$$

But, one can prove that for a $(0,2)$-tensor g and any $X,Y,Z,Z'\in\mathfrak{X}(Q)$ we have

$$\mathcal{L}_{[X,Y]}g(Z,Z') = \mathcal{L}_X(\mathcal{L}_Y g)(Z,Z') - \mathcal{L}_Y(\mathcal{L}_X g)(Z,Z'). \tag{6.2.5}$$

Hence, using this fact and Lemma 2.1.3, we conclude that

$$\begin{aligned}\mathcal{L}_W(\nabla^g P^*\alpha)(X,X) &= \frac{1}{2}\mathcal{L}_{[W,Y]}g(X,X) + \frac{1}{2}\mathcal{L}_Y(\mathcal{L}_W g)(X,X) \\ &= g(\nabla^g_X[W,Y],X) + \frac{1}{2}\mathcal{L}_Y(\mathcal{L}_W g)(X,X),\end{aligned}$$

but

$$\frac{1}{2}\mathcal{L}_Y(\mathcal{L}_W g)(X,X) = \frac{1}{2}Y((\mathcal{L}_W g)(X,X)) - (\mathcal{L}_W g)([Y,X],X)$$

vanishes because W satisfies hypothesis (ii) and (iii). Thus,

$$\left(\mathcal{L}_W(\nabla^g P^*\alpha)\right)(X,X) = g(\nabla^g_X[W,Y],X).$$

As for the second term in (6.2.3), we proceed by unwinding the definitions

$$\nabla^{nh}\mathcal{L}_W P^*\alpha(X,X) = \nabla^{nh}_X(\mathcal{L}_W\flat_g(Y))X = X(\mathcal{L}_W\flat_g(Y)(X)) - \mathcal{L}_W\flat_g(Y)(\nabla^{nh}_X X).$$

For any $Z\in\Gamma(\mathcal{D})$ one has that

$$\mathcal{L}_W\flat_g(Y)(Z) = \flat_g([W,Y])(Z) + (\mathcal{L}_W g)(Y,Z) = \flat_g([W,Y])(Z). \tag{6.2.6}$$

Therefore,

$$\begin{aligned}\nabla^{nh}\mathcal{L}_W P^*\alpha(X,X) &= X(\flat_g([W,Y])(X)) - \flat_g([W,Y])(\nabla^{nh}_X X) \\ &= X(g([W,Y],X)) - g([W,Y],\nabla^{nh}_X X) \\ &= X(g([W,Y],X)) - g([W,Y],P\nabla^g_X X).\end{aligned}$$

So, using that $[W,Y] \in \Gamma(\mathcal{D})$, it follows that

$$\begin{aligned}\nabla^{nh}\mathcal{L}_W P^*\alpha(X,X) &= X(g([W,Y],X)) - g([W,Y],\nabla^g_X X)\\ &= g(\nabla^g_X[W,Y],X).\end{aligned}$$

Hence both terms in equation (6.2.3) cancel and $[W^{\mathbf{c}}|_{\mathcal{D}}, \Gamma_{L_g,\mathcal{D}}](\widehat{\alpha}) = 0$. □

From Theorem 6.2.2, it follows that

Corollary 6.2.4. *Let $(L_g,\mathcal{D})$ be a purely kinematic nonholonomic system on the manifold and $c_v : I \to Q$ a nonholonomic trajectory with initial velocity $v \in \mathcal{D}$. If W is an infinitesimal symmetry of the system $(L_g,\mathcal{D})$, that is, W is a Killing vector field for the Riemannian metric g (i.e., $\mathcal{L}_W g = 0$) and an infinitesimal symmetry of $\mathcal{D}$ (that is $[W,\Gamma(\mathcal{D})] \subseteq \Gamma(\mathcal{D})$) then $W \circ c_v : I \to TQ$ is a nonholonomic Jacobi field for the system $(L_g,\mathcal{D})$.*

Remark 6.2.5. If the system $(L_g,\mathcal{D})$ is unconstrained (that is, $\mathcal{D} = TQ$), then using Corollary 6.2.4, we recover a well-known result in Riemannian geometry (see, for example, Lemma 26, Chapter 9 in [O'N83]): the restriction of a Killing vector field to a geodesic is a Jacobi field for the Riemannian metric.

Example 6.2.6. We show, by applying the previous corollary, that the vector field $W = \frac{\partial}{\partial z}$ is a Jacobi field for the nonholonomic particle, along any nonholonomic solution.

It is clear that the first condition in the theorem is satisfied, since the vector field $\frac{\partial}{\partial z}$ commutes with the vector fields $e_1 = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}$ and $e_2 = \frac{\partial}{\partial y}$ generating the module of sections $\Gamma(\mathcal{D})$.

On the other hand, $\frac{\partial}{\partial z}$ is a Killing vector field for the euclidean metric g on $\mathbb{R}^3$, so it satisfies the hypothesis in Corollary 6.2.4.

Therefore, by Corollary 6.2.4, the vector field $\frac{\partial}{\partial z}$ is a Jacobi field along any trajectory of the nonholonomic system $(L_g,\mathcal{D})$. However, it is clear that $\frac{\partial}{\partial z}$ is not a section of $\mathcal{D}$. △

Example 6.2.7. A more physical example is the vertical rolling disk, which models the motion of a rolling penny on a plane. It is a nonholonomic system with a Lagrangian function of kinetic type given by $L_g : T(\mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1) \to \mathbb{R}$, with

$$L_g(x,y,\theta,\varphi,\dot{x},\dot{y},\dot{\theta},\dot{\varphi}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + I\dot{\theta}^2 + J\dot{\varphi}^2),$$

where I and J are real numbers known as moment of inertia and nonholonomic constraints imposed by the equations

$$\dot{x} = R\dot{\theta}\cos(\varphi), \quad \dot{y} = R\dot{\theta}\sin(\varphi),$$

where R is the radius of the disk (for more details see [Blo15]). Now, it is easy to see that the constraints form a distribution $\mathcal{D}$ with $\operatorname{rank}(\mathcal{D}) = 2$ generated by the vector fields

$$e_1 = R\cos(\varphi)\frac{\partial}{\partial x} + R\sin(\varphi)\frac{\partial}{\partial y} + \frac{\partial}{\partial \theta}, \quad e_2 = \frac{\partial}{\partial \varphi}.$$

It is easy to see that the vector field $W = \frac{\partial}{\partial \theta}$ is an infinitesimal symmetry of $\mathcal{D}$. Moreover, W is a Killing vector field for the Riemannian metric g on $\mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$ associated to the Lagrangian L_g. So, W is a nonholonomic Jacobi field along any nonholonomic trajectory. △

Example 6.2.8. We will consider again the nonholonomic particle. However, now we will obtain an example of a Jacobi field which is not a Killing vector field for g and another one of a Jacobi field which is not a symmetry of the distribution.

Let $c_{v(s)} : I \to \mathbb{R}^3$ be a trajectory of the nonholonomic particle with $c_v(0) = (x_0, y_0, z_0)$ and initial velocity $v(s) = (\dot{x}_0(s), \dot{y}_0(s), y_0\dot{x}_0(s))$ for each $s \in (-\varepsilon, \varepsilon)$.

On one hand, suppose that $\dot{y}_0(s) \equiv 0$ and so the trajectory has the local expression given by equation (4.3.3).

The curve $W : I \to TQ$ defined by

$$W(t) = \left.\frac{d}{ds}\right|_{s=0} c_{v(s)} = ut \cdot \left(\frac{\partial}{\partial x} + y_0\frac{\partial}{\partial z}\right)$$

is a Jacobi field along $c_{v(0)}$ by definition, where u denotes $\frac{d}{ds}|_{s=0}\dot{x}_0(s)$. Supposing that $\dot{x}_0(0)$ is not zero then the vector field $\widetilde{W} \in \mathfrak{X}(\mathbb{R}^3)$ defined by

$$\widetilde{W}(x, y, z) = u \cdot \left(\frac{x - x_0}{\dot{x}_0(0)}\right) \cdot \left(\frac{\partial}{\partial x} + y\frac{\partial}{\partial z}\right)$$

extends $W(t)$ over the curve $c_{v(0)}$, that is,

$$W(t) = (\widetilde{W} \circ c_{v(0)})(t).$$

However, as it is clear, $\widetilde{W}$ is not a symmetry of the distribution.

On the other hand, suppose that $\dot{y}_0(s)$ does not vanish. Then the local expression of the trajectory is given by equation (4.3.2).

Suppose that

$$\left.\frac{d}{ds}\right|_{s=0} \dot{x}_0(s) = u \quad \text{and} \quad \left.\frac{d}{ds}\right|_{s=0} \dot{y}_0(s) = 0.$$

Then the vector field $W : I \to TQ$ defined as before is a Jacobi field over the trajectory $c_{v(0)}$ and has the local expression

$$\begin{aligned} W(t) = \frac{u}{\dot{y}_0(0)} \sqrt{y_0^2 + 1} \cdot \Bigg[&(\operatorname{arcsinh}(\dot{y}_0(0)t + y_0) - \operatorname{arcsinh}(y_0)) \frac{\partial}{\partial x} \\ &+ \left(\sqrt{(\dot{y}_0(0)t + y_0)^2 + 1} - \sqrt{y_0^2 + 1} \right) \frac{\partial}{\partial z} \Bigg]. \end{aligned}$$

Following the same construction as before, supposing that $\dot{x}_0(0)$ does not vanish, then the vector field

$$\widetilde{W}(x, y, z) = \frac{u}{\dot{x}_0(0)} \left((x - x_0) \frac{\partial}{\partial x} + (z - z_0) \frac{\partial}{\partial z} \right)$$

extends $W(t)$, in the same sense than before. However, it is easy to check that

$$(\mathcal{L}_{\widetilde{W}} g) \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial x} \right) = \frac{2u}{\dot{x}_0(0)}, \tag{6.2.7}$$

hence $\widetilde{W}$ is not a Killing vector field for g. △

Example 6.2.9. Let us find a similar counterexample for the vertical rolling disk dynamics.

Let $c_{v(s)} : I \to \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$ be a trajectory of the vertical rolling disk with $c_v(0) = (x_0, y_0, \theta_0, \varphi_0)$ and initial velocity in $\mathcal{D}$ given by $v(s) = (\dot{x}_0(s), \dot{y}_0(s), \Omega(s), \omega(s))$ for each $s \in (-\varepsilon, \varepsilon)$.

The explicit solution of the nonholonomic dynamics is discussed in Example 4.3.5 (see also [Blo15]), where we find that

$$\begin{cases} \theta_s(t) = \Omega(s)t + \theta_0 \\ \varphi_s(t) = \omega(s)t + \varphi_0 \end{cases} \tag{6.2.8}$$

and the expression for the variables x and y is determined by integrating the constraints.

Suppose that $\omega(s) \equiv 0$, in which case the trajectory is given by the local expressions (6.2.8) and

$$\begin{cases} x_s(t) = \Omega(s) t R \cos(\varphi_0) + x_0 \\ y_s(t) = \Omega(s) t R \sin(\varphi_0) + y_0. \end{cases}$$

Now, let

$$\left.\frac{d}{ds}\right|_{s=0} \Omega(s) = u \quad \text{and} \quad \Omega(0) = \Omega_0,$$

with Ω_0 different from zero. Then the vector field $W : I \to TQ$ obtained by

$$W(t) = \left.\frac{d}{ds}\right|_{s=0} c_{v(s)} = ut \cdot \left(R\cos(\varphi_0)\frac{\partial}{\partial x} + R\sin(\varphi_0)\frac{\partial}{\partial y} + \frac{\partial}{\partial \theta} \right)$$

is a Jacobi field along $c_{v(0)}$ by definition.

Moreover, the vector field $\widetilde{W} \in \mathfrak{X}(\mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1)$ defined by

$$\widetilde{W}(x, y, \theta, \varphi) = u \cdot \left(\frac{\theta - \theta_0}{\Omega_0} \right) \cdot \left(R\cos(\varphi_0)\frac{\partial}{\partial x} + R\sin(\varphi_0)\frac{\partial}{\partial y} + \frac{\partial}{\partial \theta} \right)$$

extends $W(t)$ over the curve $c_{v(0)}$, that is,

$$W(t) = (\widetilde{W} \circ c_{v(0)})(t).$$

However it is easy to see that $\widetilde{W}$ is not an infinitesimal symmetry of the distribution and it is not a Killing vector field with respect to the metric g. △

6.2.1 The lift of the kinematic nonholonomic system and the nonholonomic Jacobi fields

Before starting this section, we recommend the reader to take a look into Section A.1 in Appendix A, to recall the definition and the main features of the complete lift of a kinetic Lagrangian system.

Denote by $g^{\mathbf{c}}$ the complete lift of the Riemannian metric g (see equations (A.1.1) in Appendix A). Then $g^{\mathbf{c}}$ is a pseudo-Riemannian metric on TQ and we may consider the Lagrangian function $L_{g^{\mathbf{c}}} : TTQ \to \mathbb{R}$ associated with $g^{\mathbf{c}}$. We recall that (see Lemma A.1.1)

$$L_{g^{\mathbf{c}}} = L_g^{\mathbf{c}} \circ \kappa_Q,$$

where $L_g^{\mathbf{c}}$ is the complete lift of the Lagrangian function L_g and $\kappa_Q : TTQ \to TTQ$ is the canonical involution of the double tangent bundle TTQ (see Section 2.4.1 to recall the definition of canonical involution).

Now, consider the complete lift $\mathcal{D}^{\mathbf{c}}$ of the distribution $\mathcal{D}$ as a distribution on TQ, whose space of sections is

$$\Gamma(\mathcal{D}^{\mathbf{c}}) = \langle \{X^{\mathbf{c}}, X^{\mathbf{v}} |\ X \in \Gamma(\mathcal{D})\} \rangle.$$

Here, $X^{\mathbf{c}}$ and $X^{\mathbf{v}}$ are the complete and vertical lifts of the vector field $X \in \Gamma(\mathcal{D})$. The distribution $\mathcal{D}^{\mathbf{c}}$ was considered in [YI73].

$\mathcal{D}^{\mathbf{c}}$ is not only a vector subbundle (over TQ) of the vector bundle $\tau_{TQ} : TTQ \to TQ$ but also a vector bundle over $\mathcal{D}$ with vector bundle projection $(T\tau_Q)|_{\mathcal{D}^{\mathbf{c}}} : \mathcal{D}^{\mathbf{c}} \to \mathcal{D}$. In fact, if $X \in \Gamma(\mathcal{D})$ then

$$(T\tau_Q)(X^{\mathbf{c}}) = X \circ \tau_Q, \quad (T\tau_Q)(X^{\mathbf{v}}) = 0 \circ \tau_Q,$$

where $0 : Q \to TQ$ is the zero section.

On the other hand, the tangent bundle $T\mathcal{D}$ to $\mathcal{D}$ is also a double vector bundle. Indeed, besides the canonical vector bundle structure $\tau_{\mathcal{D}} : T\mathcal{D} \to \mathcal{D}$, there is also a vector bundle structure over TQ with vector bundle projection $T(\tau_Q|_{\mathcal{D}}) : T\mathcal{D} \to TQ$.

In addition, using that κ_Q is an involution from the vector bundle $\tau_{TQ} : TTQ \to TQ$ to the vector bundle $T\tau_Q : TTQ \to TQ$ (see Section 2.4.1), it follows that the restriction of κ_Q to $\mathcal{D}^{\mathbf{c}} \subseteq TTQ$ is also an isomorphism between the vector bundle $\tau_{\mathcal{D}^{\mathbf{c}}} : \mathcal{D}^{\mathbf{c}} \to TQ$ and $T(\tau_Q|_{\mathcal{D}}) : T\mathcal{D} \to TQ$ (respectively, between $(T\tau_Q)|_{\mathcal{D}^{\mathbf{c}}} : \mathcal{D}^{\mathbf{c}} \to \mathcal{D}$ and $\tau_{\mathcal{D}} : T\mathcal{D} \to \mathcal{D}$) over the identity of TQ (respectively, over the identity of $\mathcal{D}$). The diagram in Figure 6.1 illustrates the situation. Note that the inverse morphism of this double vector bundle isomorphism is $(\kappa_Q)|_{T\mathcal{D}} : T\mathcal{D} \to \mathcal{D}^{\mathbf{c}}$.

Definition 6.2.10. The nonholonomic system $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$ is the *complete lift* of the nonholonomic system of kinetic type $(L_g, \mathcal{D})$.

The aim of this section is to prove the following theorem:

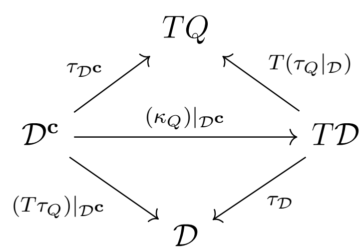


Figure 6.1: Commutative diagram showing how the restriction of the canonical involution to $\mathcal{D}^{\mathbf{c}}$ commutes with the projections to $\mathcal{D}$ and TQ.

Theorem 6.2.11. *Let $(L_g, \mathcal{D})$ be a nonholonomic system of kinetic type and $\Gamma_{(L_g,\mathcal{D})}$ the associated nonholonomic SODE. Then*

(i) The complete lift $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$ is a regular nonholonomic system.

(ii) Let $\Gamma_{(L_{g^{\mathbf{c}}},\mathcal{D}^{\mathbf{c}})} \in \mathfrak{X}(\mathcal{D}^{\mathbf{c}})$ be the nonholonomic SODE associated with the system $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$ and $\kappa_Q : TTQ \to TTQ$ the canonical involution. Then

$$\Gamma_{(L_{g^{\mathbf{c}}},\mathcal{D}^{\mathbf{c}})} = T\kappa_Q|_{T\mathcal{D}} \circ \Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}} \tag{6.2.9}$$

and so we have

(a) $\Gamma_{(L_{g^{\mathbf{c}}},\mathcal{D}^{\mathbf{c}})}$ is $T\tau_Q|_{\mathcal{D}^{\mathbf{c}}}$-projectable over $\Gamma_{(L_g,\mathcal{D})}$;

(b) The trajectories of $\Gamma_{(L_{g^{\mathbf{c}}},\mathcal{D}^{\mathbf{c}})}$ are just the Jacobi fields for the nonholonomic system $(L_g, \mathcal{D})$.

If $c_v : I \to Q$ is a trajectory of the nonholonomic dynamics and $W : I \to TQ$ is a vector field on Q along c_v then an immediate corollary of this theorem is that

Corollary 6.2.12. *W is a Jacobi field for the nonholonomic system $(L_g, \mathcal{D})$ if and only if*

1. *$\dot{W}(t) \in \mathcal{D}^{\mathbf{c}}_{W(t)}$, for every $t \in I$;*

2. *$i_{\ddot{W}}\omega_{L_{g^{\mathbf{c}}}}(\dot{W}) - dL_{g^{\mathbf{c}}}(\dot{W}) \in ((\mathcal{D}^{\mathbf{c}})^o)^{\boldsymbol{v}}$,*

where $\omega_{L_{g^{\mathbf{c}}}}$ is the Poincaré-Cartan 2-form associated with the Lagrangian function $L_{g^{\mathbf{c}}}$.

First we show that the complete lift $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$ on TQ obtained from the nonholonomic system of kinetic type $(L_g, \mathcal{D})$ on Q is always regular.

Proposition 6.2.13. *If L_g is the Lagrangian function associated to the Riemannian metric g, then the nonholonomic system $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$ is regular.*

Proof. Let $Z \in \mathcal{D}^{\mathbf{c}} \cap (\mathcal{D}^{\mathbf{c}})^{\perp}$. Let $\{X^a\}$ be an orthonormal basis of sections on $\mathcal{D}$. The set $\{(X^a)^{\mathbf{v}}, (X^a)^{\mathbf{c}}\}$ is then a basis of sections of $\mathcal{D}^{\mathbf{c}}$ and Z may be written as

$$Z = \lambda_a (X^a)^{\mathbf{c}} + \mu_a (X^a)^{\mathbf{v}}.$$

Since Z is in the intersection of $\mathcal{D}^{\mathbf{c}}$ with its $g^{\mathbf{c}}-$orthogonal distribution then, using (A.1.1) in Appendix A, we have that for every $Y \in \Gamma(D^{\mathbf{c}})$ expressed as $Y = f_a (X^a)^{\mathbf{c}} + g_a (X^a)^{\mathbf{v}}$ in the same basis,

$$\begin{aligned} 0 &= g^{\mathbf{c}}(Z, Y) \\ &= \lambda_a f_b (g(X^a, X^b))^{\mathbf{c}} + (\lambda_a g_b + \mu_a f_b)(g(X^a, X^b))^{\mathbf{v}} \\ &= \lambda_a g_a + \mu_a f_a, \end{aligned}$$

since we are taking an orthonormal basis of $\mathcal{D}$. Since the functions f_a and g_a are arbitrary, we deduce that $\lambda_a = \mu_a = 0$, hence, $Z = 0$. Therefore, by Theorem 5.1.7 the nonholonomic system $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$ is regular. □

The last proposition proves item (i) in Theorem 6.2.11. Therefore, from now on we can refer to the nonholonomic SODE $\Gamma_{(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})}$ associated with the complete lifted nonholonomic system $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$.

In order to prove item (ii) in Theorem 6.2.11 we will characterize further the distribution $\mathcal{D}^{\mathbf{c}}$. Our main purpose is to identify a local basis of the vector subbundle $((\mathcal{D}^{\mathbf{c}})^o)^{\mathbf{v}} \to TQ$ of $T^*TQ \to TQ$.

If μ is a 1-form on Q, we will denote by $\mu^{\mathbf{c}} \in \Omega^1(TQ)$ and $\mu^{\mathbf{v}} \in \Omega^1(TQ)$ the complete and vertical lifts, respectively, of μ to TQ (see (2.4.7) and (2.4.10)). Moreover, as in the case of TQ, if $V \to Q$ is a vector subbundle of $\pi_Q : T^*Q \to Q$, then we can define the complete lift $V^{\mathbf{c}}$ of V as a vector subbundle of $\pi_{TQ} : T^*(TQ) \to TQ$ over TQ which is characterized by the following condition

$$\Gamma(V^{\mathbf{c}}) = \langle \{\alpha^{\mathbf{v}}, \alpha^{\mathbf{c}} \mid \alpha \in \Gamma(V)\} \rangle \tag{6.2.10}$$

Lemma 6.2.14. *1. We have that*

$$\Gamma((\mathcal{D}^{\mathbf{c}})^o) = \langle \{\mu^{\mathbf{c}}, \mu^{\mathbf{v}} | \mu \in \Gamma(\mathcal{D}^o)\} \rangle \quad \textit{and} \quad \Gamma((\mathcal{D}^o)^{\mathbf{v}}) = \langle \{\mu^{\mathbf{v}} | \mu \in \Gamma(\mathcal{D}^o)\} \rangle.$$

2. *Moreover,*

$$\Gamma(((\mathcal{D}^o)^{\mathbf{v}})^{\mathbf{c}}) = \langle\{(\mu^{\mathbf{v}})^{\mathbf{c}}, (\mu^{\mathbf{v}})^{\mathbf{v}} | \mu \in \Gamma(\mathcal{D}^o)\}\rangle$$
$$\Gamma(\kappa_Q^*(((\mathcal{D}^o)^{\mathbf{v}})^{\mathbf{c}})) = \langle\{(\mu^{\mathbf{c}})^{\mathbf{v}}, (\mu^{\mathbf{v}})^{\mathbf{v}} | \mu \in \Gamma(\mathcal{D}^o)\}\rangle = \Gamma(((\mathcal{D}^{\mathbf{c}})^o)^{\mathbf{v}}).$$

Proof. 1. For every $X \in \Gamma(\mathcal{D})$ and $\mu \in \Gamma(\mathcal{D}^o)$ we have that $\langle \mu, X\rangle = 0$. Moreover, note that the following identities hold

$$\langle \mu^{\mathbf{c}}, X^{\mathbf{c}}\rangle = (\langle \mu, X\rangle)^{\mathbf{c}} = 0,$$
$$\langle \mu^{\mathbf{c}}, X^{\mathbf{v}}\rangle = (\langle \mu, X\rangle)^{\mathbf{v}} = 0,$$
$$\langle \mu^{\mathbf{v}}, X^{\mathbf{c}}\rangle = (\langle \mu, X\rangle)^{\mathbf{v}} = 0,$$
$$\langle \mu^{\mathbf{v}}, X^{\mathbf{v}}\rangle = 0,$$

hence the elements in $\{\mu^{\mathbf{c}}, \mu^{\mathbf{v}}\}$ annihilate $\Gamma(\mathcal{D}^{\mathbf{c}})$. Therefore, by dimensional reasons they must span $\Gamma((\mathcal{D}^{\mathbf{c}})^o)$.

On the other hand, using (2.4.10), we deduce that

$$\Gamma((\mathcal{D}^o)^{\mathbf{v}}) = \langle\{\mu^{\mathbf{v}} | \mu \in \Gamma(\mathcal{D}^o)\}\rangle.$$

2. This part follows from the previous item, (6.2.10) and the relations in (2.4.18).

□

Proof of Theorem 6.2.11. (i) By Proposition 6.2.13, the complete lift nonholonomic system $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$ is regular.

(ii) Recall that the nonholonomic vector field $\Gamma_{(L_g,\mathcal{D})}$ is defined by the equations

$$\begin{cases} \left(i_{\Gamma_{(L_g,\mathcal{D})}}\omega_{L_g} - dE_{L_g}\right)\Big|_{\mathcal{D}} \in \Gamma((\mathcal{D}^o)^{\mathbf{v}}) \\ \Gamma_{(L_g,\mathcal{D})} \in \mathfrak{X}(\mathcal{D}). \end{cases}$$

Using the complete lift and (2.4.8) and (2.4.9) in Section 2.4.1, we can obtain the following equation

$$\left(i_{\Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})}}\omega^{\mathbf{c}}_{L_g} - dE^{\mathbf{c}}_{L_g}\right)\Big|_{T\mathcal{D}} \in \Gamma(((\mathcal{D}^o)^{\mathbf{v}})^{\mathbf{c}})$$

If we pullback the previous equation by κ_Q and use Lemma 6.2.14, we deduce that

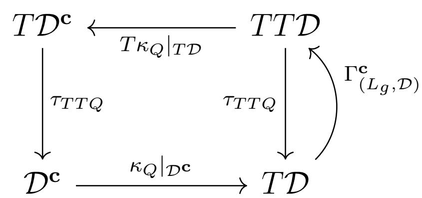
$$\left(i_{(\kappa_Q)_*\Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})}}\kappa_Q^*\omega^{\mathbf{c}}_{L_g} - d(\kappa_Q^* E^{\mathbf{c}}_{L_g})\right)\Big|_{\mathcal{D}^{\mathbf{c}}} \in \Gamma(\kappa_Q^*(((\mathcal{D}^o)^{\mathbf{v}})^{\mathbf{c}})) = \Gamma(((\mathcal{D}^{\mathbf{c}})^o)^{\mathbf{v}})$$

Applying Proposition A.1.2 in Appendix A, the equation reduces to

$$\left(i_{(\kappa_Q)_*\Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})}}\omega_{L_{g^{\mathbf{c}}}} - d(E_{L_{g^{\mathbf{c}}}})\right)\Big|_{\mathcal{D}^{\mathbf{c}}} \in \Gamma(((\mathcal{D}^{\mathbf{c}})^o)^{\mathbf{v}}).$$

Notice that since $\Gamma_{(L_g,\mathcal{D})}$ is a vector field in the submanifold $\mathcal{D}$, its complete lift satisfies $\Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})} \in \mathfrak{X}(T\mathcal{D})$.

Therefore we may form the commutative diagram below



Hence, $(\kappa_Q)_*\Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})} = T\kappa_Q|_{T\mathcal{D}} \circ \Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}}$ is a vector field on $\mathcal{D}^{\mathbf{c}}$. Moreover, since the nonholonomic system $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$ is regular, by uniqueness of nonholonomic vector field, it coincides with $\Gamma_{(L_{g^{\mathbf{c}}},\mathcal{D}^{\mathbf{c}})}$, i.e.,

$$\Gamma_{(L_{g^{\mathbf{c}}},\mathcal{D}^{\mathbf{c}})} = T\kappa_Q|_{T\mathcal{D}} \circ \Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}}. \tag{6.2.11}$$

Then the statements in item (ii) are just consequences of the properties of the complete lift and the canonical involution. Indeed,

$$\begin{aligned} T(T\tau_Q|_{\mathcal{D}^{\mathbf{c}}})(\Gamma_{(L_{g^{\mathbf{c}}},\mathcal{D}^{\mathbf{c}})}) &= T(T\tau_Q|_{\mathcal{D}^{\mathbf{c}}} \circ \kappa_Q|_{T\mathcal{D}}) \circ \Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}} \\ &= T(\tau_{TQ}|_{T\mathcal{D}})(\Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}}) \\ &= \Gamma_{(L_g,\mathcal{D})} \circ \tau_{TQ}|_{T\mathcal{D}} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}} \\ &= \Gamma_{(L_g,\mathcal{D})} \circ (T\tau_Q|_{\mathcal{D}^{\mathbf{c}}}), \end{aligned}$$

where we have used that $\tau_{TQ}|_{T\mathcal{D}} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}} = T\tau_Q|_{\mathcal{D}^{\mathbf{c}}}$. This proves the first statement.

The second statement in item (ii), may be seen from the fact that if $W : I \to TQ$ is a trajectory of $\Gamma_{(L_{g^{\mathbf{c}}},\mathcal{D}^{\mathbf{c}})}$, then its tangent lift $\dot{W} : I \to \mathcal{D}^{\mathbf{c}}$ is an integral curve of $\Gamma_{(L_{g^{\mathbf{c}}},\mathcal{D}^{\mathbf{c}})}$ and, thus, $\kappa_Q \circ \dot{W} : I \to T\mathcal{D}$ is an integral curve of $\Gamma^{\mathbf{c}}_{(L_g,\mathcal{D})}$. Therefore we may write it as

$$\kappa_Q \circ \dot{W}(t) = \left(T_{W(0)}\phi_t^{\Gamma_{(L_g,\mathcal{D})}}\right)(\kappa_Q \circ \dot{W}(0)).$$

So,

$$W(t) = T\tau_Q(\kappa_Q(\dot{W}(t))) = T\tau_Q\left(\left(T_{W(0)}\phi_t^{\Gamma_{(L_g,\mathcal{D})}}\right)(\kappa_Q \circ \dot{W}(0))\right)$$

and

$$W(t) = \left(T_{W(0)}(\tau_Q \circ \phi_t^{\Gamma_{(L_g,\mathcal{D})}})\right)(\kappa_Q \circ \dot{W}(0)).$$

Let now $v : J \subseteq \mathbb{R} \to \mathcal{D}$ be a curve such that its initial velocity is $v'(0) = \kappa_Q \circ \dot{W}(0)$, where v' denotes differentiation with respect to the variable $s \in J$. Then

$$W(t) = \left.\frac{d}{ds}\right|_{s=0}\left(\tau_Q \circ \phi_t^{\Gamma_{(L_g,\mathcal{D})}}\right)(v(s)).$$

Hence, W is a nonholonomic Jacobi field for $\Gamma_{(L_g,\mathcal{D})}$, since it is an infinitesimal variation of nonholonomic trajectories of $\Gamma_{(L_g,\mathcal{D})}$.

□

Remark 6.2.15. As a consequence of the last theorem if $W : I \longrightarrow TQ$ is a Jacobi field for the nonholonomic dynamics $(L_g, \mathcal{D})$ it must satisfy the constraint:

$$\dot{W}(t) \in D^{\mathbf{c}}_{W(t)}, \quad \text{for every } t \in I.$$

Example 6.2.16. Let us check that the lifted nonholonomic system obtained from the nonholonomic particle is regular.

By Theorem 5.1.7 it is enough to check that $\mathcal{D}^{\mathbf{c}} \cap (\mathcal{D}^{\mathbf{c}})^{\perp} = \{0\}$. This is equivalent to show that the matrix C^{ab} defined in (3.6.7) is non-singular at points of $\mathcal{D}^{\mathbf{c}}$. If we were to compute this matrix we would find it was

$$\begin{pmatrix} 0 & y^2+1 \\ y^2+1 & 2vy \end{pmatrix}$$

which is clearly non-singular.

In this example the constraint distribution $\mathcal{D}$ is generated by the vectors $e_1 = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}$ and $e_2 = \frac{\partial}{\partial y}$. The orthogonal distribution $\mathcal{D}^{\perp}$ for the euclidean metric is generated by $e_3 = y\frac{\partial}{\partial x} - \frac{\partial}{\partial z}$.

The lifted distribution $\mathcal{D}^{\mathbf{c}}$, by definition, is generated by the vectors $e_1^{\mathbf{c}}, e_2^{\mathbf{c}}, e_1^{\mathbf{v}}, e_2^{\mathbf{v}}$. The set $\{e_3^{\mathbf{c}}, e_3^{\mathbf{v}}\}$ is linearly independent and it is easily proven to be $g^{\mathbf{c}}-$orthogonal to $\mathcal{D}^{\mathbf{c}}$, hence, by dimensional reasons, it generates the orthogonal distribution $(\mathcal{D}^{\mathbf{c}})^{\perp}$.

Moreover, since $\{e_1^{\mathbf{c}}, e_2^{\mathbf{c}}, e_1^{\mathbf{v}}, e_2^{\mathbf{v}}, e_3^{\mathbf{c}}, e_3^{\mathbf{v}}\}$ is a basis of sections of $\mathfrak{X}(TQ)$, the intersection of $\mathcal{D}^{\mathbf{c}}$ and $(\mathcal{D}^{\mathbf{c}})^{\perp}$ must be zero. △

6.2.2 Nonholonomic Jacobi equation

Theorem 5.1.8 asserts that if $c_v : I \to Q$ is a trajectory of $\Gamma_{(L_g,\mathcal{D})}$, then

$$\nabla^{nh}_{\dot{c}_v}\dot{c}_v = 0 \quad \text{and} \quad \dot{c}_v(t) \in \mathcal{D}_{c_v(t)}, \quad \text{for every } t \in I.$$

Consider the complete lift of the metric g denoted by $g^{\mathbf{c}}$, which is a symmetric non-degenerate $(0,2)$-tensor on TQ. The kinetic Lagrangian $L_{g^{\mathbf{c}}}$ associated to $g^{\mathbf{c}}$ satisfies $L_g^{\mathbf{c}} \circ \kappa_Q = L_{g^{\mathbf{c}}}$ (see Lemma A.1.1 in Appendix A). Moreover, from Theorem 6.2.11, we have that $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$ is a regular nonholonomic system.

Since the Lagrangian function $L_{g^{\mathbf{c}}}$ is kinetic, its trajectories are geodesics of the nonholonomic connection ∇^{NH} defined by

$$\nabla^{NH}_X Y := P^T(\nabla^{g^{\mathbf{c}}}_X Y) + \nabla^{g^{\mathbf{c}}}_X[P'^T(Y)], \quad \text{for } X, Y \in \mathfrak{X}(TQ), \tag{6.2.12}$$

where $\nabla^{g^{\mathbf{c}}}$ is the Levi-Civita connection of $g^{\mathbf{c}}$, $P^T : TTQ \to \mathcal{D}^{\mathbf{c}}$ is the associated orthogonal projector onto the distribution $\mathcal{D}^{\mathbf{c}}$ and $P'^T : TTQ \to (\mathcal{D}^{\mathbf{c}})^\perp$ is the orthogonal projector onto $(\mathcal{D}^{\mathbf{c}})^\perp$, the orthogonal distribution with respect to $g^{\mathbf{c}}$.

Lemma 6.2.17. *The following identities are satisfied:*

1. $\nabla^{g^{\mathbf{c}}} = (\nabla^g)^{\mathbf{c}}$;

2. $\kappa_Q \circ TP \circ \kappa_Q(X^{\mathbf{c}}) = (P(X))^{\mathbf{c}}$, *for any* $X \in \mathfrak{X}(Q)$;

3. $\kappa_Q \circ TP \circ \kappa_Q(X^{\mathbf{v}}) = (P(X))^{\mathbf{v}}$, *for any* $X \in \mathfrak{X}(Q)$;

4. $P^T = \kappa_Q \circ TP \circ \kappa_Q$;

5. $P'^T = \kappa_Q \circ TP' \circ \kappa_Q$.

Proof. The first item is proved in Corollary 2.6.6. in [LR89]. To prove item 2, just use the properties of the canonical involution in Section 2.4.1 (see (2.4.16) in Section 2.4.1)

$$\kappa_Q \circ TP \circ \kappa_Q(X^{\mathbf{c}}) = \kappa_Q \circ TP(TX) = \kappa_Q(T(P \circ X)) = (P(X))^{\mathbf{c}}.$$

We may prove item 3 in a similar way. Given $u_q \in T_qQ$, we have

$$\begin{aligned}\kappa_Q \circ TP \circ \kappa_Q(X^{\mathbf{v}})(u_q) &= (\kappa_Q \circ TP)(T_q0(u_q) + (X(q))^{\mathbf{v}}_{0_q}) \\ &= (\kappa_Q \circ TP)(T_q0(u_q) + \left.\frac{d}{dt}\right|_{t=0}(tX(q))) \\ &= \kappa_Q(T_q0(u_q) + \left.\frac{d}{dt}\right|_{t=0}(tP(X(q)))) \\ &= (PX)^{\mathbf{v}}(u_q).\end{aligned}$$

As a consequence of the two previous items we have that

$$\begin{aligned}&\kappa_Q \circ TP \circ \kappa_Q(X^{\mathbf{c}}) = X^{\mathbf{c}}, \quad \kappa_Q \circ TP \circ \kappa_Q(X^{\mathbf{v}}) = X^{\mathbf{v}}, \quad X \in \Gamma(\mathcal{D}) \\ &\kappa_Q \circ TP \circ \kappa_Q(Y^{\mathbf{c}}) = 0, \quad \kappa_Q \circ TP \circ \kappa_Q(Y^{\mathbf{v}}) = 0, \quad Y \in \Gamma(\mathcal{D}^{\perp}).\end{aligned}$$

Note that while $\{X^{\mathbf{c}}, X^{\mathbf{v}} | X \in \Gamma(\mathcal{D})\}$ spans $\Gamma(\mathcal{D}^{\mathbf{c}})$, the set $\{Y^{\mathbf{c}}, Y^{\mathbf{v}} | Y \in \Gamma(\mathcal{D}^{\perp})\}$ spans $\Gamma((\mathcal{D}^{\mathbf{c}})^{\perp})$, where the orthogonal is taken with respect to the pseudo-Riemannian metric $g^{\mathbf{c}}$. Hence, $\kappa_Q \circ TP \circ \kappa_Q$ is the identity on $\mathcal{D}^{\mathbf{c}}$ and vanishes on $(\mathcal{D}^{\mathbf{c}})^{\perp}$. Therefore, it must be the orthogonal projector P^T.

The argument to prove item 5. is completely analogous, just substitute P by P'. □

The last Lemma simplifies the proof of the next Proposition, relating both nonholonomic connections by the complete lift. Before, the statement let us recall some properties of the complete lift of a linear connection ∇ (see [LR89] or [YI73]):

$$\nabla^{\mathbf{c}}_{X^{\mathbf{c}}} Y^{\mathbf{c}} = (\nabla_X Y)^{\mathbf{c}}, \quad \nabla^{\mathbf{c}}_{X^{\mathbf{c}}} Y^{\mathbf{v}} = \nabla^{\mathbf{c}}_{X^{\mathbf{v}}} Y^{\mathbf{c}} = (\nabla_X Y)^{\mathbf{v}}, \quad \nabla^{\mathbf{c}}_{X^{\mathbf{v}}} Y^{\mathbf{v}} = 0, \tag{6.2.13}$$

for any $X, Y \in \mathfrak{X}(Q)$.

Proposition 6.2.18. *The nonholonomic connection constructed from the Levi-Civita connection associated to $g^{\mathbf{c}}$ and from the projectors P^T, P'^T is the complete lift of the nonholonomic connection constructed from the Levi-Civita for g and from the projector P, and P'. In other words,*

$$\nabla^{NH} = (\nabla^{nh})^{\mathbf{c}}.$$

Proof. We will prove the identity on complete and vertical lifts. Using the definition of ∇^{NH} we get

$$\nabla^{NH}_{X^{\mathbf{c}}} Y^{\mathbf{c}} = P^T (\nabla^g)^{\mathbf{c}}_{X^{\mathbf{c}}} Y^{\mathbf{c}} + (\nabla^g)^{\mathbf{c}}_{X^{\mathbf{c}}} [P'^T (Y^{\mathbf{c}})].$$

Using the properties stated in equations (6.2.13) and in Lemma 6.2.17 we deduce

$$\nabla^{NH}_{X^{\mathbf{c}}} Y^{\mathbf{c}} = P^T (\nabla^g_X Y)^{\mathbf{c}} + (\nabla^g)^{\mathbf{c}}_{X^{\mathbf{c}}} (P' Y)^{\mathbf{c}}.$$

Again applying the same combination of arguments we may reduce the previous line to

$$\nabla^{NH}_{X^{\mathbf{c}}} Y^{\mathbf{c}} = (P \nabla^g_X Y)^{\mathbf{c}} + (\nabla^g_X P' Y)^{\mathbf{c}},$$

which is just the complete lift of ∇^{nh}. So,

$$\nabla^{NH}_{X^{\mathbf{c}}} Y^{\mathbf{c}} = (\nabla^{nh}_X Y)^{\mathbf{c}} = (\nabla^{nh})^{\mathbf{c}}_{X^{\mathbf{c}}} Y^{\mathbf{c}}.$$

The very same arguments are still valid to prove

$$\begin{aligned}
\nabla^{NH}_{X^{\mathbf{c}}} Y^{\mathbf{v}} &= P^T (\nabla^g)^{\mathbf{c}}_{X^{\mathbf{c}}} Y^{\mathbf{v}} + (\nabla^g)^{\mathbf{c}}_{X^{\mathbf{c}}} [P'^T (Y^{\mathbf{v}})] \\
&= P^T (\nabla^g_X Y)^{\mathbf{v}} + (\nabla^g)^{\mathbf{c}}_{X^{\mathbf{c}}} (P' Y)^{\mathbf{v}} \\
&= (P \nabla^g_X Y)^{\mathbf{v}} + (\nabla^g_X P' Y)^{\mathbf{v}} \\
&= (\nabla^{nh}_X Y)^{\mathbf{v}} = (\nabla^{nh})^{\mathbf{c}}_{X^{\mathbf{c}}} Y^{\mathbf{v}},
\end{aligned}$$

and also to prove

$$\begin{aligned}
\nabla^{NH}_{X^{\mathbf{v}}} Y^{\mathbf{v}} &= P^T (\nabla^g)^{\mathbf{c}}_{X^{\mathbf{v}}} Y^{\mathbf{v}} + (\nabla^g)^{\mathbf{c}}_{X^{\mathbf{v}}} [P'^T (Y^{\mathbf{v}})] \\
&= (\nabla^g)^{\mathbf{c}}_{X^{\mathbf{v}}} (P' Y)^{\mathbf{v}} = 0 = (\nabla^{nh})^{\mathbf{c}}_{X^{\mathbf{v}}} Y^{\mathbf{v}}.
\end{aligned}$$

□

Remark 6.2.19. If $W : I \to TQ$ is a trajectory of the nonholonomic system $(L_{g^{\mathbf{c}}}, \mathcal{D}^{\mathbf{c}})$, it is also by Theorem 6.2.11 a Jacobi field for the nonholonomic system $(L_g, \mathcal{D})$, and it is a geodesic for the nonholonomic connection ∇^{NH} by Theorem 5.1.8. Hence, by the last proposition W satisfies

$$(\nabla^{nh})^{\mathbf{c}}_{\dot{W}} \dot{W} = 0.$$

Proposition 6.2.20. *Let $W : I \to TQ$ be a vector field along $c : I \to Q$, a nonholonomic trajectory of $\Gamma_{(L_g, \mathcal{D})}$. Then the coordinate expression of $(\nabla^{nh})^{\mathbf{c}}_{\dot{W}} \dot{W}$ is*

$$(\nabla^{nh})^{\mathbf{c}}_{\dot{W}} \dot{W} = \left(\frac{d^2 W^k}{dt^2} + \dot{q}^i \dot{q}^j W^l \frac{\partial \Gamma^k_{ij}}{\partial q^l} + \dot{q}^j \frac{dW^i}{dt} (\Gamma^k_{ij} + \Gamma^k_{ji}) \right) \frac{\partial}{\partial \dot{q}^k}, \qquad (6.2.14)$$

where (q^i) are local coordinates on Q with respect to which the local expression of W is

$$W(t) = W^i(t)\left.\frac{\partial}{\partial q^i}\right|_{c(t)},$$

$(q^i, \dot{q}^i)$ is the corresponding local expression of $\dot{c}$ on TQ and Γ^k_{ij} are the Christoffel symbols for the nonholonomic connection ∇^{nh}, i.e.,

$$\nabla^{nh}_{\frac{\partial}{\partial q^i}}\frac{\partial}{\partial q^j} = \Gamma^k_{ij}\frac{\partial}{\partial q^k}.$$

Proof. Denote by $\dot{W} : I \to TTQ$ the tangent lift of $W : I \to TQ$. Then we have that

$$\dot{W}(t) = \dot{q}^i(t)\left.\frac{\partial}{\partial q^i}\right|_{W(t)} + \dot{W}^i(t)\left.\frac{\partial}{\partial \dot{q}^i}\right|_{W(t)}.$$

Observe that the coordinate vector fields on TQ denoted by $\frac{\partial}{\partial q^i}$ and $\frac{\partial}{\partial \dot{q}^i}$ are just the complete and the vertical lift of the corresponding coordinate vector field on Q, i.e.,

$$\frac{\partial}{\partial q^i}(v_q) = \left(\frac{\partial}{\partial q^i}\right)^{\mathbf{c}}(v_q), \quad \frac{\partial}{\partial \dot{q}^i}(v_q) = \left(\frac{\partial}{\partial q^i}\right)^{\mathbf{v}}(v_q).$$

With these properties in mind and using (6.2.13), it is easy to prove that,

$$\begin{aligned}
(\nabla^{nh})^{\mathbf{c}}_{\frac{\partial}{\partial q^i}}\frac{\partial}{\partial q^j} &= \Gamma^k_{ij}\frac{\partial}{\partial q^k} + \dot{q}^l\frac{\partial \Gamma^k_{ij}}{\partial q^l}\frac{\partial}{\partial \dot{q}^k}\\
(\nabla^{nh})^{\mathbf{c}}_{\frac{\partial}{\partial \dot{q}^i}}\frac{\partial}{\partial q^j} &= (\nabla^{nh})^{\mathbf{c}}_{\frac{\partial}{\partial q^i}}\frac{\partial}{\partial \dot{q}^j} = \Gamma^k_{ij}\frac{\partial}{\partial \dot{q}^k}\\
(\nabla^{nh})^{\mathbf{c}}_{\frac{\partial}{\partial \dot{q}^i}}\frac{\partial}{\partial \dot{q}^j} &= 0.
\end{aligned}$$

Thus, one can also compute $(\nabla^{nh})^{\mathbf{c}}_{\dot{W}}\dot{W}$ to be

$$\begin{aligned}
(\nabla^{nh})^{\mathbf{c}}_{\dot{W}}\dot{W} &= \left(\ddot{q}^k + \Gamma^k_{ij}\dot{q}^i\dot{q}^j\right)\frac{\partial}{\partial q^k}\\
&\quad + \left(\ddot{W}^k + \dot{q}^i\dot{q}^j W^l\frac{\partial \Gamma^k_{ij}}{\partial q^l} + \dot{q}^j\dot{W}^i(\Gamma^k_{ij} + \Gamma^k_{ji})\right)\frac{\partial}{\partial \dot{q}^k}.
\end{aligned} \tag{6.2.15}$$

The first term vanishes since c is a geodesic for ∇^{nh} by Theorem 5.1.8. Hence, we get the expected result. □

Denote by T^{nh} and R^{nh} the torsion and curvature tensors, respectively, associated with the nonholonomic connection ∇^{nh}, that is,

$$T^{nh}(X,Y) = \nabla^{nh}_X Y - \nabla^{nh}_Y X - [X,Y],$$
$$R^{nh}(X,Y)Z = \nabla^{nh}_X \nabla^{nh}_Y Z - \nabla^{nh}_Y \nabla^{nh}_X Z - \nabla^{nh}_{[X,Y]} Z,$$

for $X, Y, Z \in \mathfrak{X}(Q)$. Then, using T^{nh} and R^{nh}, we will obtain a characterization of the nonholonomic Jacobi fields with an equation which may be considered as the version for kinematic nonholonomic systems of the Jacobi equation in Riemannian geometry.

Theorem 6.2.21. *Let $(L_g, \mathcal{D})$ be a kinematic nonholonomic system, ∇^{nh} the nonholonomic connection on Q with torsion and curvature tensors denoted by T^{nh} and R^{nh}, respectively, and $W : I \to TQ$ a vector field along a nonholonomic trajectory $c : I \to Q$. Then W is a nonholonomic Jacobi field if and only if*

$$\nabla^{nh}_{\dot c}\nabla^{nh}_{\dot c} W + \nabla^{nh}_{\dot c} T^{nh}(W,\dot c) + R^{nh}(W,\dot c)\dot c = 0, \quad \dot W(t) \in \mathcal{D}^{\mathbf{c}}_{W(t)}. \tag{6.2.16}$$

Proof. Using the same notation introduced both in the statement of the last proposition as well as in its proof, let us compute the coordinate expression of the left-hand side of equation (6.2.16).

It is easy to see that

$$\nabla^{nh}_{\dot c} W = \left(\dot W^k + \dot q^i W^j \Gamma^k_{ij}\right)\frac{\partial}{\partial q^k}.$$

Computing the second covariant derivative we obtain

$$\begin{aligned}\nabla^{nh}_{\dot c}\nabla^{nh}_{\dot c} W = \Big(\ddot W^m + 2\dot W^j \dot q^i \Gamma^m_{ij} + \ddot q^i W^j \Gamma^m_{ij} + \dot q^i W^j \Gamma^k_{ij} \dot q^l \Gamma^m_{lk} \\ + \dot q^i W^j \frac{\partial \Gamma^m_{ij}}{\partial q^l}\dot q^l\Big)\frac{\partial}{\partial q^m}.\end{aligned} \tag{6.2.17}$$

Now, the term with the curvature tensor appearing in equation (6.2.16) is

$$R^{nh}(W,\dot c)\dot c = W^i \dot q^j \dot q^l \left(\frac{\partial \Gamma^m_{jl}}{\partial q^i} + \Gamma^k_{jl}\Gamma^m_{ik} - \frac{\partial \Gamma^m_{il}}{\partial q^j} - \Gamma^k_{il}\Gamma^m_{jk}\right)\frac{\partial}{\partial q^m}, \tag{6.2.18}$$

while the torsion tensor is

$$T^{nh}(W,\dot c) = W^i \dot q^j T^m_{ij}\frac{\partial}{\partial q^m}, \quad \text{with } T^m_{ij} = \Gamma^m_{ij} - \Gamma^m_{ji},$$

and the term involving the covariant derivative of the torsion tensor is

$$\nabla^{nh}_{\dot{c}}T^{nh}(W,\dot{c}) = \left(\dot{W}^i\dot{q}^jT^m_{ij} + W^i\ddot{q}^jT^m_{ij} + W^i\dot{q}^j\frac{\partial T^m_{ij}}{\partial q^l}\dot{q}^l + W^i\dot{q}^jT^k_{ij}\dot{q}^l\Gamma^m_{lk}\right)\frac{\partial}{\partial q^m} \tag{6.2.19}$$

Adding the three terms appearing in equation (6.2.16), we deduce that their sum is equal to

$$\left(\ddot{W}^m + \dot{q}^i\dot{q}^jW^l\frac{\partial \Gamma^m_{ij}}{\partial q^l} + 2\dot{q}^i\dot{W}^j\Gamma^m_{ij} + \dot{W}^i\dot{q}^jT^m_{ij}\right)\frac{\partial}{\partial q^m}, \tag{6.2.20}$$

which implies that

$$(\nabla^{nh})^{\mathbf{c}}_{\dot{W}}\dot{W} = \left(\nabla^{nh}_{\dot{c}}\nabla^{nh}_{\dot{c}}W + \nabla^{nh}_{\dot{c}}T^{nh}(W,\dot{c}) + R^{nh}(W,\dot{c})\dot{c}\right)^{\mathbf{v}}.$$

Indeed note that adding the third term in (6.2.17) with the second one in (6.2.18) we get

$$\ddot{q}^iW^j\Gamma^m_{ij} + W^i\dot{q}^j\dot{q}^l\Gamma^k_{jl}\Gamma^m_{ik} = W^i\dot{q}^j\dot{q}^l\Gamma^k_{jl}T^m_{ik},$$

adding the fourth term in (6.2.17) with the last term in (6.2.18)

$$\dot{q}^i\dot{q}^lW^j\Gamma^k_{ij}\Gamma^m_{lk} - W^i\dot{q}^j\dot{q}^l\Gamma^k_{il}\Gamma^m_{jk} = W^j\dot{q}^i\dot{q}^lT^k_{ij}\Gamma^m_{lk}$$

and adding the last term in (6.2.17) with the first and third terms in (6.2.18) we get

$$\dot{q}^i\dot{q}^lW^j\frac{\partial \Gamma^m_{ij}}{\partial q^l} + W^i\dot{q}^j\dot{q}^l\left(\frac{\partial \Gamma^m_{jl}}{\partial q^i} - \frac{\partial \Gamma^m_{il}}{\partial q^j}\right) = \dot{q}^i\dot{q}^lW^j\left(\frac{\partial T^m_{ij}}{\partial q^l} + \frac{\partial \Gamma^m_{il}}{\partial q^j}\right).$$

The sum of $\nabla^{nh}_{\dot{c}}\nabla^{nh}_{\dot{c}}W$ and $R^{nh}(W,\dot{c})\dot{c}$ is

$$\left[\ddot{W}^m + 2\dot{W}^j\dot{q}^i\Gamma^m_{ij} + W^i\dot{q}^j\dot{q}^l\Gamma^k_{jl}T^m_{ik} + W^j\dot{q}^i\dot{q}^lT^k_{ij}\Gamma^m_{lk} + \dot{q}^i\dot{q}^lW^j\left(\frac{\partial T^m_{ij}}{\partial q^l} + \frac{\partial \Gamma^m_{il}}{\partial q^j}\right)\right]\frac{\partial}{\partial q^m}.$$

Comparing the expression above with our goal, which is to prove that the sum of the three terms is equal to (6.2.20), the result would be proven if we establish that

$$\left(\dot{W}^i\dot{q}^jT^m_{ij} - W^i\dot{q}^j\dot{q}^l\Gamma^k_{jl}T^m_{ik} - W^j\dot{q}^i\dot{q}^lT^k_{ij}\Gamma^m_{lk} - \dot{q}^i\dot{q}^lW^j\frac{\partial T^m_{ij}}{\partial q^l}\right)\frac{\partial}{\partial q^m} = \nabla^{nh}_{\dot{c}}T^{nh}(W,\dot{c})$$

Using that c is a geodesic, so

$$\ddot{q}^i = -\Gamma^i_{jk}\dot{q}^j\dot{q}^k,$$

and the identity

$$T^m_{ik} = -T^m_{ki}$$

we get

$$\begin{aligned}\nabla^{nh}_{\dot{c}} T^{nh}(W,\dot{c}) = \Big(\dot{W}^i\dot{q}^j T^m_{ij} - W^i\Gamma^j_{lk}\dot{q}^l\dot{q}^k T^m_{ij} - W^i\dot{q}^j\frac{\partial T^m_{ji}}{\partial q^l}\dot{q}^l \\ -W^i\dot{q}^j T^k_{ji}\dot{q}^l\Gamma^m_{lk}\Big)\frac{\partial}{\partial q^m},\end{aligned} \tag{6.2.21}$$

which is what we expected. Hence, by Proposition 6.2.18 and the remark following it, we proved that W is a nonholonomic Jacobi field if and only if it satisfies equations (6.2.16).

□

The coordinate expression of the nonholonomic Jacobi equation is still a second-order differential equation. Indeed, the local expression of (6.2.16) is

$$\left(\ddot{W}^k + \dot{q}^i\dot{q}^j W^l\frac{\partial\Gamma^k_{ij}}{\partial q^l} + 2\dot{q}^i\dot{W}^j\Gamma^k_{ij} - \dot{q}^i\dot{W}^j T^k_{ij}\right)\frac{\partial}{\partial q^k}. \tag{6.2.22}$$

Equation (6.2.16) is called the *nonholonomic Jacobi equation* for the nonholonomic geodesic problem.

In order to compute the nonholonomic Christoffel symbols, consider the following local expressions for the orthogonal projectors

$$P(\partial q^i) = P^i_j\partial q^j, \quad (P')(\partial q^i) = (P')^i_j\partial q^j. \tag{6.2.23}$$

We have the following lemma, which is proven using the definition of the nonholonomic connection and the properties of linear connections.

Lemma 6.2.22. *The nonholonomic Christoffel symbols are given by*

$$\Gamma^k_{ij} = (\overset{g}{\Gamma})^l_{ij}P^k_l + \frac{\partial (P')^k_j}{\partial q^i} + (P')^l_j(\overset{g}{\Gamma})^k_{il}, \tag{6.2.24}$$

where $\overset{g}{(\Gamma)}{}^k_{ij}$ *are the Levi-Civita Christoffel symbols. Moreover, if the Riemannian metric* g *is flat, we have that*

$$\Gamma^k_{ij} = \frac{\partial (P')^k_j}{\partial q^i}. \tag{6.2.25}$$

Example 6.2.23. Recall the nonholonomic particle given by

$$L(x, y, z, \dot{x}, \dot{y}, \dot{z}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

and subjected to the nonholonomic constraint $\dot{z} - y\dot{x} = 0$.

As we have seen before, nonholonomic Jacobi fields may be obtained using two different geometric frameworks: either as the trajectories of the lifted nonholonomic system $\Gamma_{(L_{g^c}, \mathcal{D}^c)}$ or as the solution of the nonholonomic Jacobi equation (6.2.16). Let us explore both these characterizations in this particular example.

(i) We are going to obtain the Lagrange-d'Alembert equations for the nonholonomic system $L_{g^{\mathbf{c}}}$ with constraint distribution $\mathcal{D}^{\mathbf{c}}$.

The Lagrangian function is

$$L_{g^{\mathbf{c}}}(q, r, \dot{q}, \dot{r}) = \dot{x}\dot{u} + \dot{y}\dot{v} + \dot{z}\dot{w}.$$

where $q = (x, y, z)$, $r = (u, v, w)$ and the lifted distribution $\mathcal{D}^{\mathbf{c}}$ is given by the span of the vectors

$$\mathcal{D}^{\mathbf{c}} = \left\langle \left\{ \frac{\partial}{\partial u} + y\frac{\partial}{\partial w}, \frac{\partial}{\partial v}, \frac{\partial}{\partial x} + y\frac{\partial}{\partial z} + v\frac{\partial}{\partial w}, \frac{\partial}{\partial y} \right\} \right\rangle$$

with annihilator $(\mathcal{D}^{\mathbf{c}})^o$

$$(\mathcal{D}^{\mathbf{c}})^o = \langle \{-y dx + dz, -v dx - y du + dw\} \rangle .$$

Hence, the new nonholonomic constraints are $\dot{z} - y\dot{x} = 0$ and $\dot{w} - v\dot{x} - y\dot{u} = 0$. The Lagrange-d'Alembert equations are then

$$\begin{aligned} \ddot{x} &= -y\lambda_2 & \ddot{u} &= -y\lambda_1 - v\lambda_2 \\ \ddot{y} &= 0 & \ddot{v} &= 0 \\ \ddot{z} &= \lambda_2 & \ddot{w} &= \lambda_1 \\ \dot{z} - y\dot{x} &= 0, & \dot{w} - v\dot{x} - y\dot{u} &= 0, \end{aligned}$$

and solving for the Lagrange multipliers' λ_1 and λ_2, we obtain

$$\lambda_1 = \frac{(\dot{u}\dot{y} + \dot{x}\dot{v})(1+y^2) - 2yv\dot{x}\dot{y}}{(1+y^2)^2}, \quad \lambda_2 = \frac{\dot{x}\dot{y}}{1+y^2}.$$

(ii) We will compute the nonholonomic Jacobi equation using the local expression deduced in (6.2.22). The only non-vanishing Christoffel symbols relative to the nonholonomic connection ∇^{nh} are

$$\Gamma^x_{yx} = \frac{2y}{(1+y^2)^2}, \quad \Gamma^z_{yx} = \Gamma^x_{yz} = \frac{y^2-1}{(1+y^2)^2}, \quad \Gamma^z_{yz} = -\frac{2y}{(1+y^2)^2},$$

which implies that the non-vanishing torsion entries are

$$T^x_{yx} = \frac{2y}{(1+y^2)^2}, \quad T^z_{yx} = T^x_{yz} = \frac{y^2-1}{(1+y^2)^2}, \quad T^z_{yz} = -\frac{2y}{(1+y^2)^2},$$

along with the corresponding skew-symmetric entries. If the vector field W is given by

$$W = u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z},$$

then Jacobi equation together with the constraint that $\dot{W} \in \mathcal{D}^{\mathbf{c}}$ gives

$$\begin{aligned}
&\ddot{u} + v\left(\dot{x}\dot{y}\frac{\partial \Gamma^x_{yx}}{\partial y} + \dot{z}\dot{y}\frac{\partial \Gamma^x_{yz}}{\partial y}\right) + 2(\dot{u}\dot{y}\Gamma^x_{yx} + \dot{w}\dot{y}\Gamma^x_{yz}) - (\dot{u}\dot{y}T^x_{yx} + \dot{w}\dot{y}T^x_{yz}) = 0\\
&\ddot{v} = 0\\
&\ddot{w} + v\left(\dot{x}\dot{y}\frac{\partial \Gamma^z_{yx}}{\partial y} + \dot{z}\dot{y}\frac{\partial \Gamma^z_{yz}}{\partial y}\right) + 2(\dot{u}\dot{y}\Gamma^z_{yx} + \dot{w}\dot{y}\Gamma^z_{yz}) - (\dot{u}\dot{y}T^z_{yx} + \dot{w}\dot{y}T^z_{yz}) = 0\\
&\dot{w} - v\dot{x} - y\dot{u} = 0.
\end{aligned} \tag{6.2.26}$$

The fact that W is a vector field along a nonholonomic geodesic satisfying $\dot{z} = y\dot{x}$ simplifies the equation. Moreover, since $T^x_{yx} = \Gamma^x_{yx}$, $T^z_{yx} = T^x_{yz} = \Gamma^z_{yx} = \Gamma^x_{yz}$ and $T^z_{yz} = \Gamma^z_{yz}$ simplifies even more the equation that reduces to

$$\begin{aligned}
&\ddot{u} + v\dot{x}\dot{y}\left(\frac{\partial \Gamma^x_{yx}}{\partial y} + y\frac{\partial \Gamma^x_{yz}}{\partial y}\right) + \dot{u}\dot{y}\Gamma^x_{yx} + \dot{w}\dot{y}\Gamma^x_{yz} = 0\\
&\ddot{v} = 0\\
&\ddot{w} + v\dot{x}\dot{y}\left(\frac{\partial \Gamma^z_{yx}}{\partial y} + y\frac{\partial \Gamma^z_{yz}}{\partial y}\right) + \dot{u}\dot{y}\Gamma^z_{yx} + \dot{w}\dot{y}\Gamma^z_{yz} = 0\\
&\dot{w} - v\dot{x} - y\dot{u} = 0.
\end{aligned} \tag{6.2.27}$$

It is easy to see now that both approaches coincide.

△

Example 6.2.24. Let us consider again the nonholonomic particle. We proved before that $\frac{\partial}{\partial z}$ was a Jacobi field. Let us check that it satisfies the nonholonomic Jacobi equation.

In fact, since the component functions of $\frac{\partial}{\partial z}$ in the coordinate basis are constant, the local expression of the left-hand side of the nonholonomic Jacobi equation reduces to

$$\nabla_{\dot{c}}^{nh}\nabla_{\dot{c}}^{nh}W + \nabla_{\dot{c}}^{nh}T^{nh}(W,\dot{c}) + R^{nh}(W,\dot{c})\dot{c} = \dot{q}^i\dot{q}^j\frac{\partial\Gamma_{ij}^k}{\partial z}\frac{\partial}{\partial q^k},$$

where $(q^i(t))$ are the coordinate expression of a fixed geodesic. However, since the only non-vanishing Christoffel symbols relative to the nonholonomic connection are

$$\Gamma_{yx}^x = \frac{2y}{(1+y^2)^2}, \quad \Gamma_{yx}^z = \Gamma_{yz}^x = \frac{y^2-1}{(1+y^2)^2}, \quad \Gamma_{yz}^z = -\frac{2y}{(1+y^2)^2},$$

there is no dependence on the coordinate z, hence the expression above vanishes, i.e., the nonholonomic Jacobi equations is satisfied. △

Example 6.2.25. For the vertical rolling disk (see Example 6.2.7), apart from $W_3 = \dfrac{\partial}{\partial\theta}$, the following vector fields

$$W_1 = \frac{\partial}{\partial x}, \quad W_2 = \frac{\partial}{\partial y},$$

$$Z = \frac{\partial}{\partial\phi} - y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$$

are nonholonomic Jacobi vector fields.

In fact, the nonholonomic Christoffel symbols are computed with (6.2.25), obtaining the following non-vanishing expressions

$$\Gamma_{\phi x}^x = \cos\phi\sin\phi, \quad \Gamma_{\phi x}^y = \frac{1}{2}(\sin^2\phi - \cos^2\phi), \quad \Gamma_{\phi x}^\theta = \frac{\sin\phi}{2},$$

$$\Gamma_{\phi y}^x = \frac{1}{2}(\sin^2\phi - \cos^2\phi), \quad \Gamma_{\phi y}^y = -\cos\phi\sin\phi, \quad \Gamma_{\phi y}^\theta = -\frac{\cos\phi}{2},$$

$$\Gamma^x_{\phi x} = \frac{\sin\phi}{2}, \quad \Gamma^y_{\phi x} = -\frac{\cos\phi}{2}.$$

Now, since W_i are constant vector fields, its time derivatives vanish and since none of the Christoffel symbols has explicit dependence on the variables θ, x and y, equation (6.2.22) is trivially satisfied. For Z, we have a slightly longer computation since the coefficients do not vanish, but with the help of a symbolic computation software, we eventually conclude that Z is a nonholonomic Jacobi field restricted to any nonholonomic trajectory. △

Remark 6.2.26. A straightforward computation proves that the vector fields Z and W_i, $i = 1, 2, 3$, are infinitesimal symmetries for the vertical rolling disk in the terminology of Corollary 6.2.4. Thus, one can directly deduce that they are nonholonomic Jacobi fields.

Example 6.2.27. Given (x, y) are coordinates on the plane and (ϕ, ψ, θ) Euler angles determining a rotation on $SO(3)$, consider the Lagrangian function $L : T(\mathbb{R}^2 \times SO(3)) \to \mathbb{R}$ describing the kinetic energy of a homogeneous ball rolling on a plane, locally given by

$$L(x, y, \phi, \psi, \theta, \dot{x}, \dot{y}, \dot{\phi}, \dot{\psi}, \dot{\theta}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}(\dot{\phi}^2 + \dot{\psi}^2 + \dot{\theta}^2 + 2\cos(\theta)\dot{\phi}\dot{\psi}),$$

where that the mass m and the radius r of the ball are $m = r = 1$ and we assume that the matrix of the inertia tensor is the identity. The ball rolls on the plane without slipping which is equivalent to the constraints

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

The vector fields

$$W_1 = \frac{\partial}{\partial x}, \quad W_2 = \frac{\partial}{\partial y}, \quad W_3 = \frac{\partial}{\partial \psi}$$

$$Z_1 = \frac{\sin\psi}{\sin\theta}\frac{\partial}{\partial\phi} - \frac{\cos\theta\sin\psi}{\sin\theta}\frac{\partial}{\partial\psi} + \cos\psi\frac{\partial}{\partial\theta}$$

$$Z_2 = -\frac{\cos\psi}{\sin\theta}\frac{\partial}{\partial\phi} + \frac{\cos\theta\cos\psi}{\sin\theta}\frac{\partial}{\partial\psi} + \sin\psi\frac{\partial}{\partial\theta}$$

are nonholonomic Jacobi fields.

In fact, the non-vanishing Levi-Civita Christoffel symbols are

$$\Gamma^{\phi}_{\phi\theta} = \frac{1}{2}\frac{\cos\theta}{\sin\theta}, \quad \Gamma^{\phi}_{\psi\theta} = -\frac{1}{2\sin\theta}, \quad \Gamma^{\psi}_{\phi\theta} = -\frac{1}{2\sin\theta},$$

$$\Gamma^{\psi}_{\psi\theta} = \frac{1}{2}\frac{\cos\theta}{\sin\theta}, \quad \Gamma^{\theta}_{\phi\psi} = \frac{1}{2}\sin\theta,$$

and, thus, we may use equation (6.2.24) to compute the nonholonomic Christoffel symbols. Then, we may check that along any curve satisfying the constraints, i.e., $\dot{q} \in \mathcal{D}$, the vector fields $W_i(q(t))$ and $Z_i(q(t))$ satisfy the nonholonomic Jacobi equation (6.2.22) and thus they are indeed nonholonomic Jacobi vector fields along any trajectory of the nonholonomic system.

Remark 6.2.28. As in the previous example, a long but straightforward computation proves that the vector fields Z_i, $i = 1, 2$, and W_j, $j = 1, 2, 3$, also are infinitesimal symmetries for the ball rolling on the plane.

△

6.3 Nonholonomic Jacobi fields for mechanical systems

Now, we will introduce the complete lift of a mechanical nonholonomic system following the same ideas that in the case of a kinetic nonholonomic system.

Definition 6.3.1. Let $L_{(g,V)}$ be the mechanical Lagrangian function on TQ associated with a Riemannian metric g on Q and the potential energy $V \in C^\infty(Q)$. The nonholonomic system $(L_{(g^c,V^c)}, \mathcal{D}^{\mathbf{c}})$ is the complete lift of the nonholonomic system $(L_{(g,V)}, \mathcal{D})$, with the Lagrangian function of mechanical type $L_{(g^c,V^c)} : TTQ \to \mathbb{R}$ defined by

$$L_{(g^c,V^c)} = L_{g^{\mathbf{c}}} - V^{\mathbf{c}} \circ \tau_{TQ},$$

where $\tau_{TQ} : T(TQ) \to TQ$ is the canonical projection and $\mathcal{D}^{\mathbf{c}}$ is the complete lift of the distribution $\mathcal{D}$.

Next, we will prove some results which will be used later.

Lemma 6.3.2. *If V is a function on Q, its complete lift satisfies*

$$(V \circ \tau_Q)^{\mathbf{c}} \circ \kappa_Q = V^{\mathbf{c}} \circ \tau_{TQ}. \tag{6.3.1}$$

Proof. Observe that for any $Y \in T_{v_q}TQ$, with $v_q \in T_qQ$, we have that

$$(V \circ \tau_Q)^{\mathbf{c}} \circ \kappa_Q(Y) = \langle d(V \circ \tau_Q)(\tau_{TQ}(\kappa_Q(Y))), \kappa_Q(Y)\rangle$$

by (2.4.1). Then we deduce that

$$(V \circ \tau_Q)^{\mathbf{c}} \circ \kappa_Q(Y) = \langle dV(q), T\tau_Q(\kappa_Q(Y))\rangle = \langle dV(q), \tau_{TQ}(Y)\rangle = V^{\mathbf{c}} \circ \tau_{TQ}(Y),$$

where we used the fact that κ_Q is a morphism between the vector bundles $T\tau_Q$ and τ_{TQ} and (2.4.1). $\square$

Lemma 6.3.3. *Let g be a Riemannian metric, $g^{\mathbf{c}}$ its complete lift and V a smooth function on Q. We have that*

$$\mathrm{grad}_{g^{\mathbf{c}}} V^{\mathbf{c}} = (\mathrm{grad}_g V)^{\mathbf{c}}. \tag{6.3.2}$$

Proof. On one hand, for an arbitrary $Z \in \mathfrak{X}(Q)$ we have that

$$g^{\mathbf{c}}(\mathrm{grad}_{g^{\mathbf{c}}} V^{\mathbf{c}}, Z^{\mathbf{c}}) = d(V^{\mathbf{c}})(Z^{\mathbf{c}}) = (dV(Z))^{\mathbf{c}},$$

where we used the definition of $\mathrm{grad}_{g^{\mathbf{c}}} V^{\mathbf{c}}$, (2.4.7) and (2.4.8). Then, the definition of $\mathrm{grad}_g V$ implies that

$$g^{\mathbf{c}}(\mathrm{grad}_{g^{\mathbf{c}}} V^{\mathbf{c}}, Z^{\mathbf{c}}) = (g(\mathrm{grad}_g V, Z))^{\mathbf{c}} = g^{\mathbf{c}}((\mathrm{grad}_g V)^{\mathbf{c}}, Z^{\mathbf{c}}),$$

where we used (A.1.1). On the other hand, using the same arguments we deduce

$$g^{\mathbf{c}}(\mathrm{grad}_{g^{\mathbf{c}}} V^{\mathbf{c}}, Z^{\mathbf{v}}) = d(V^{\mathbf{c}})(Z^{\mathbf{v}}) = (dV(Z))^{\mathbf{v}} = (g(\mathrm{grad}_g V, Z))^{\mathbf{v}}.$$

So, from (A.1.1) we conclude that

$$g^{\mathbf{c}}(\mathrm{grad}_{g^{\mathbf{c}}} V^{\mathbf{c}}, Z^{\mathbf{v}}) = g^{\mathbf{c}}((\mathrm{grad}_g V)^{\mathbf{c}}, Z^{\mathbf{v}}).$$

Since $g^{\mathbf{c}}$ is non-degenerate and $Z^{\mathbf{v}}$, $Z^{\mathbf{c}}$, with $Z \in \mathfrak{X}(Q)$, (locally) generate $\mathfrak{X}(TQ)$, we have finished the proof. $\square$

Now, we will prove a similar result to Theorem 6.2.11 for the more general case of a mechanical nonholonomic system.

Theorem 6.3.4. *Let $(L_{(g,V)}, \mathcal{D})$ be a nonholonomic system of mechanical type and $\Gamma_{(L_{(g,V)},\mathcal{D})}$ the associated nonholonomic SODE. Then*

(i) The complete lift $(L_{(g^c,V^c)}, \mathcal{D}^{\mathbf{c}})$ is a regular nonholonomic system.

(ii) Let $\Gamma_{(L_{(g^c,V^c)},\mathcal{D}^{\mathbf{c}})} \in \mathfrak{X}(\mathcal{D}^{\mathbf{c}})$ be the nonholonomic SODE associated with the system $(L_{(g^c,V^c)}, \mathcal{D}^{\mathbf{c}})$ and $\kappa_Q : TTQ \to TTQ$ the canonical involution. Then

$$\Gamma_{(L_{(g^c,V^c)},\mathcal{D}^{\mathbf{c}})} = T\kappa_Q|_{T\mathcal{D}} \circ \Gamma^{\mathbf{c}}_{(L_{(g,V)},\mathcal{D})} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}} \tag{6.3.3}$$

and so we have

(a) $\Gamma_{(L_{(g^c,V^c)},\mathcal{D}^{\mathbf{c}})}$ is $T\tau_Q|_{\mathcal{D}^{\mathbf{c}}}$-projectable over $\Gamma_{(L_{(g,V)},\mathcal{D})}$;

(b) The trajectories of $\Gamma_{(L_{(g^c,V^c)},\mathcal{D}^{\mathbf{c}})}$ are just the Jacobi fields for the nonholonomic system $(L_{(g,V)}, \mathcal{D})$.

Proof. Item (i) is a consequence of Theorem 5.1.12 together with the proof of Proposition 6.2.13, where we see that the distribution $\mathcal{D}^{\mathbf{c}}$ is non-degenerate.

Before proving item (ii), note that by Lemma 5.1.11 and Proposition A.1.2 we have that

$$\omega_{L_{(g^c,V^c)}} = \omega_{L_{g^{\mathbf{c}}}} = \kappa_Q^* \omega^{\mathbf{c}}_{L_g} = \kappa_Q^* \omega^{\mathbf{c}}_{L_{(g,V)}}.$$

Moreover, from Lemma 6.3.2, it follows that

$$E_{L_{(g^c,V^c)}} = E_{L_{g^{\mathbf{c}}}} + V^{\mathbf{c}} \circ \tau_{TQ} = (E_{L_g})^{\mathbf{c}} \circ \kappa_Q + (V \circ \tau_Q)^{\mathbf{c}} \circ \kappa_Q = E^{\mathbf{c}}_{L_{(g,V)}} \circ \kappa_Q.$$

Now we may follow the proof of Theorem 6.2.11 and we conclude by following exactly the same steps that $(\kappa_Q)_* \Gamma^{\mathbf{c}}_{(L_{(g,V)},\mathcal{D})}$ is a vector field on $\mathcal{D}^{\mathbf{c}}$ and by uniqueness of nonholonomic vector field, it coincides with $\Gamma_{(L_{(g^c,V^c)},\mathcal{D}^{\mathbf{c}})}$, i.e.,

$$\Gamma_{(L_{(g^c,V^c)},\mathcal{D}^{\mathbf{c}})} = T\kappa_Q|_{T\mathcal{D}} \circ \Gamma^{\mathbf{c}}_{(L_{(g,V)},\mathcal{D})} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}}. \tag{6.3.4}$$

The remaining statements in item *(ii)* are just consequences of the properties of the complete lift and the canonical involution and we can follow the proof of Theorem 6.2.11 with the necessary changes:

$$\begin{aligned} T(T\tau_Q|_{\mathcal{D}^{\mathbf{c}}})(\Gamma_{(L_{(g^c,V^c)},\mathcal{D}^{\mathbf{c}})}) &= T(T\tau_Q|_{\mathcal{D}^{\mathbf{c}}} \circ \kappa_Q|_{T\mathcal{D}}) \circ \Gamma^{\mathbf{c}}_{(L_{(g,V)},\mathcal{D})} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}} \\ &= T(\tau_{TQ}|_{T\mathcal{D}})(\Gamma^{\mathbf{c}}_{(L_{(g,V)},\mathcal{D})} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}}) \\ &= \Gamma_{(L_{(g,V)},\mathcal{D})} \circ \tau_{TQ}|_{T\mathcal{D}} \circ \kappa_Q|_{\mathcal{D}^{\mathbf{c}}} \\ &= \Gamma_{(L_{(g,V)},\mathcal{D})} \circ T\tau_Q|_{\mathcal{D}^{\mathbf{c}}}. \end{aligned}$$

Now if $W : I \to TQ$ is a trajectory of $\Gamma_{(L_{(g^c,V^c)},\mathcal{D}^{\mathbf{c}})}$, then $\kappa_Q \circ \dot{W} : I \to T\mathcal{D}$ is an integral curve of $\Gamma^{\mathbf{c}}_{(L_{(g,V)},\mathcal{D})}$. Therefore we may write it as

$$\kappa_Q \circ \dot{W}(t) = \left(T_{W(0)} \phi_t^{\Gamma_{(L_{(g,V)},\mathcal{D})}} \right) (\kappa_Q \circ \dot{W}(0)).$$

So,

$$W(t) = T\tau_Q(\kappa_Q \circ \dot{W}) = T\tau_Q \left(\left(T_{W(0)} \phi_t^{\Gamma_{(L_{(g,V)},\mathcal{D})}} \right) (\kappa_Q \circ \dot{W}(0)) \right)$$

and

$$W(t) = \left(T_{W(0)} (\tau_Q \circ \phi_t^{\Gamma_{(L_{(g,V)},\mathcal{D})}}) \right) (\kappa_Q \circ \dot{W}(0)).$$

Let now $v : I \to \mathcal{D}$ be a curve such that its initial velocity is $v'(0) = \kappa_Q \circ \dot{W}(0)$. Then

$$W(t) = \left.\frac{d}{ds}\right|_{s=0} \left(\tau_Q \circ \phi_t^{\Gamma_{(L_{(g,V)},\mathcal{D})}} \right) (v(s)).$$

Hence, W is a nonholonomic Jacobi field for $\Gamma_{(L_{(g,V)},\mathcal{D})}$, since it is an infinitesimal variation of nonholonomic trajectories of $\Gamma_{(L_{(g,V)},\mathcal{D})}$.

□

Finally, we will present the Jacobi equation for the nonholonomic Jacobi fields associated with a mechanical nonholonomic system.

Theorem 6.3.5. *Let $(L_{(g,V)}, \mathcal{D})$ be a mechanical nonholonomic system, ∇^{nh} the nonholonomic connection on Q with torsion and curvature tensors denoted by T^{nh} and R^{nh}, respectively, and $W : I \to TQ$ a vector field along a nonholonomic trajectory $c : I \to Q$. Then W is a nonholonomic Jacobi field if and only if*

$$\begin{aligned} &\nabla^{nh}_{\dot{c}} \nabla^{nh}_{\dot{c}} W + \nabla^{nh}_{\dot{c}} T^{nh}(W, \dot{c}) + R^{nh}(W, \dot{c})\dot{c} + \nabla^{nh}_W (P(\text{grad}_g V \circ c)) = 0, \\ &\dot{W}(t) \in \mathcal{D}^{\mathbf{c}}_{W(t)}. \end{aligned} \tag{6.3.5}$$

Proof. We already know by Theorem 5.1.13 that if $c_v : I \to Q$ is a trajectory of $\Gamma_{(L_{(g,V)},\mathcal{D})}$, then it satisfies equations (5.1.8). Moreover, by Theorem 6.3.4 if $W : I \to TQ$ is a Jacobi field for the nonholonomic system $(L_{(g,V)}, \mathcal{D})$, then it is a trajectory of the nonholonomic SODE $\Gamma_{(L_{(g^c,V^c)},\mathcal{D}^{\mathbf{c}})}$. As a result, W must satisfy the equations

$$\nabla^{NH}_{\dot{W}}\dot{W} = -P^T(\mathrm{grad}_{g^{\mathbf{c}}}V^{\mathbf{c}} \circ W), \quad \dot{W} \in \mathcal{D}^{\mathbf{c}},$$

where ∇^{NH} is the linear connection on TQ defined by

$$\nabla^{NH}_X Y := P^T(\nabla^{g^{\mathbf{c}}}_X Y) + \nabla^{g^{\mathbf{c}}}_X[P'^T(Y)], \quad \text{for } X, Y \in \mathfrak{X}(TQ),$$

with $\nabla^{g^{\mathbf{c}}}$ the Levi-Civita connection of $g^{\mathbf{c}}$, $P^T : TTQ \to \mathcal{D}^{\mathbf{c}}$ the associated orthogonal projector onto the distribution $\mathcal{D}^{\mathbf{c}}$ and $P'^T : TTQ \to (\mathcal{D}^{\mathbf{c}})^{\perp}$ the orthogonal projector onto $(\mathcal{D}^{\mathbf{c}})^{\perp}$, the orthogonal distribution.

On one hand, by Proposition 6.2.18, $\nabla^{NH} = (\nabla^{nh})^{\mathbf{c}}$. On the other hand, by Lemma 6.3.3, we have that

$$\mathrm{grad}_{g^{\mathbf{c}}}V^{\mathbf{c}} = (\mathrm{grad}_g V)^{\mathbf{c}}$$

and by Lemma 6.2.17 we have that

$$P^T(X^{\mathbf{c}}) = (P(X))^{\mathbf{c}}.$$

Hence, W must satisfy

$$(\nabla^{nh})^{\mathbf{c}}_{\dot{W}}\dot{W} = -(P(\mathrm{grad}_g V))^{\mathbf{c}} \circ W, \quad \dot{W}(t) \in \mathcal{D}^{\mathbf{c}}_{W(t)}.$$

Now, we will follow Proposition 6.2.20 and keep the same notation that was introduced in the corresponding proof. Suppose that the local expression of $P(\mathrm{grad}_g V)$ is

$$P(\mathrm{grad}_g V) = (P(\mathrm{grad}_g V))^i \frac{\partial}{\partial q^i}.$$

Then, equation (6.2.15) together with

$$(P(\mathrm{grad}_g V))^{\mathbf{c}}(W(t)) = (P(\mathrm{grad}_g V))^i \left.\frac{\partial}{\partial q^i}\right|_{W(t)} + W^j(t)\frac{\partial(P(\mathrm{grad}_g V))^i}{\partial q^j} \left.\frac{\partial}{\partial \dot{q}^i}\right|_{W(t)}$$

and the fact that c_v satisfies the equations (5.1.8) imply that

$$(\nabla^{nh})^{\mathbf{c}}_{\dot{W}}\dot{W} + (P(\mathrm{grad}_g V))^{\mathbf{c}} \circ W = \left(\ddot{W}^k + \dot{q}^i\dot{q}^j W^l \frac{\partial \Gamma^k_{ij}}{\partial q^l} \right.$$
$$\left. + \dot{q}^j\dot{W}^i(\Gamma^k_{ij} + \Gamma^k_{ji}) + W^j(t)\frac{\partial(P(\mathrm{grad}_g V))^k}{\partial q^j}\right)\frac{\partial}{\partial \dot{q}^k}.$$

Using similar techniques to those applied in the proof of Theorem 6.2.21, we are able to prove that

$$
\begin{aligned}
&(\nabla^{nh})^{\mathbf{c}}_{\dot{W}}\dot{W} + (P(\operatorname{grad}_g V))^{\mathbf{c}} \circ W = \\
&\qquad \left(\nabla^{nh}_{\dot{c}}\nabla^{nh}_{\dot{c}}W + \nabla^{nh}_{\dot{c}}T^{nh}(W,\dot{c}) + R^{nh}(W,\dot{c})\dot{c} + \nabla^{nh}_{W}(P(\operatorname{grad}_g V \circ c))\right)^{\mathbf{v}}
\end{aligned}
$$

Indeed, by following its proof and having in mind that now the curve c_v locally satisfies the equation

$$
\ddot{q}^i = -\Gamma^i_{jk}\dot{q}^j\dot{q}^k - (P(\operatorname{grad}_g V))^i,
$$

we deduce that the sum of $\nabla^{nh}_{\dot{c}}\nabla^{nh}_{\dot{c}}W$ and $R^{nh}(W,\dot{c})\dot{c}$ is

$$
\begin{aligned}
&\Big[\ddot{W}^m + 2\dot{W}^j\dot{q}^i\Gamma^m_{ij} + W^i\dot{q}^j\dot{q}^l\Gamma^k_{jl}T^m_{ik} + W^j\dot{q}^i\dot{q}^l T^k_{ij}\Gamma^m_{lk} \\
&+\dot{q}^i\dot{q}^l W^j\left(\frac{\partial T^m_{ij}}{\partial q^l} + \frac{\partial \Gamma^m_{il}}{\partial q^j}\right) - (P(\operatorname{grad}_g V))^i W^j\Gamma^m_{ij}\Big]\frac{\partial}{\partial q^m}.
\end{aligned}
$$

Since

$$
\begin{aligned}
\nabla^{nh}_{\dot{c}}T^{nh}(W,\dot{c}) = \Big(&\dot{W}^i\dot{q}^j T^m_{ij} - W^i\Gamma^j_{lk}\dot{q}^l\dot{q}^k T^m_{ij} - W^i(P(\operatorname{grad}_g V))^j T^m_{ij} \\
&-W^i\dot{q}^j\frac{\partial T^m_{ji}}{\partial q^l}\dot{q}^l - W^i\dot{q}^j T^k_{ji}\dot{q}^l\Gamma^m_{lk}\Big)\frac{\partial}{\partial q^m}
\end{aligned}
\tag{6.3.6}
$$

and

$$
\nabla^{nh}_{W}(P(\operatorname{grad}_g V)) = \left(W^i\frac{\partial (P(\operatorname{grad}_g V))^j}{\partial q^i} + W^i(P(\operatorname{grad}_g V))^k\Gamma^j_{ik}\right)\frac{\partial}{\partial q^j}
$$

we obtain the expected result and the theorem follows. □

Chapter 7

Discrete nonholonomic mechanics

In this chapter, we will discuss discrete nonholonomic mechanics (see also [CM01; MP06; FID08; BZ15; FBO12; Cel+19; Igl+08; MV19; CM01; LDSM04; PL19]). Our first goal is to be able to identify the exact discrete flow of nonholonomic mechanics which will be a map of the form

$$\Phi^e_{h,nh} : \mathcal{M}^{e,nh}_h \to \mathcal{M}^{e,nh}_h,$$

where $\mathcal{M}^{e,nh}_h$ is the submanifold defined in Definition 4.3.3 in Section 4.3, satisfying the property that if c is the unique nonholonomic trajectory satisfying $c(0) = q_0$ and $c(h) = q_1$ then the pair $\Phi^e_{h,nh}(q_0, q_1) = (q_1, q_2)$ satisfies $q_2 = c(2h)$.

Then we will propose a set of discrete equations satisfying a new discrete "variational" principle which is able to incorporate the exact discrete flow as a particular solution for the appropriate choice of objects (discrete Lagrangian function, discrete space of constraints and discrete forces).

As it will become clear below, this principle is based on the assumption that we are given objects such as a discrete Lagrangian function and a discrete constraint space on the ambient manifold $Q \times Q$, which is the discrete version of TQ. Nonetheless, we take a few steps in the direction of finding an "intrinsic" version of our discrete principle using objects that are defined on $\mathcal{M}^{e,nh}_h$ right from the beginning.

Finally, we show how the proposed discrete equations perform in numerical simulations.

7.1 Modified Lagrange-d'Alembert principle

In this section, we introduce a modification of the Lagrange-d'Alembert principle. Later, using the construction of the nonholonomic exponential map in Section 4.3, we will define the exact discrete version of nonholonomic mechanics and show that it actually satisfies the modified Lagrange-d'Alembert principle.

Let $\mathcal{D}$ be a distribution on the manifold Q, $L_d : Q \times Q \longrightarrow \mathbb{R}$ a discrete Lagrangian function, $F_d^{\pm} : Q \times Q \longrightarrow T^*Q$ discrete forces and $\mathcal{M}_h^d \subseteq Q \times Q$ a submanifold called the *discrete constraint space*. We remark that $\pi_Q \circ F_d^+ = \mathrm{pr}_2$ and $\pi_Q \circ F_d^- = \mathrm{pr}_1$, where $\pi_Q : T^*Q \to Q$ and $\mathrm{pr}_{1,2} : Q \times Q \to Q$ are the canonical projections (see [MW01]).

Thus, a discrete nonholonomic system on the configuration manifold Q is determined by a tuple $(L_d, \mathcal{D}, \mathcal{M}_h^d, F_d^{\pm})$. In the following definition we will use the discrete action defined in (3.7.1) in Section 3.7.

Definition 7.1.1. Given $L_d, \mathcal{D}, \mathcal{M}_h^d, F_d^{\pm}$ as before, a sequence $(q_0, ..., q_N)$ in Q satisfies the *modified Lagrange-d'Alembert principle* associated with the discrete nonholonomic system $(L_d, \mathcal{D}, \mathcal{M}_h^d, F_d^{\pm})$ if it extremizes

$$\begin{aligned} &\delta S_d(q_d) \cdot \delta q_d + \sum_{k=1}^{N-1} \left[F_d^+(q_{k-1}, q_k) + F_d^-(q_k, q_{k+1}) \right] \cdot \delta q_k = 0 \\ &(q_k, q_{k+1}) \in \mathcal{M}_h^d, \ 0 \leqslant k \leqslant N-1 \end{aligned} \tag{7.1.1}$$

for all variations lying in the distribution $\delta q_k \in \mathcal{D}_{q_k}$, $\delta q_d = (\delta q_0, ..., \delta q_N) \in T_{q_d}\mathcal{C}_d(q_0, q_N)$ and $\delta q_0 = \delta q_N = 0$.

Remark 7.1.2. Observe that this principle is exactly the same than the discrete Lagrange-d'Alembert principle for forced systems when $\mathcal{D} = TQ$ and $\mathcal{M}_h^d = Q \times Q$ (see Section 3.7.7). It is also the discrete Lagrange-d'Alembert principle for nonholonomic systems introduced in [CM01] when $F_d^+ = F_d^- = 0$ (see also Section 3.7.8). Also, in this context we find the methods proposed by [LDSM04], using a discretization of the forces for a nonholonomic system and a discrete submanifold derived from the continuous constraints and the forced discrete Legendre transformations. Recently, a similar principle was introduced in [PL19] to study discretizations of Dirac mechanics.

Now, as in the case of forced systems, we have that

Proposition 7.1.3. *Given $L_d, \mathcal{D}, \mathcal{M}_h^d, F_d^{\pm}$ as defined previously, a sequence $(q_0, ..., q_N)$ in Q satisfies the modified Lagrange-d'Alembert principle associated with the discrete nonholonomic system $(L_d, \mathcal{D}, \mathcal{M}_h^d, F_d^{\pm})$ if and only if it satisfies modified Lagrange-d'Alembert equations*

$$\begin{aligned} &D_2L_d(q_{k-1}, q_k) + D_1L_d(q_k, q_{k+1}) + F_d^+(q_{k-1}, q_k) + F_d^-(q_k, q_{k+1}) \in \mathcal{D}^o_{q_k} \\ &\omega^a(q_k, q_{k+1}) = 0, \ 0 \leqslant k \leqslant N-1, \quad 0 \leqslant a \leqslant n-k \end{aligned} \tag{7.1.2}$$

where $\mathcal{M}_d$ is determined by the zeros of a set of constraint functions $\omega^a : Q \times Q \longrightarrow \mathbb{R}$.

Proof. The proof develops in the same lines as the proof of Theorem 3.7.17, the only essential difference being the introduction of the force maps. □

7.2 The exact discrete nonholonomic flow

For the remaining of the section, consider the regular nonholonomic system $(L, \mathcal{D})$, with associated nonholonomic vector field $\Gamma_{(L,\mathcal{D})}$, whose flow and nonholonomic exponential map are $\phi_t^{\Gamma_{(L,\mathcal{D})}}$ and $\exp_h^{\Gamma_{(L,\mathcal{D})}}$, respectively. Let $R^{e-}_{h,nh}$ be the inverse map of the nonholonomic exponential map. It is called the exact (negative) inverse retraction. Let the map $R^{e+}_{h,nh}$ defined by

$$R^{e+}_{h,nh} = \phi_h^{\Gamma_{(L,\mathcal{D})}} \circ R^{e-}_{h,nh}$$

be called the exact positive inverse retraction.

If we denote the inclusion of $\mathcal{D}$ in TQ by $i_{\mathcal{D}} : \mathcal{D} \hookrightarrow TQ$, we induce the dual projection $i^*_{\mathcal{D}} : T^*Q \to \mathcal{D}^*$ defined by

$$\langle i^*_{\mathcal{D}}(\mu_q), v_q \rangle = \langle \mu_q, i_{\mathcal{D}}(v_q) \rangle, \quad \mu_q \in T^*_qQ, \ v_q \in \mathcal{D}_q.$$

The Legendre transformations of the Lagrangian functions $L : TQ \to \mathbb{R}$ and $l = L|_{\mathcal{D}} : \mathcal{D} \to \mathbb{R}$ satisfy the following relation

$$i^*_{\mathcal{D}} \circ \mathbb{F}L \circ i_{\mathcal{D}} = \mathbb{F}l, \tag{7.2.1}$$

where $\mathbb{F}l : \mathcal{D} \to \mathcal{D}^*$ is the restricted Legendre transformation defined from l

$$\langle \mathbb{F}l(u_q), v_q \rangle = \left.\frac{d}{dt}\right|_{t=0} l(u_q + tv_q), \quad u_q, v_q \in \mathcal{D}_q.$$

Definition 7.2.1. The maps $\mathbb{F}^{\pm}_{h,nh}l : \mathcal{M}^{e,nh}_h \to \mathcal{D}^*$ defined by

$$\begin{aligned}\mathbb{F}^{-}_{h,nh}l(q_0,q_1) &= \mathbb{F}l \circ R^{e-}_{h,nh}(q_0,q_1) \in \mathcal{D}^*_{q_0}\\ \mathbb{F}^{+}_{h,nh}l(q_0,q_1) &= \mathbb{F}l \circ R^{e+}_{h,nh}(q_0,q_1) \in \mathcal{D}^*_{q_1}.\end{aligned}$$

are said to be the *exact discrete nonholonomic Legendre transformations*.

Note that $\mathbb{F}^{\pm}_{h,nh}l$ are (local) diffeomorphisms, since they are the composition of (local) diffeomorphisms.

As we will see below, the condition of momentum matching gives the *exact discrete nonholonomic equations*:

$$\begin{gathered}\mathbb{F}^{+}_{h,nh}l(q_0,q_1) = \mathbb{F}^{-}_{h,nh}l(q_1,q_2)\\ (q_0,q_1),(q_1,q_2) \in \mathcal{M}^{e,nh}_h.\end{gathered} \tag{7.2.2}$$

We will see in a theorem below that the term “exact” is not naïve.

Remark 7.2.2. Alternatively we can define the subset

$$S^e_{nh} = \{(\mathbb{F}l \circ R^{e-}_{h,nh}(q_0,q_1), \mathbb{F}l \circ R^{e+}_{h,nh}(q_0,q_1)) \mid (q_0,q_1) \in \mathcal{M}^{e,nh}_h\}$$

and we can think $S^e_{nh} \subset \mathcal{D}^* \times \mathcal{D}^*$ as an implicit difference equation [IP+13] producing the exact discrete nonholonomic dynamics.

Observe that, since both $R^{e-}_{h,nh}$ and $\mathbb{F}l$ are local diffeomorphisms, then equations (7.2.2) implicitly define an *exact discrete flow*: $\Phi^e_{h,nh} : \mathcal{M}^{e,nh}_h \to \mathcal{M}^{e,nh}_h$ by

$$\Phi^e_{h,nh}(q_0,q_1) = \exp^{\Gamma_{(L,\mathcal{D})}}_h \circ R^{e+}_{h,nh}(q_0,q_1). \tag{7.2.3}$$

Moreover, it produces a well-defined flow on $\mathcal{D}^*$, denoted by $\varphi^e_{h,nh} : \mathcal{D}^* \to \mathcal{D}^*$, which is defined by

$$\varphi^e_{h,nh}(\mu_{q_0}) = \mathbb{F}^{+}_{h,nh}l \circ (\mathbb{F}^{-}_{h,nh}l)^{-1}(\mu_{q_0}), \quad \mu_{q_0} \in \mathcal{D}^*_{q_0}.$$

The interplay between both discrete flows and the nonholonomic Legendre transformations may be summarized in the following commutative diagram

Next, we state the main theorem in this section: the exact discrete nonholonomic flow, as the name indicates, exactly reproduces the continuous flow of the nonholonomic system at any step h.

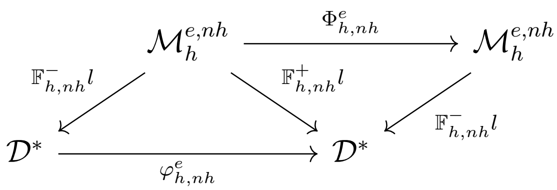


Figure 7.1: Commutative diagram: Exact discrete and continuous nonholonomic flows.

Theorem 7.2.3. *Given* $(q_0, q_1) \in \mathcal{M}_h^{e,nh}$ *and* $h > 0$*, consider the sequence* $(q_0, q_1, ..., q_N)$ *satisfying the exact discrete nonholonomic equations* (7.2.2)*, in the sense that*

$$\mathbb{F}^+_{h,nh}l(q_{k-1}, q_k) - \mathbb{F}^-_{h,nh}l(q_k, q_{k+1}) = 0, \quad (q_k, q_{k+1}) \in \mathcal{M}_h^{e,nh}, \tag{7.2.4}$$

for $0 \leqslant k \leqslant N-1$*.*

Then, we have that the sequence $(q_0, q_1, ..., q_N)$ *exactly matches the trajectories of* $\Gamma_{(L,\mathcal{D})}$ *in the sense that*

$$q_k = q_{0,1}(kh), \tag{7.2.5}$$

where $q_{0,1}$ *is the unique trajectory of* $\Gamma_{(L,\mathcal{D})}$ *satisfying* $q_{0,1}(0) = q_0$ *and* $q_{0,1}(h) = q_1$*.*

Proof. The theorem is a direct consequence of the definition of the exact discrete flow in (7.2.3).

□

In order to relate the exact discrete nonholonomic equations with the modified Lagrange-d'Alembert equations, which is the objective of the next section, we will need another alternative expression of equations (7.2.4). In particular, using (7.2.1) we can rewrite these equations in a very similar way to the modified Lagrange-d'Alembert equations given by equations (7.1.2) as

$$\begin{aligned} i_D^*\left((\mathbb{F}L \circ i_{\mathcal{D}} \circ R^{e+}_{h,nh})(q_0, q_1) - (\mathbb{F}L \circ i_{\mathcal{D}} \circ R^{e-}_{h,nh})(q_1, q_2)\right) &= 0 \\ (q_0, q_1), (q_1, q_2) &\in \mathcal{M}_h^{e,nh}. \end{aligned}$$

Note that the projection $i_{\mathcal{D}}^* : T^*Q \to \mathcal{D}^*$ satisfies

$$\ker(i_{\mathcal{D}}^*) = \mathcal{D}^o. \tag{7.2.6}$$

Thus, we conclude that

$$
\begin{aligned}
(\mathbb{F}L \circ R^{e+}_{h,nh}(q_0,q_1) - \mathbb{F}L \circ R^{e-}_{h,nh})(q_1,q_2) \in \mathcal{D}^o_{q_1} \\
(q_0,q_1),(q_1,q_2) \in \mathcal{M}^{e,nh}_h,
\end{aligned}
\tag{7.2.7}
$$

where we omit $i_{\mathcal{D}}$ since $R^{e+}_{h,nh}(q_0,q_1)$ and $R^{e-}_{h,nh}(q_1,q_2)$ are vectors in the distribution $\mathcal{D}$ and may be identified with its inclusion.

7.3 The exact discrete Lagrangian function

Given a regular nonholonomic system determined by $(L,\mathcal{D})$, we have seen how to derive the nonholonomic force $F_{nh}: \mathcal{D} \to T^*Q$ in Section 3.6 (see equation (3.6.8)).

Consider now an arbitrary extension $\widetilde{F_{nh}}: TQ \to T^*Q$ of F_{nh}. It is clear that the solutions of the forced system determined by $(L,\widetilde{F_{nh}})$ with initial conditions in $\mathcal{D}$, remain in $\mathcal{D}$ and match the trajectories of the nonholonomic system. In fact, if $\Gamma_{(L,\mathcal{D})}$ is the nonholonomic dynamics and $\Gamma_{(L,\widetilde{F_{nh}})}$ is the forced dynamics, then it is clear that $\Gamma_{(L,\mathcal{D})} = \Gamma_{(L,\widetilde{F_{nh}})}|_{\mathcal{D}}$.

If $R^{e-}_{h,\widetilde{F_{nh}}}$ is the exact retraction associated with the forced SODE $\Gamma_{(L,\widetilde{F_{nh}})}$ then, as in Section 3.7.7, we may define the exact discrete Lagrangian function $L^{e,h}_{d,\widetilde{F_{nh}}}: Q \times Q \to \mathbb{R}$ given by

$$
L^{e,h}_{d,\widetilde{F_{nh}}}(q_0,q_1) = \int_0^h \left(L \circ \phi_t^{\Gamma_{(L,\widetilde{F_{nh}})}} \circ R^{e-}_{h,\widetilde{F_{nh}}} \right)(q_0,q_1)\, dt,
$$

and the exact discrete forces $(\widetilde{F_{nh}})^{e,\pm}_d : Q \times Q \to T^*Q$ given by

$$
\begin{aligned}
\langle (\widetilde{F_{nh}})^{e,+}_d(q_0,q_1), X_{q_1} \rangle &= \langle F^{e,h}_d(q_0,q_1), (0_{q_0}, X_{q_1}) \rangle \\
\langle (\widetilde{F_{nh}})^{e,-}_d(q_0,q_1), X_{q_0} \rangle &= \langle F^{e,h}_d(q_0,q_1), (X_{q_0}, 0_{q_1}, \rangle,
\end{aligned}
$$

where $F^{e,h}_d : Q \times Q \to T^*(Q \times Q)$ is the double exact discrete force given by

$$
\langle F^{e,h}_d(q_0,q_1), (X_{q_0}, X_{q_1}) \rangle = \int_0^h \langle \left(\widetilde{F_{nh}} \circ \phi_t^{\Gamma_{(L,\widetilde{F_{nh}})}} \circ R^{e-}_{h,\widetilde{F_{nh}}} \right)(q_0,q_1), X_{0,1}(t) \rangle\, dt
$$

where $X_{0,1}(t) = T_{(q_0,q_1)}(\tau_Q \circ \phi_t^{\Gamma_{(L,\widetilde{F_{nh}})}} \circ R^{e-}_{h,\widetilde{F_{nh}}})(X_{q_0}, X_{q_1})$, for $(X_{q_0}, X_{q_1}) \in T_{q_0}Q \times T_{q_1}Q$. Note that $(\widetilde{F_{nh}})^{e,+}_d(q_0,q_1) \in T^*_{q_1}Q$ and $(\widetilde{F_{nh}})^{e,-}_d(q_0,q_1) \in T^*_{q_0}Q$.

Following the notation in [MW01], we may rewrite these maps as

$$\begin{aligned}
L^{e,h}_{d,\widetilde{F_{nh}}}(q_0,q_1) &= \int_0^h L(q_{0,1}(t),\dot q_{0,1}(t))\ dt\ ,\\
(\widetilde{F_{nh}})^{e,+}_d(q_0,q_1) &= \int_0^h \langle(\widetilde{F_{nh}})(q_{0,1}(t),\dot q_{0,1}(t)),\frac{\partial q_{0,1}(t)}{\partial q_1}\rangle\ dt\ ,\\
(\widetilde{F_{nh}})^{e,-}_d(q_0,q_1) &= \int_0^h \langle(\widetilde{F_{nh}})(q_{0,1}(t),\dot q_{0,1}(t)),\frac{\partial q_{0,1}(t)}{\partial q_0}\rangle\ dt\ .
\end{aligned}$$

where now $q_{0,1}:[0,h]\to Q$ is the solution of the forced Euler-Lagrange equations for $(L,\widetilde{F_{nh}})$ verifying $q_{0,1}(0)=q_0$ and $q_{0,1}(h)=q_1$.

We now prove that when we apply the modified Lagrange-d'Alembert principle to the exact discrete objects defined above, we obtain the exact discrete nonholonomic equations.

Theorem 7.3.1. *Let $(L,\mathcal{D})$ be a regular continuous-time nonholonomic problem with regular Lagrangian L. Consider the exact discrete Lagrangian function $L^{e,h}_{d,\widetilde{F_{nh}}}$ defined above, as well as the exact discrete forces $(\widetilde{F_{nh}})^{e,-}_d$ and $(\widetilde{F_{nh}})^{e,+}_d$. Also let $\mathcal{M}^{e,nh}_h$ be the exact discrete space associated to $(L,\mathcal{D})$. Then the modified Lagrange-d'Alembert principle induces modified Lagrange-d'Alembert equations*

$$\begin{aligned}
&D_2L^{e,h}_{d,\widetilde{F_{nh}}}(q_{k-1},q_k)+D_1L^{e,h}_{d,\widetilde{F_{nh}}}(q_k,q_{k+1})\\
&+(\widetilde{F_{nh}})^{e,+}_d(q_{k-1},q_k)+(\widetilde{F_{nh}})^{e,-}_d(q_k,q_{k+1})\in\mathcal{D}^o_{q_k},\\
&(q_k,q_{k+1})\in\mathcal{M}^{e,nh}_h,\ 0\leqslant k\leqslant N-1,
\end{aligned}\tag{7.3.1}$$

which are equivalent to the exact discrete nonholonomic equations (7.2.2).

Proof. The terms appearing in equations (7.3.1) are the restriction to $\mathcal{M}^{e,nh}_h$ of the exact discrete Legendre transformations for the forced system $(L,\widetilde{F_{nh}})$:

$$\begin{aligned}
&\mathbb{F}^{f+}L^{e,h}_{d,\widetilde{F_{nh}}}(q_{k-1},q_k)-\mathbb{F}^{f-}L^{e,h}_{d,\widetilde{F_{nh}}}(q_k,q_{k+1})\in\mathcal{D}^o_{q_k}\\
&(q_k,q_{k+1})\in\mathcal{M}^{e,nh}_h,\ 0\leqslant k\leqslant N-1.
\end{aligned}$$

Thus, using Lemma 3.7.15, the equations above are equivalent to

$$\begin{aligned}
&\mathbb{F}L\circ R^{e,+}_{h,\widetilde{F_{nh}}}(q_{k-1},q_k)-\mathbb{F}L\circ R^{e,-}_{h,\widetilde{F_{nh}}}(q_k,q_{k+1})\in\mathcal{D}^o_{q_k}\\
&(q_k,q_{k+1})\in\mathcal{M}^{e,nh}_h,\ 0\leqslant k\leqslant N-1.
\end{aligned}\tag{7.3.2}$$

Observe that, since the restriction of the forced dynamics to $\mathcal{D}$ matches the nonholonomic dynamics, then also the restriction of the forced inverse retraction maps to $\mathcal{M}_h^{e,nh}$ matches the nonholonomic retraction maps $R_{h,nh}^{e,\pm}$. Thus, projecting the first expression of (7.3.2) to $\mathcal{D}^*$ through $i_{\mathcal{D}}^*$ and using (7.2.6), we obtain

$$
\begin{aligned}
&i_{\mathcal{D}}^* \left(\mathbb{F}L|_{\mathcal{D}} \circ R_{h,nh}^{e,+}(q_{k-1}, q_k) - \mathbb{F}L|_{\mathcal{D}} \circ R_{h,nh}^{e,-}(q_k, q_{k+1})\right) = 0 \\
&(q_k, q_{k+1}) \in \mathcal{M}_h^{e,nh}, \ 0 \leqslant k \leqslant N-1
\end{aligned}
$$

and so, using (7.2.1), we obtain

$$
\mathbb{F}l \circ R_{h,nh}^{e,+}(q_{k-1}, q_k) - \mathbb{F}l \circ R_{h,nh}^{e,-}(q_k, q_{k+1}) = 0
$$

Now, if the sequence $(q_0, ..., q_N)$ satisfies (7.3.1), then since $\mathbb{F}l$ is a diffeomorphism, we deduce that

$$
\begin{aligned}
&R_{h,nh}^{e,+}(q_{k-1}, q_k) = R_{h,nh}^{e,-}(q_k, q_{k+1}) \\
&(q_k, q_{k+1}) \in \mathcal{M}_h^{e,nh}, \ 0 \leqslant k \leqslant N-1,
\end{aligned}
$$

and therefore the exact discrete nonholonomic equations (7.2.2) are satisfied.

Conversely, if the sequence $(q_0, ..., q_N)$ satisfies the exact discrete nonholonomic equations (7.2.2), we may reverse the argument and obtain (7.3.1). Note that, by (7.2.6), we have that

$$
i_{\mathcal{D}}^* \left(\mathbb{F}L|_{\mathcal{D}} \circ R_{h,nh}^{e,+}(q_{k-1}, q_k) - \mathbb{F}L|_{\mathcal{D}} \circ R_{h,nh}^{e,-}(q_k, q_{k+1})\right) = 0
$$

if and only if

$$
\mathbb{F}L|_{\mathcal{D}} \circ R_{h,nh}^{e,+}(q_{k-1}, q_k) - \mathbb{F}L|_{\mathcal{D}} \circ R_{h,nh}^{e,-}(q_k, q_{k+1}) \in \mathcal{D}_{q_k},
$$

from where the conclusion follows. □

Observe that, we are restricting to pairs of points in $\mathcal{M}_d^{e,nh}$ and applying the modified Lagrange-d'Alembert principle

$$
\begin{aligned}
&\delta S_d(q_d) \cdot \delta q_d + \sum_{k=1}^{N-1} \left[(\widetilde{F_{nh}})_d^{e,+}(q_{k-1}, q_k) + (\widetilde{F_{nh}})_d^{e,-}(q_k, q_{k+1})\right] \delta q_k = 0 \\
&(q_k, q_{k+1}) \in \mathcal{M}_d^{e,nh},
\end{aligned}
$$

with $\delta q_d = (\delta q_0, ..., \delta q_N)$ for all variations $\delta q_k \in \mathcal{D}_{q_k}$ verifying $\delta q_0 = \delta q_N = 0$ and

$$
S_d(q_d) = \sum_{k=0}^{N-1} L_{d,\widetilde{F_{nh}}}^{e,h}(q_k, q_{k+1}).
$$

7.4 Towards an intrinsic version of the exact discrete nonholonomic equations

Assume again that we have a regular nonholonomic system $(L, \mathcal{D})$. We can introduce the intrinsic version of the exact discrete objects appearing in the exact discrete nonholonomic equations. That is, instead of defining the various maps in the ambient manifold $Q \times Q$, we wish to define them directly on $\mathcal{M}_h^{e,nh}$.

With the help of the *constrained exact inverse retraction*, defined by $R_{h,nh}^{e-} : \mathcal{M}_h^{e,nh} \to \mathcal{U}_h \subseteq \mathcal{D}$ introduced in Section 4.3, we define the *nonholonomic exact discrete Lagrangian* as a function on the exact discrete space $l_{h,nh}^e : \mathcal{M}_h^{e,nh} \to \mathbb{R}$ given by

$$l_{h,nh}^e(q_0, q_1) = \int_0^h \left(L \circ \phi_t^{\Gamma_{(L,\mathcal{D})}} \circ R_{h,nh}^{e-} \right)(q_0, q_1)\, dt. \tag{7.4.1}$$

where $\{\phi_t^{\Gamma_{(L,\mathcal{D})}}\}$ is the flow of $\Gamma_{(L,\mathcal{D})}$.

To ease the notation let us introduce the following objects:

1. given $(q_0, q_1) \in \mathcal{M}_h^{e,nh}$, define the following curves on $\mathcal{D}$ and Q, respectively:

$$\gamma_0(t) := \left(\phi_t^{\Gamma_{(L,\mathcal{D})}} \circ R_{h,nh}^{e-} \right)(q_0, q_1) \text{ and } c_0(t) := \tau_Q \circ \gamma_0(t);$$

2. a variation of the former curve is denoted by

$$\gamma_s(t) = \left(\phi_t^{\Gamma_{(L,\mathcal{D})}} \circ R_{h,nh}^{e-} \right)(q_0(s), q_1(s)) \text{ and } c_s(t) := \tau_Q \circ \gamma_s(t)$$

3. the infinitesimal variation vector field on the configuration manifold is

$$X_{0,1}(t) = \left.\frac{d}{ds}\right|_{s=0} c_s(t).$$

Next, we will prove a result which we will use later. The proof of this result involves the canonical involution $\kappa_Q : TTQ \to TTQ$ of the double tangent bundle defined on Section 2.4.1.

Lemma 7.4.1. *Given a SODE Γ, if γ_s is a one-parameter family of integral curves of Γ, then the infinitesimal variation vector field of γ_s is the complete lift of the infinitesimal variation vector field of the one-parameter family of curves formed by the base integral curves of Γ, that is $c_s = \tau_Q \circ \gamma_s$.*

Proof. If γ_s is a one-parameter family of integral curves of Γ, it has the form $\gamma_s = \frac{d}{dt}c_s$. Let

$$X_{01}(t) = \left.\frac{d}{ds}\right|_{s=0} \tau_Q \circ \gamma_s(t)$$

be the infinitesimal variation vector field of c_s. Then the infinitesimal variation vector field of γ_s is

$$\begin{aligned} \left.\frac{d}{ds}\right|_{s=0} \gamma_s(t) &= \left.\frac{d}{ds}\right|_{s=0} \frac{dc_s}{dt}(t) \\ &= \kappa_Q \left(\frac{d}{dt} \left.\frac{d}{ds}\right|_{s=0} c(s,t) \right) = \kappa_Q \left(\frac{d}{dt} X_{01}(t) \right) = X_{01}^{\mathbf{c}}(t). \end{aligned}$$

□

Next, we will obtain an interesting expression for the differential of the nonholonomic exact discrete Lagrangian function $l^e_{h,nh}$.

Proposition 7.4.2. *The differential of the nonholonomic exact discrete Lagrangian satisfies*

$$\begin{aligned} \langle dl^e_{h,nh}(q_0,q_1), (X_{q_0}, X_{q_1}) \rangle &= -\langle \beta_{nh}(q_0,q_1), (X_{q_0}, X_{q_1}) \rangle \\ &\quad + \langle \mathbb{F}L \circ R^{e+}_{h,nh}(q_0,q_1), X_{q_1} \rangle - \langle \mathbb{F}L \circ R^{e-}_{h,nh}(q_0,q_1), X_{q_0} \rangle, \end{aligned}$$

where

$$\langle \beta_{nh}(q_0,q_1), (X_{q_0}, X_{q_1}) \rangle = \int_0^h \langle F_{nh}(\gamma_0(t)), X_{01}(t) \rangle \, dt$$

and we are identifying the vector $(X_{q_0}, X_{q_1}) \in T_{(q_0,q_1)}\mathcal{M}_h^{e,nh}$ with its image by $Ti : T\mathcal{M}_h^{e,nh} \hookrightarrow T(Q \times Q)$, $i : \mathcal{M}_h^{e,nh} \hookrightarrow Q \times Q$ being the canonical inclusion. The smooth curve $X_{01} : [0,h] \to TQ$ is defined as

$$X_{01}(t) = T_{(q_0,q_1)}(\tau_Q \circ \phi_t^{\Gamma_{(L,\mathcal{D})}} \circ R^{e-}_{h,nh})(X_{q_0}, X_{q_1}) \; .$$

Proof. Let $v : (-s, s) \to \mathcal{M}_h^{e,nh}$ be a smooth curve denoted by $v(s) = (q_0(s), q_1(s))$ such that $v(0) = (q_0, q_1) \in \mathcal{M}_h^{e,nh}$ and $v'(0) = (X_{q_0}, X_{q_1}) \in T_{(q_0,q_1)}\mathcal{M}_h^{e,nh}$ and

$$\gamma_s(t) = \left(\phi_t^{\Gamma_{(L,\mathcal{D})}} \circ R_{h,nh}^{e-}\right)(q_0(s), q_1(s)).$$

Then, using Lemma 7.4.1, we have that

$$\begin{aligned} \langle dl_{h,nh}^e(q_0, q_1), \left.\frac{d}{ds}\right|_{s=0} (q_0(s), q_1(s))\rangle = \\ &= \int_0^h \langle dL(\gamma_0(t)), \left.\frac{d}{ds}\right|_{s=0} \gamma_s(t)\rangle dt \\ &= \int_0^h \langle dL(\gamma_0(t)), X_{01}^{\mathbf{c}}(t)\rangle dt. \end{aligned} \tag{7.4.2}$$

Note that $X_{01}^{\mathbf{c}}(t)$ is a vector field on TQ along $\gamma_0(t)$, hence using (3.4.4), it follows that

$$\begin{aligned} \langle dl_{h,nh}^e(q_0, q_1), (X_{q_0}, X_{q_1})\rangle &= X_{01}^{\mathbf{v}}(h)(L) - X_{01}^{\mathbf{v}}(0)(L) - \int_0^h \langle F_{nh}(\gamma_0(t)), X_{01}(t)\rangle dt \\ &= \langle \mathbb{F}L(\gamma_0(h)), X_{01}(h)\rangle - \langle \mathbb{F}L(\gamma_0(0)), X_{01}(0)\rangle - \int_0^h \langle F_{nh}(\gamma_0(t)), X_{01}(t)\rangle dt. \end{aligned} \tag{7.4.3}$$

By unwinding the definition of X_{01} and identifying (X_{q_0}, X_{q_1}) with its image by $Ti : T\mathcal{M}_h^{e,nh} \hookrightarrow T(Q \times Q)$, we see that

$$\begin{aligned} X_{01}(h) &= T_{(q_0,q_1)}(\tau_Q \circ R_{h,nh}^{e+})(X_{q_0}, X_{q_1}) = X_{q_1}, \\ X_{01}(0) &= T_{(q_0,q_1)}(\tau_Q(R_{h,nh}^{e-})(X_{q_0}, X_{q_1}) = X_{q_0}, \end{aligned}$$

since

$$\tau_Q \circ R_{h,nh}^{e+} = \mathrm{pr}_2|_{\mathcal{M}_h^{e,nh}} \quad \text{and} \quad \tau_Q \circ R_{h,nh}^{e-} = \mathrm{pr}_1|_{\mathcal{M}_h^{e,nh}},$$

where $\mathrm{pr}_{1,2} : Q \times Q \to Q$ are the projection onto the first and second factor, respectively. □

Observe that in the previous Proposition, the intrinsic discrete objects associated to the nonholonomic problem are dl_h^e, $\beta_{nh} \in \Lambda^1\mathcal{M}_h^{e,nh}$. Then, σ_{nh} given by

$$\sigma_{nh}(X_{q_0}, X_{q_1}) = \langle(\mathbb{F}L \circ R_{h,nh}^{e+})(q_0, q_1), X_{q_1}\rangle - \langle(\mathbb{F}L \circ R_{h,nh}^{e-})(q_0, q_1), X_{q_0}\rangle \tag{7.4.4}$$

is also a 1-form in $\mathcal{M}_h^{e,nh}$, where (X_{q_0}, X_{q_1}) is identified with its image by Ti. From the definition of the Legendre transform $\mathbb{F}L : TQ \to T^*Q$, it is easy to see that this map can be extended to a map

$$\widetilde{\sigma_{nh}} : \mathcal{M}_h^{e,nh} \longrightarrow T^*(Q \times Q)$$

defined by expression (7.4.4) but applying it to an arbitrary vector $(X_{q_0}, X_{q_1}) \in T_{(q_0,q_1)}(Q \times Q)$ with $(q_0, q_1) \in \mathcal{M}_h^{e,nh}$.

Finally, we will relate the exact discrete objects associated with the forced system $(L, \widetilde{F_{nh}})$ and the intrinsic exact discrete objects we have just defined.

Proposition 7.4.3. *The restriction to $\mathcal{M}_h^{e,nh}$ of the forced exact discrete Lagrangian function $L^{e,h}_{d,\widetilde{F_{nh}}}$ is just the non-holonomic exact discrete Lagrangian function $l^e_{h,nh}$, that is,*

$$L^{e,h}_{d,\widetilde{F_{nh}}}\Big|_{\mathcal{M}_h^{e,nh}} = l^e_{h,nh}.$$

Moreover, if $(q_0, q_1) \in \mathcal{M}_h^{e,nh}$ and $(X_{q_0}, X_{q_1}) \in T_{(q_0,q_1)}\mathcal{M}_h^{e,nh}$ then

$$\langle ((\widetilde{F_{nh}})_d^{e,-}(q_0, q_1), (\widetilde{F_{nh}})_d^{e,+}(q_0, q_1)), (X_{q_0}, X_{q_1})\rangle = \langle \beta_{nh}(q_0, q_1), (X_{q_0}, X_{q_1})\rangle.$$

Proof. Given a pair of points $(q_0, q_1) \in \mathcal{M}_h^{e,nh}$, since the unique trajectory of $\Gamma_{(L,\mathcal{D})}$ connecting the two points is also the unique trajectory of the forced problem $(L, \widetilde{F_{nh}})$ connecting these points, the expressions of $L^{e,h}_{d,\widetilde{F_{nh}}}\Big|_{\mathcal{M}_h^{e,nh}}$ and $l^e_{h,nh}$ match.

According to Proposition 7.4.2 and the observations following it we have that

$$dl^e_{h,nh} + \beta_{nh} = \sigma_{nh}.$$

Then, since $\sigma_{nh} = i^*\widetilde{\sigma}_{nh}$ we have that

$$i^* dL^{e,h}_{d,\widetilde{F_{nh}}} + \beta_{nh} = i^*\widetilde{\sigma}_{nh},$$

where $i : \mathcal{M}_h^{e,nh} \to Q \times Q$ is the inclusion. So,

$$\beta_{nh} = i^*(\widetilde{\sigma}_{nh} - dL^{e,h}_{d,\widetilde{F_{nh}}}).$$

Observe that

$$\widetilde{\sigma}_{nh} - dL^{e,h}_{d,\widetilde{F_{nh}}} = (-\mathbb{F}L \circ R^{e,-}_{h,nh} - D_1 L^{e,h}_{d,\widetilde{F_{nh}}}, \mathbb{F}L \circ R^{e,+}_{h,nh} - D_2 L^{e,h}_{d,\widetilde{F_{nh}}}).$$

Therefore, using Lemma 3.7.15, we conclude

$$\widetilde{\sigma}_{nh} - dL^{e,h}_{d,\widetilde{F_{nh}}} = ((\widetilde{F_{nh}})_d^{e,-}, (\widetilde{F_{nh}})_d^{e,+}).$$

□

7.5 Construction of integrators and numerical examples

To construct variational integrators we consider discretizations (L_d, F_d^-, F_d^+) of $(L^{e,h}_{d,\widetilde{F_{nh}}}, (\widetilde{F_{nh}})^{e,-}_d, (\widetilde{F_{nh}})^{e,+}_d)$ as a typical forced integrator and then we consider a discretization $\mathcal{M}_h^d$ of $\mathcal{M}_h^{e,nh}$ to derive the *modified discrete Lagrange-d'Alembert equations*:

$$\begin{aligned}&D_2L_d(q_{k-1},q_k)+D_1L_d(q_k,q_{k+1})+F_d^+(q_{k-1},q_k)+F_d^-(q_k,q_{k+1})\in\mathcal{D}^o_{q_k}\\&(q_k,q_{k+1})\in\mathcal{M}_h^d,\ 0\leqslant k\leqslant N-1,\end{aligned}\tag{7.5.1}$$

We remark that, by (7.2.6), we have that (7.5.1) is equivalent to the projection onto $\mathcal{D}^*$, i.e.,

$$\begin{aligned}&i^*_{\mathcal{D}}\left(D_2L_d(q_{k-1},q_k)+D_1L_d(q_k,q_{k+1})+F_d^+(q_{k-1},q_k)+F_d^-(q_k,q_{k+1})\right)=0\\&(q_k,q_{k+1})\in\mathcal{M}_h^d,\ 0\leqslant k\leqslant N-1,\end{aligned}\tag{7.5.2}$$

This projection motivates the definition of the Legendre transforms $\mathbb{F}^\pm l_d : \mathcal{M}_h^d \to \mathcal{D}^*$ given by

$$\begin{aligned}\mathbb{F}^+l_d&=i^*_{\mathcal{D}}\circ\mathbb{F}^{f+}L_d|_{\mathcal{M}_h^d}\\\mathbb{F}^-l_d&=i^*_{\mathcal{D}}\circ\mathbb{F}^{f-}L_d|_{\mathcal{M}_h^d}.\end{aligned}$$

Example 7.5.1. Consider once more the nonholonomic particle. We introduce a discretization of the discrete space $\mathcal{M}_h^{e,nh}$

$$\mathcal{M}_h^d=\{z_1=z_0+\left(\frac{y_1+y_0}{2}\right)(x_1-x_0)\},\tag{7.5.3}$$

and a discrete Lagrangian

$$L_d(x_0,y_0,z_0,x_1,y_1,z_1)=\frac{1}{2h}\left[(x_1-x_0)^2+(y_1-y_0)^2+(z_1-z_0)^2\right].$$

Moreover we need two discrete forces

$$F_d^+(q_0,q_1)=\frac{2}{h}\frac{(x_1-x_0)(y_1-y_0)}{4+(y_1+y_0)^2}\left(-\frac{y_1+y_0}{2}dx_1+dz_1\right)$$

and

$$F_d^-(q_0,q_1) = \frac{2}{h}\frac{(x_1-x_0)(y_1-y_0)}{4+(y_1+y_0)^2}\left(-\frac{y_1+y_0}{2}dx_0+dz_0\right).$$

The forced discrete Legendre transforms which appear also in the modified Lagrange-d'Alembert equations are

$$\begin{aligned}\mathbb{F}^{f-}L_d(q_0,q_1) = &\left(\frac{x_1-x_0}{h}+\frac{1}{h}\frac{(x_1-x_0)(y_1-y_0)(y_1+y_0)}{4+(y_1+y_0)^2}\right)dx_0 \\ &+\frac{y_1-y_0}{h}dy_0+\left(\frac{z_1-z_0}{h}-\frac{2}{h}\frac{(x_1-x_0)(y_1-y_0)}{4+(y_1+y_0)^2}\right)dz_0\end{aligned}$$

and

$$\begin{aligned}\mathbb{F}^{f+}L_d(q_0,q_1) = &\left(\frac{x_1-x_0}{h}-\frac{1}{h}\frac{(x_1-x_0)(y_1-y_0)(y_1+y_0)}{4+(y_1+y_0)^2}\right)dx_1 \\ &+\frac{y_1-y_0}{h}dy_1+\left(\frac{z_1-z_0}{h}+\frac{2}{h}\frac{(x_1-x_0)(y_1-y_0)}{4+(y_1+y_0)^2}\right)dz_1.\end{aligned}$$

Now projecting the forced Legendre transforms onto $\mathcal{D}^*$ by means of $i_{\mathcal{D}}^*$ and restricting to $\mathcal{M}_h^d$ we get

$$\mathbb{F}^-l_d(q_0^i,q_1^a) = \frac{x_1-x_0}{h}\left(1+\frac{1}{2}y_0(y_1+y_0)+\frac{(y_1-y_0)^2}{4+(y_1+y_0)^2}\right)e^1+\left(\frac{y_1-y_0}{h}\right)e^2$$

and

$$\mathbb{F}^+l_d(q_0^i,q_1^a) = \frac{x_1-x_0}{h}\left(1+\frac{1}{2}y_1(y_1+y_0)+\frac{(y_1-y_0)^2}{4+(y_1+y_0)^2}\right)e^1+\left(\frac{y_1-y_0}{h}\right)e^2,$$

where the local frame $\{e^a\}\subseteq\mathcal{D}^*$ is dual to the local frame $\{e_a\}$ spanning $\mathcal{D}$, where $e_1=\frac{\partial}{\partial x}+y\frac{\partial}{\partial z}$ and $e_2=\frac{\partial}{\partial y}$.

Now solving equations (7.5.2) for this example we get

$$\begin{aligned}x_2 &= x_1+(x_1-x_0)\frac{1+\frac{1}{2}y_1(y_1+y_0)+\frac{(y_1-y_0)^2}{4+(y_1+y_0)^2}}{1+\frac{1}{2}y_1(3y_1-y_0)+\frac{(y_1-y_0)^2}{4+(3y_1-y_0)^2}}\\ y_2 &= 2y_1-y_0.\end{aligned}$$

We can see in Figures 7.2 and 7.3 a comparison between the proposed integrator (MLA) and the more standard Discrete Lagrange-d'Alembert (DLA) integrator (see [CM01]), where we use the same discrete Lagrangian L_d, discrete forces $F_d^{\pm}$ and discrete space $\mathcal{M}_h^d$ than above. We compare the error in both integrators as well as the energy behaviour of both. We observe the proposed integrator as good behaviour in both aspects and it even behaves slightly better than DLA. Notice that the Hamiltonian function $H|_{\mathcal{D}^*}$ given by

$$H|_{\mathcal{D}^*}(x, y, z, p_1, p_2) = \frac{1}{2}\left(\frac{p_1^2}{1+y^2} + p_2^2\right)$$

becomes approximately constant along the discrete flow, after the first steps. To run the simulation we set the initial position at the origin $q_0 = 0$ and $q_1 = (0.4, 0.4, z_1)$, with z_1 being determined by (7.5.3). The step is $h = 0.5$ and the total number of steps is $N = 1200$.

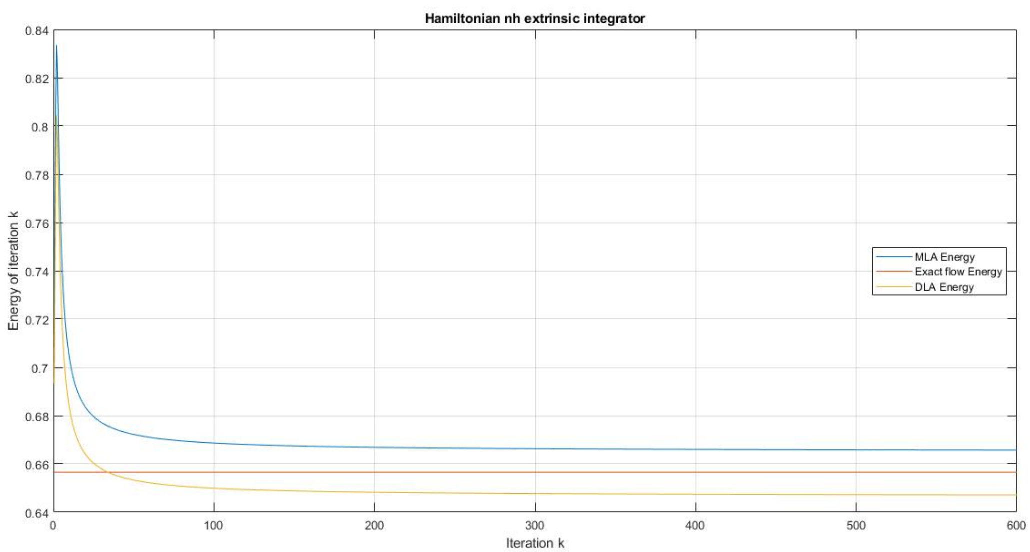


Figure 7.2: Comparison of the value of the Hamiltonian function between DLA and MLA integrators.

We also draw in Figure 7.4 the discrete constraint space $\mathcal{M}_h^d$ and compare it with its exact version $\mathcal{M}_h^{e,nh}$. △

Example 7.5.2. Let us introduce another typical example of nonholonomic system (see [Blo15]): the knife edge. Choosing appropriate constants, its

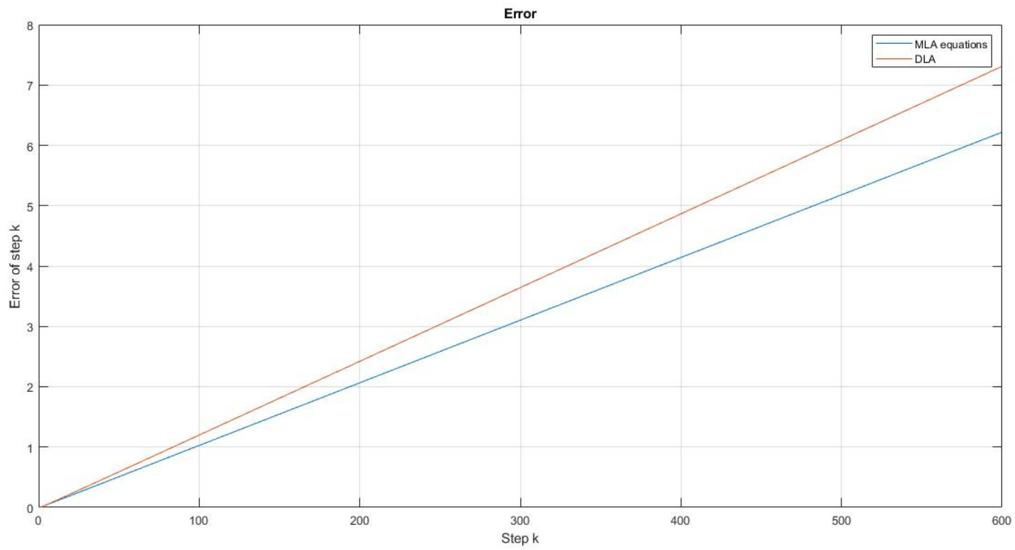
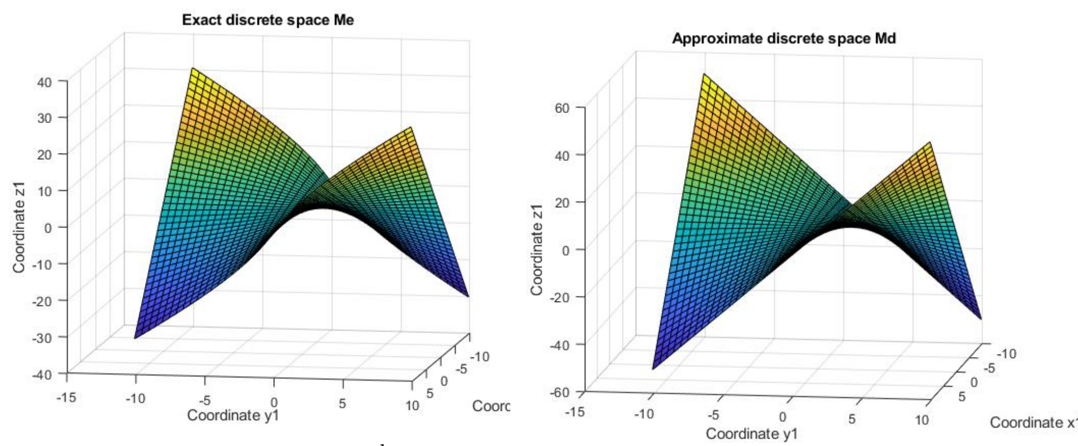


Figure 7.3: Evolution of the error in DLA and MLA integrators.



(a) Exact discrete space $\mathcal{M}_h^{e,nh}$ given by (4.3.4).

(b) Discrete space $\mathcal{M}_h^d$ given by (7.5.3).

Figure 7.4: Graph of the defining function for the respective spaces. We have fixed the origin as the initial point $q_0 = 0$ and plotted the coordinate z_1 as a function of x_1 and y_1.

Lagrangian function is described by the function $L : T(Q \times \mathbb{S}^1) \to \mathbb{R}$

$$L(x, y, \varphi, \dot{x}, \dot{y}, \dot{\varphi}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{\varphi}^2) + \frac{x}{2},$$

and it is subjected to the nonholonomic constraint

$$\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y} = 0.$$

We introduce the following discretization of the constraint space

$$\mathcal{M}_h^d = \left\{ \sin\left(\frac{\varphi_1 + \varphi_0}{2}\right) \frac{x_1 - x_0}{h} - \cos\left(\frac{\varphi_1 + \varphi_0}{2}\right) \frac{y_1 - y_0}{h} = 0 \right\}.$$

The natural discretization of the Lagrangian compatible with the above discrete constraint space is then

$$L_d(x_0, y_0, \varphi_0, x_1, y_1, \varphi_1) = \frac{1}{2h}((x_1 - x_0)^2 + (y_1 - y_0)^2 + (\varphi_1 - \varphi_0)^2) + h \cdot \frac{x_1 + x_0}{4}$$

Moreover the discrete forces are given by

$$F_d^+(q_0, q_1) = \frac{h}{2}\lambda\left(\mu_x dx_1 + \mu_y dy_1\right), \quad F_d^-(q_0, q_1) = \frac{h}{2}\lambda\left(\mu_x dx_0 + \mu_y dy_0\right),$$

with

$$\lambda = -\frac{\varphi_1 - \varphi_0}{h^2}\left((x_1 - x_0)\cos\left(\frac{\varphi_1 + \varphi_0}{2}\right) + (y_1 - y_0)\sin\left(\frac{\varphi_1 + \varphi_0}{2}\right)\right) - \frac{1}{2}\sin\left(\frac{\varphi_1 + \varphi_0}{2}\right)$$

and

$$\mu_x = \sin\left(\frac{\varphi_1 + \varphi_0}{2}\right), \quad \mu_y = \cos\left(\frac{\varphi_1 + \varphi_0}{2}\right).$$

With these ingredients we obtained an integrator with a nearly preservation of the energy (see Figure 7.5), where we use the Hamiltonian function

$$H|_{\mathcal{D}^*}(x, \varphi, y, p_1, p_2) = \frac{1}{2}\left(\frac{p_1^2}{A(\varphi)} + p_2^2 - x\right), \quad A(\varphi) = 1 + \frac{\sin^2(\varphi)}{\cos^2(\varphi)}.$$

△

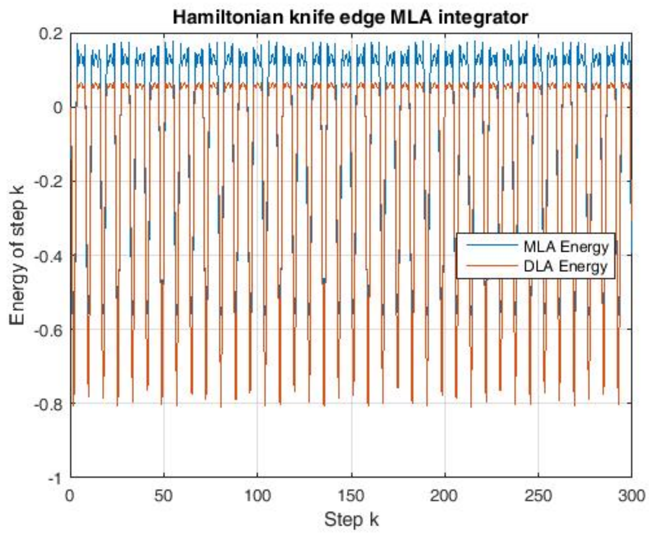


Figure 7.5: Experiment with the knife edge example: the initial positions are the origin $q_0 = 0$ and $q_1 = (0.4, 0.4, y_1)$, the step is $h = 0.5$ and the total number of steps is $N = 600$. As before, we use the same discrete Lagrangian L_d, discrete forces $F_d^{\pm}$ and discrete space $\mathcal{M}_h^d$ in both integrators.

Example 7.5.3. We now slightly perturb the knife edge system by introducing the nonholonomic constraint (see [MV19])

$$\sin(\varphi)\dot{x} - (\cos(\varphi) - \varepsilon)\dot{y} = 0, \quad \varepsilon > 0.$$

We obtain an integrator for the perturbed system that no longer preserves energy. Anyway, it still behaves clearly better than standard DLA algorithm (check Figure 7.6), for the Hamiltonian function

$$H|_{\mathcal{D}^*}(x, \varphi, y, p_1, p_2) = \frac{1}{2}\left(\frac{p_1^2}{A(\varphi, \varepsilon)} + p_2^2 - x\right), \quad A(\varphi, \varepsilon) = 1 + \frac{\sin^2(\varphi)}{(\cos(\varphi) - \varepsilon)^2}.$$

△

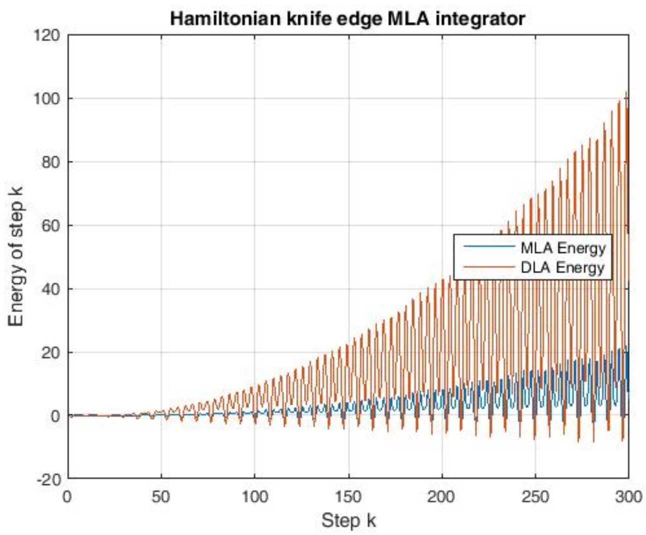


Figure 7.6: Experiment with the perturbed knife edge example with $\varepsilon = 0.1$: the initial positions are the origin $q_0 = 0$ and $q_1 = (0.4, 0.4, y_1)$, the step is $h = 0.5$ and the total number of steps is $N = 600$. As before, we use the same discrete Lagrangian L_d, discrete forces $F_d^{\pm}$ and discrete space $\mathcal{M}_h^d$ in both integrators.

Chapter 8

The geometry of discrete contact mechanics

In this chapter, we will introduce the discrete variational principle replacing Herglotz variational principle in the discrete setting. The discretization of Lagrangian contact systems is a very recent subject (see [VBS19] and also [Bra+20]). In fact, Lagrangian contact mechanics (see Section 3.5) was developed in the last few years (see [Bra17; Bra18; LL21; LLV19a]) and so it is natural to wonder if the constructions of discrete Lagrangian mechanics (see Section 3.7 and references there in) may be generalized to this context.

We will use the results described in Chapters 4 and 7 to define a contact exponential map and the discrete contact Lagrangian function, using the fact that Herglotz equations for contact Lagrangian systems on $TQ \times \mathbb{R}$ may be seen as the equations of motions of a nonholonomic Lagrangian system on $T(Q \times \mathbb{R})$ subjected to nonlinear constraints.

Section 8.1 is devoted to construct the discrete version of contact Lagrangian dynamics for a discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$, where Q is the configuration manifold. We consider the discrete Herglotz principle to obtain the so-called discrete Herglotz equations. The Legendre transformations $F^{-}L_d$ and $F^{+}L_d$ are defined, and consequently the discrete flow (at the Lagrangian and Hamiltonian levels); the main result is that the discrete flow is a conformal contactomorphism.

In the next section, we define the contact exponential map for the Herglotz vector field and prove that it is a local diffeomorphism, as an application of the nonholonomic exponential map which also also allow us to define the exact discrete contact Lagrangian function generating the exact discrete contact flow.

Finally, we consider several examples to illustrate our theoretical developments.

8.1 Discrete contact mechanics

Let $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$ be a discrete Lagrangian function. In our point of view $Q \times Q \times \mathbb{R}$ will be the discrete space corresponding to the manifold $TQ \times \mathbb{R}$, where continuous contact Lagrangian mechanics takes place. We fix a *time-step* $h > 0$, on which L_d depends, though we will omit this explicit dependence.

For each $N \in \mathbb{N}$, let us define the *discrete path space* as the space containing sequences on Q with length $N + 1$, i.e.,

$$\mathcal{C}_d^N(Q) = \{(q_0, q_1, ..., q_N) | q_k \in Q,\ k = 0, \ldots, N\}.$$

The set $\mathcal{C}_d^N(Q)$ is a manifold and it is canonically identified with the product space Q^{N+1}.

To each $q_d \in \mathcal{C}_d^N(Q)$ and each $z_0 \in \mathbb{R}$ we will associate another sequence $(z_k) \in \mathbb{R}^{N+1}$ defined by

$$z_{k+1} - z_k = L_d(q_k, q_{k+1}, z_k), \quad k = 0, ..., N - 1. \tag{8.1.1}$$

In the sequel, for each $1 \leqslant k \leqslant N$, we will denote by $\mathcal{Z}_k$ the function $\mathcal{Z}_k : Q \times Q \times \mathbb{R} \longrightarrow \mathbb{R}$

$$\mathcal{Z}_k(q_{k-1}, q_k, z_{k-1}) = z_{k-1} + L_d(q_{k-1}, q_k, z_{k-1}).$$

We define the *contact discrete action* to be the functional that for each point $q_d \in \mathcal{C}_d^N(Q)$ and each real number z_0 returns as output the real number z_N obtained recursively from (8.1.1), i.e.,

$$\begin{aligned} \mathcal{A}_d :\ \mathcal{C}_d^N(Q) \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (q_d, z_0) &\mapsto z_N. \end{aligned} \tag{8.1.2}$$

A *variation* of a sequence $q_d \in \mathcal{C}_d^N(Q)$ is a curve $\widetilde{q}_d : (-\epsilon, \epsilon) \to \mathcal{C}_d^N(Q)$ satisfying $\widetilde{q}_d(0) = q_d$. Given such a variation, we will define its *infinitesimal variation* by

$$\delta q_d := \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \widetilde{q}_d(\epsilon) = (\delta q_0, ..., \delta q_N),$$

where $\delta q_k := \frac{d}{d\epsilon}\big|_{\epsilon=0} \widetilde{q}_k(\epsilon)$.

Proposition 8.1.1. *Let L_d be a smooth discrete Lagrangian. Then, if we fix $z_0 \in \mathbb{R}$, we obtain the functional*

$$\begin{aligned} \mathcal{A}_{d,z_0} :\ \mathcal{C}_d^N(Q) &\longrightarrow \mathbb{R} \\ q_d &\mapsto \mathcal{A}_d(q_d, z_0). \end{aligned}$$

The differential of the functional $\mathcal{A}_{d,z_0}$ is the following

$$\begin{aligned} d\mathcal{A}_{d,z_0}(q_d) =& \sigma_N \cdots \sigma_2 \frac{\partial \mathcal{Z}_1}{\partial q_0}(q_0, q_1, z_0) dq_0 \\ &+ \sum_{k=1}^{N-1} \prod_{j=k+2}^{N} \sigma_j \cdot \left(\frac{\partial \mathcal{Z}_{k+1}}{\partial q_k} + \sigma_{k+1} \frac{\partial \mathcal{Z}_k}{\partial q_k} \right) dq_k \\ &+ \frac{\partial \mathcal{Z}_N}{\partial q_N}(q_{N-1}, q_N, z_{N-1}) dq_N, \end{aligned} \tag{8.1.3}$$

where we are using the identification of $\mathcal{C}_d^N(Q)$ with Q^{N+1} and for each $1 \leqslant j \leqslant N$

$$\sigma_j = \frac{\partial \mathcal{Z}_j}{\partial z_{j-1}}(q_{j-1}, q_j, z_{j-1}) = 1 + D_z L_d(q_{j-1}, q_j, z_{j-1}).$$

Proof. Using the identification of $\mathcal{C}_d^N(Q)$ with Q^{N+1}, note that the discrete action may be rewritten as

$$\mathcal{A}_{d,z_0}(q_d) = \mathcal{Z}_N(q_{N-1}, q_N, \mathcal{Z}_{N-1}(q_{N-2}, q_{N-1}, \mathcal{Z}_{N-2}(...\mathcal{Z}_1(q_0, q_1, z_0)...)).$$

Using that

$$d\mathcal{A}_{d,z_0}(q_d) = \frac{\partial \mathcal{A}_{d,z_0}}{\partial q_0} dq_0 + \sum_{k=1}^{N-1} \frac{\partial \mathcal{A}_{d,z_0}}{\partial q_k} dq_k + \frac{\partial \mathcal{A}_{d,z_0}}{\partial q_N} dq_N.$$

and applying the chain rule, we deduce that

$$\frac{\partial \mathcal{A}_{d,z_0}}{\partial q_0} = \frac{\partial \mathcal{Z}_N}{\partial z_{N-1}} \cdots \frac{\partial \mathcal{Z}_2}{\partial z_1} \frac{\partial \mathcal{Z}_1}{\partial q_0},$$

since the function $\mathcal{Z}_1$ is the only one that depends on q_0 among all the N functions $\mathcal{Z}_k$. It is also clear that

$$\frac{\partial \mathcal{A}_{d,z_0}}{\partial q_N} = \frac{\partial \mathcal{Z}_N}{\partial q_N},$$

since none of the functions $\mathcal{Z}_k$ depend on q_N except the function $\mathcal{Z}_N$. Finally if $1 \leqslant k \leqslant N-1$ we have that

$$\frac{\partial \mathcal{A}_{d,z_0}}{\partial q_k} = \frac{\partial \mathcal{Z}_N}{\partial z_{N-1}} \cdots \frac{\partial \mathcal{Z}_{k+2}}{\partial z_{k+1}} \left(\frac{\partial \mathcal{Z}_{k+1}}{\partial q_k} + \frac{\partial \mathcal{Z}_{k+1}}{\partial z_k} \frac{\partial \mathcal{Z}_k}{\partial q_k} \right),$$

where we applied the chain rule and the fact that the functions $\mathcal{Z}_{k+1}$ and $\mathcal{Z}_k$ are the only ones that depend on q_k. Hence, we finished the proof. □

Remark 8.1.2. Let us see the special case $N = 2$, where we can directly compute the differential of the action.

Let L_d be a smooth discrete Lagrangian. In the case where $N = 2$, the differential of the discrete action function satisfies:

$$\begin{aligned} d\mathcal{A}_{d,z_0} = (D_1 L_d(q_1, q_2, z_1) + (1 + D_z L_d(q_1, q_2, z_1) D_2 L_d(q_0, q_1, z_0))\, dq_1 \\ + D_2 L_d(q_1, q_2, z_1) dq_2 + (1 + D_z L_d(q_1, q_2, z_1)) D_1 L_d(q_0, q_1, z_0) dq_0. \end{aligned} \tag{8.1.4}$$

Definition 8.1.3 (Discrete Herglotz Principle)**.** Given $z_0 \in \mathbb{R}$, a discrete path $q_d = (q_0, ..., q_N)$ in $\mathcal{C}_d^N(Q)$ is said to satisfy the *Discrete Herglotz Principle* if q_d is a critical value of the discrete action functional $\mathcal{A}_{d,z_0}$ among all paths in $\mathcal{C}_d^N(Q)$ with fixed end points q_0, q_N.

We will now obtain as a sufficient and necessary condition for a path to satisfy the discrete Herglotz principle, a set of equations called *Discrete Herglotz equations* [VBS19].

Theorem 8.1.4. *Let L_d be a discrete Lagrangian function such that $1+D_z L_d$ is non-vanishing everywhere. Given $z_0 \in \mathbb{R}$, a discrete path $q_d \in \mathcal{C}_d^N(Q)$ satisfies the discrete Herglotz principle if and only if it satisfies*

$$\begin{aligned} D_1 L_d(q_k, q_{k+1}, z_k) + (1 + D_z L_d(q_k, q_{k+1}, z_k)) D_2 L_d(q_{k-1}, q_k, z_{k-1}) = 0, \\ z_k - z_{k-1} = L_d(q_{k-1}, q_k, z_{k-1}), \end{aligned} \tag{8.1.5}$$

for $k = 1, ..., N-1$.

Proof. Let $q_d(\epsilon)$ be a variation of $q_d \in \mathcal{C}_d^N(Q)$ with fixed end-points q_0 and q_N. Then q_d is a critical value of the discrete action functional if and only if

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (\mathcal{A}_{d,z_0}(q_d(\epsilon))) = d\mathcal{A}_{d,z_0}(\delta q_d) = 0.$$

By (8.1.3) the last expression is equivalent to

$$\sum_{k=1}^{N-1} \prod_{j=k+2}^{N} \sigma_j \cdot \left(\frac{\partial \mathcal{Z}_{k+1}}{\partial q_k} + \frac{\partial \mathcal{Z}_{k+1}}{\partial z_k} \frac{\partial \mathcal{Z}_k}{\partial q_k} \right) \delta q_k = 0.$$

Since the infinitesimal variations δq_k, $1 \leq k \leq N-1$, are arbitrary we deduce

$$\prod_{j=k+2}^{N} \sigma_j \cdot \left(\frac{\partial \mathcal{Z}_{k+1}}{\partial q_k} + \frac{\partial \mathcal{Z}_{k+1}}{\partial z_k} \frac{\partial \mathcal{Z}_k}{\partial q_k} \right) = 0.$$

Note that,

$$\sigma_j = \frac{\partial \mathcal{Z}_j}{\partial z_{j-1}}(q_{j-1}, q_j, z_{j-1}) = 1 + D_z L_d(q_{j-1}, q_j, z_{j-1})$$

is non-vanishing by hypothesis and

$$\frac{\partial \mathcal{Z}_{k+1}}{\partial q_k} + \frac{\partial \mathcal{Z}_{k+1}}{\partial z_k} \frac{\partial \mathcal{Z}_k}{\partial q_k} = D_1 L_d(q_k, q_{k+1}, z_k) + \sigma_{k+1} D_2 L_d(q_{k-1}, q_k, z_{k-1}),$$

from where the result follows. □

Remark 8.1.5. The discrete principle introduced in [VBS19] is just the condition

$$\frac{\partial \mathcal{Z}_{k+1}}{\partial q_k} + \frac{\partial \mathcal{Z}_{k+1}}{\partial z_k} \frac{\partial \mathcal{Z}_k}{\partial q_k} = 0,$$

afer rewriting it in our notation. For discrete Lagrangian functions where $1 + D_z L_d$ is non-vanishing, the condition above is equivalent to the Herglotz discrete principle.

8.1.1 Discrete Lagrangian flows and discrete Legendre transformations

Given a discrete contact Lagrangian L_d, if $1+D_z L_d(q_0, q_1, z_0)$ does not vanish, we can define two maps called *discrete Legendre transformations*: $\mathbb{F}^{\pm} L_d : Q \times Q \times \mathbb{R} \to T^*Q \times \mathbb{R}$

$$\begin{aligned} \mathbb{F}^{+} L_d(q_0, q_1, z_0) &= (q_1, D_2 L_d(q_0, q_1, z_0), z_0 + L_d(q_0, q_1, z_0)) \\ \mathbb{F}^{-} L_d(q_0, q_1, z_0) &= \left(q_0, -\frac{D_1 L_d(q_0, q_1, z_0)}{1 + D_z L_d(q_0, q_1, z_0)}, z_0 \right). \end{aligned} \tag{8.1.6}$$

Lemma 8.1.6. *$\mathbb{F}^+L_d$ is a local diffeomorphism if and only if $\mathbb{F}^-L_d$ is a local diffeomorphism.*

Proof. It is a direct consequence of the implicit function theorem. □

The Legendre transformations allow us to rewrite discrete Herglotz equations (8.1.5) as a momentum matching equations as in [MW01]. Indeed, provided $1 + D_zL_d(q_0, q_1, z_0)$ is not zero, we may write the discrete Herglotz equations as

$$\mathbb{F}^+L_d(q_0, q_1, z_0) = \mathbb{F}^-L_d(q_1, q_2, z_1). \tag{8.1.7}$$

Inspired by the following theorem, we say that a discrete contact Lagrangian is *regular* if the function $1 + D_zL_d(q_0, q_1, z_0)$ does not vanish and its negative discrete Legendre transform $\mathbb{F}^-L_d$ is a local diffeomorphism. Thus, we have the following theorem

Theorem 8.1.7. *Suppose that the discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$ is regular. Then there is a well-defined discrete Lagrangian flow $\Phi_d : Q \times Q \times \mathbb{R} \to Q \times Q \times \mathbb{R}$ for the discrete Herglotz equations. Moreover Φ_d is a local diffeomorphism given by*

$$\Phi_d = (\mathbb{F}^-L_d)^{-1} \circ \mathbb{F}^+L_d.$$

Proof. Consider the points $(q_0, q_1, z_0) \in Q\times Q\times\mathbb{R}$ and $(q_1, q_2, z_1) \in Q\times Q\times\mathbb{R}$ satisfying equation (8.1.7). If $\mathbb{F}^-L_d$ is a local diffeomorphism, then the map defined by

$$\Phi_d = (\mathbb{F}^-L_d)^{-1} \circ \mathbb{F}^+L_d$$

is also a local diffeomorphism and satisfies

$$\Phi_d(q_0, q_1, z_0) = (q_1, q_2, z_1),$$

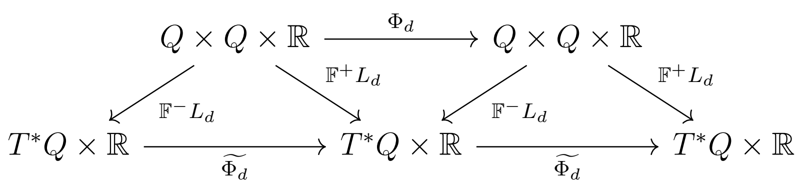
showing that it is the discrete Lagrangian flow for discrete Herglotz equations. □

The discrete Legendre transformations also allow us to define an associated *discrete Hamiltonian flow* on $T^*Q \times \mathbb{R}$. Indeed, considering a regular discrete Lagrangian function L_d, let $\widetilde{\Phi_d} : T^*Q \times \mathbb{R} \to T^*Q \times \mathbb{R}$ be defined by

$$\widetilde{\Phi_d} = \mathbb{F}^+L_d \circ \Phi_d \circ (\mathbb{F}^+L_d)^{-1}. \tag{8.1.8}$$

It is not difficult to show that the discrete Hamiltonian flow admits the alternative expressions

$$\widetilde{\Phi_d} = \mathbb{F}^- L_d \circ \Phi_d \circ (\mathbb{F}^- L_d)^{-1} \qquad \text{or} \qquad \widetilde{\Phi_d} = \mathbb{F}^+ L_d \circ (\mathbb{F}^- L_d)^{-1}. \tag{8.1.9}$$



(8.1.10)

We may define the one-forms

$$\eta^+ = (\mathbb{F}^+ L_d)^* \eta, \quad \eta^- = (\mathbb{F}^- L_d)^* \eta, \tag{8.1.11}$$

where η is the canonical contact form on $T^*Q \times \mathbb{R}$. These are contact forms on $Q \times Q \times \mathbb{R}$. If we chose natural coordinates (q^i, p_i, z) on $T^*Q \times \mathbb{R}$ where $\eta = dz - p_i dq^i$, the discrete 1-forms may be locally written as

$$\begin{aligned} \eta^+ &= dz_0 + dL_d(q_0, q_1, z_0) - D_2 L_d(q_0, q_1, z_0) dq_1, \\ \eta^- &= dz_0 + \frac{D_1 L_d(q_0, q_1, z_0)}{1 + D_z L_d(q_0, q_1, z_0)} dq_0. \end{aligned} \tag{8.1.12}$$

The one-form η^+ is further simplified to

$$\eta^+ = (1 + D_z L_d(q_0, q_1, z_0)) dz_0 + D_1 L_d(q_0, q_1, z_0) dq_0. \tag{8.1.13}$$

Given a discrete Lagrangian L_d, let $\sigma_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$ be the smooth function given by

$$\sigma_d(q_0, q_1, z_0) = 1 + D_z L_d(q_0, q_1, z_0)$$

then we have that:

Lemma 8.1.8. *The discrete contact forms $\eta^\pm$ satisfy*

(i) $\eta^+ = \sigma_d \cdot \eta^-$*;*

(ii) $(\Phi_d)^* \eta^- = \eta^+$.

Proof. For the first item, observe that (8.1.13) is equivalent to

$$\eta^+ = (1 + D_z L_d(q_0, q_1, z_0))\eta^-.$$

For the second one, note that

$$(\Phi_d)^*\eta^- = (\Phi_d)^* \circ (\mathbb{F}^- L_d)^*\eta = (\mathbb{F}^- L_d \circ \Phi_d)^*\eta = (\mathbb{F}^+ L_d)^*\eta$$

by applying Theorem 8.1.7. □

As a consequence of the last Lemma we have the following theorem:

Theorem 8.1.9. *Let L_d be a regular discrete Lagrangian function. The discrete flow Φ_d associated to L_d is a conformal contactomorphism with respect to both contact structures $\eta^\pm$. In particular, it satisfies*

$$(\Phi_d)^*\eta^+ = (\sigma_d \circ \Phi_d) \cdot \eta^+, \quad (\Phi_d)^*\eta^- = \sigma_d \cdot \eta^- \tag{8.1.14}$$

Likewise, the discrete Hamiltonian flow $\widetilde{\Phi_d}$ is also a conformal contactomorphism satisfying

$$(\widetilde{\Phi_d})^*\eta = (\sigma_d \circ (\mathbb{F}^- L_d)^{-1}) \cdot \eta. \tag{8.1.15}$$

Proof. The first two claims are trivial consequences of Lemma 8.1.8. Indeed, combining the two statements of the Lemma we get

$$(\Phi_d)^*\eta^- = \sigma_d \cdot \eta^-.$$

Then, also

$$(\Phi_d)^*\eta^+ = (\Phi_d)^*(\sigma_d \cdot \eta^-) = (\sigma_d \circ \Phi_d) \cdot (\Phi_d)^*\eta^- = (\sigma_d \circ \Phi_d) \cdot \eta^+.$$

As for the last equation, observing that the discrete Hamiltonian flow satisfies $\widetilde{\Phi_d} = \mathbb{F}^+ L_d \circ \Phi_d \circ (\mathbb{F}^+ L_d)^{-1}$ by definition, then

$$\begin{aligned}(\widetilde{\Phi_d})^*\eta =& ((\mathbb{F}^+ L_d)^{-1})^* \circ (\Phi_d)^*\eta^+ = ((\mathbb{F}^+ L_d)^{-1})^*((\sigma_d \circ \Phi_d) \cdot \eta^+) \\ =& (\sigma_d \circ \Phi_d \circ (\mathbb{F}^+ L_d)^{-1}) \cdot ((\mathbb{F}^+ L_d)^{-1})^*\eta^+,\end{aligned}$$

where the last equality comes from the properties of the pullback. Since we have that

$$\Phi_d \circ (\mathbb{F}^+ L_d)^{-1} = (\mathbb{F}^- L_d)^{-1} \quad \text{and} \quad ((\mathbb{F}^+ L_d)^{-1})^*\eta^+ = \eta,$$

the desired result follows.

Moreover, since the discrete Lagrangian function L_d is regular, the function σ_d does not vanish. Hence, the discrete flows Φ_d and $\widetilde{\Phi_d}$ are conformal contact. □

8.1.2 Discrete symmetries and dissipated quantities

Let G be a Lie group acting on Q through the map $\varphi : G \times Q \to Q$. We define the lifted action on $Q \times Q \times \mathbb{R}$ to be the diagonal action on $Q \times Q$ and the identity on $\mathbb{R}$, so that

$$\widetilde{\varphi} : G \times Q \times Q \times \mathbb{R} \to Q \times Q \times \mathbb{R}, \quad \widetilde{\varphi}_g(q_0, q_1, z_0) = (\varphi_g(q_0), \varphi_g(q_1), z_0).$$

Let us denote by $\xi_Q \in \mathfrak{X}(Q)$ the infinitesimal generator associated to a Lie algebra element $\xi \in \mathfrak{g}$ and by $\widetilde{\xi} \in \mathfrak{X}(Q \times Q \times \mathbb{R})$ the corresponding infinitesimal generator on $Q \times Q \times \mathbb{R}$.

Notice that, since $\mathrm{pr}_3(\varphi_g(q_0, q_1, z_0)) = z_0$ is constant for all $g \in G$, where $\mathrm{pr}_3 : Q \times Q \times \mathbb{R} \to \mathbb{R}$ is the projection onto the third factor, then we have that

$$T_{(q_0,q_1,z_0)}\mathrm{pr}_3(\widetilde{\xi}(q_0, q_1, z_0)) = 0.$$

In fact, the infinitesimal generator may be identified with

$$\widetilde{\xi}(q_0, q_1, z_0) = (\xi_Q(q_0), \xi_Q(q_1), 0_{z_0}) \in T_{q_0}Q \times T_{q_1}Q \times T_{z_0}\mathbb{R}, \tag{8.1.16}$$

where $0 : \mathbb{R} \to T\mathbb{R}$ is the zero section of $T\mathbb{R}$.

Lemma 8.1.10. *If $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$ is an invariant discrete Lagrangian function, i.e., $L_d \circ \widetilde{\varphi}_g = L_d$ for all $g \in G$, then it satisfies the equation*

$$D_1 L_d(q_0, q_1, z_0)\xi_Q(q_0) + D_2 L_d(q_0, q_1, z_0)\xi_Q(q_1) = 0. \tag{8.1.17}$$

Proof. Since the discrete Lagrangian function is invariant for the lifted action, it satisfies

$$\langle dL_d(q_0, q_1, z_0), \widetilde{\xi}(q_0, q_1, z_0)\rangle = 0, \ \forall (q_0, q_1, z_0) \in Q \times Q \times \mathbb{R}.$$

Then using equation (8.1.16), one immediately gets the desired expression. □

Now, consider the discrete momentum map J_d given by

$$\begin{aligned} &J_d : Q \times Q \times \mathbb{R} \to \mathfrak{g}^*, \\ &\langle J_d(q_0, q_1, z_0), \xi\rangle = \langle \eta^-(q_0, q_1, z_0), \widetilde{\xi}(q_0, q_1, z_0)\rangle. \end{aligned} \tag{8.1.18}$$

Theorem 8.1.11. *Let L_d be an invariant discrete Lagrangian function for the lifted action $\widetilde{\varphi}$. Then $\widetilde{\varphi}$ acts by contactomorphisms on $Q \times Q \times \mathbb{R}$ and the function $\hat{J}_d(\xi) : Q \times Q \times \mathbb{R} \to \mathbb{R}$ given by*

$$\hat{J}_d(\xi)(q_0, q_1, z_0) = \langle J_d(q_0, q_1, z_0), \xi \rangle$$

is dissipated along the discrete flow of Herglotz equations in the sense that

$$\hat{J}_d(\xi)(\varphi_d(q_0, q_1, z_0)) = \sigma_d(q_0, q_1, z_0) \hat{J}_d(\xi)(q_0, q_1, z_0),$$

where $\sigma_d(q_0, q_1, z_0) = 1 + D_z L_d(q_0, q_1, z_0)$.

Proof. The fact that $\widetilde{\varphi}$ acts by contactomorphisms is immediately checked by computing the pullback of either the 1-forms $\eta^{\pm}$:

$$(\widetilde{\varphi}_g)^* \eta^{\pm} = \eta^{\pm}.$$

Indeed, it is a direct consequence of the G-invariance of L_d. Following a similar proof as in Subsection 1.3.3 in [MW01] (where the authors show that, in the symplectic context, G-invariance implies that the action map preserves the discrete Lagrangian one-forms), we differentiate the equality $L_d \circ \widetilde{\varphi}_g = L_d$ with respect to z_0 and obtain

$$D_z L_d(\widetilde{\varphi}_g(q_0, q_1, z_0)) = D_z L_d(q_0, q_1, z_0),$$

while differentiation with respect to q_0 implies

$$(\widetilde{\varphi}_g)^*(D_1 L_d(q_0, q_1, z_0) dq_0) = D_1 L_d(q_0, q_1, z_0) dq_0.$$

Then, from the local expressions (8.1.12) and (8.1.13) and noting that $(\widetilde{\varphi}_g)^* dz_0 = dz_0$, the result follows.

In order to simplify the notation, let $P_0 = (q_0, q_1, z_0)$ and $P_1 = \varphi_d(q_0, q_1, z_0)$. By definition we have that

$$\hat{J}_d(\xi)(P_1) = \langle \eta^-(P_1), \widetilde{\xi}_Q(P_1) \rangle.$$

Now, applying the definition of η^- and equation (8.1.16) we get

$$\hat{J}_d(\xi)(P_1) = \frac{1}{\sigma_d(P_1)} \langle D_1 L_d(P_1), \xi_Q(q_1) \rangle.$$

Using the discrete Herglotz equations, the right-hand side reduces to

$$\hat{J}_d(\xi)(P_1) = -\langle D_2 L_d(P_0), \xi_Q(q_1)\rangle.$$

From the infinitesimal symmetry formula in equation (8.1.17), we deduce

$$\hat{J}_d(\xi)(P_1) = \langle D_1 L_d(P_0), \xi_Q(q_0)\rangle.$$

Now inserting $\sigma_d(P_0)$ so that

$$\hat{J}_d(\xi)(P_1) = \sigma_d(P_0)\langle \frac{D_1 L_d(P_0)}{\sigma_d(P_0)}, \xi_Q(q_0)\rangle,$$

we deduce

$$\hat{J}_d(\xi)(P_1) = \sigma_d(P_0)\langle \eta^-(P_0), \widetilde{\xi}(P_0)\rangle$$

and so we have proved that

$$\hat{J}_d(\xi)(P_1) = \sigma_d(P_0)\hat{J}_d(\xi)(P_0).$$

□

8.2 Exact discrete Lagrangian for contact systems

In this section, we will define the exact discrete Lagrangian function for contact systems (see Section 3.7.6 for the standard case) and prove that the associated discrete flow generated by discrete Herglotz equations is indeed the exact discrete flow.

In order to do that, we will need to define the contact exponential map which will relate the continuous and the discrete contact phase spaces.

8.2.1 The contact exponential map

Given a contact regular Lagrangian $L : TQ \times \mathbb{R} \to \mathbb{R}$, consider the corresponding Lagrangian vector field ξ_L and denote its flow by $\phi_t^{\xi_L}$ (see Section 3.5 to recall the definition).

Define the open subset U_h of $TQ \times \mathbb{R}$ given by

$$U_h = \{(q_0, \dot{q}_0, z_0) \in TQ \times \mathbb{R} \mid \phi_t^{\xi_L} \text{ is defined for } t \in [0, h]\}$$

and let the *contact exponential map* be defined by

$$\begin{aligned} \exp_h^{\xi_L} : U_h \subseteq TQ \times \mathbb{R} &\to Q \times Q \times \mathbb{R} \\ (q_0, \dot{q}_0, z_0) &\mapsto (q_0, q_1, z_0), \end{aligned} \tag{8.2.1}$$

where $q_1 = p_Q \circ \phi_h^{\xi_L}(q_0, \dot{q}_0, z_0)$ and $p_Q : TQ \times \mathbb{R} \to Q$ is the projection onto Q given by $p_Q(v_q, z) = q$ for $v_q \in T_qQ$.

We will prove that the contact exponential map is a local diffeomorphism, using the fact that the nonholonomic exponential map is a local embedding (see Theorem 4.3.2).

Indeed, to every regular contact system, one can associate a nonholonomic Lagrangian system on $T(Q\times\mathbb{R})$ with nonlinear constraints, using the singular Lagrangian function

$$\widetilde{L} : T(Q \times \mathbb{R}) \to \mathbb{R}, \quad \widetilde{L} = L \circ \pi, \tag{8.2.2}$$

where $\pi : T(Q \times \mathbb{R}) \to TQ \times \mathbb{R}$ is a projection onto $TQ \times \mathbb{R}$. Also, we take the nonlinear constraints

$$M_L = \{(q, z, \dot{q}, \dot{z}) \in T(Q \times \mathbb{R}) \mid \dot{z} = L(q, \dot{q}, z)\}. \tag{8.2.3}$$

Observe that M_L is the zero level set of the real-valued function $\Phi : T(Q \times \mathbb{R}) \to \mathbb{R}$ given by $\Phi(q, z, \dot{q}, \dot{z}) = \dot{z} - L(q, \dot{q}, z)$.

The pair $(\widetilde{L}, M_L)$ forms a Lagrangian nonholonomic system with nonlinear constraints determined by the submanifold M_L and dynamics given by Chetaev's principle (see [Blo15; LD96] and references therein). According to this principle the equations of motion are

$$\begin{aligned} &\frac{d}{dt}\left(\frac{\partial \widetilde{L}}{\partial \dot{q}^i}\right) - \frac{\partial \widetilde{L}}{\partial q^i} = \lambda \frac{\partial \Phi}{\partial \dot{q}^i} \\ &\frac{d}{dt}\left(\frac{\partial \widetilde{L}}{\partial \dot{z}}\right) - \frac{\partial \widetilde{L}}{\partial z} = \lambda \frac{\partial \Phi}{\partial \dot{z}} \\ &\Phi(q^i, z, \dot{q}^i, \dot{z}) = 0, \end{aligned} \tag{8.2.4}$$

with Lagrange multiplier λ. As $\widetilde{L}$ does not depend on $\dot{z}$ it is straightforward to check that the Lagrange multiplier is just

$$\lambda = -\frac{\partial L}{\partial z}$$

and that equations (8.2.4) are equivalent to the Herglotz equations for L.

Moreover, since L is regular, we may check that the nonholonomic system formed by $(\widetilde{L}, M_L)$ is regular and so we can define a SODE vector field $\Gamma_{(\widetilde{L},M_L)} \in \mathfrak{X}(M_L)$ as the unique vector field on M_L whose integral curves satisfy equations (8.2.4). Hence, we deduce

$$T\pi(\Gamma_{(\widetilde{L},M_L)}) = \xi_L \circ \pi. \tag{8.2.5}$$

Let us denote the flow of the vector field $\Gamma_{(\widetilde{L},M_L)}$ by $\phi_t^{\Gamma_{(\widetilde{L},M_L)}} : M_L \to M_L$. Consider now the submanifold of M_L given by

$$M_{L,h} = \{(q_0, \dot{q}_0, z_0, \dot{z}_0) \in M_L \mid \phi_t^{\Gamma_{(\widetilde{L},M_L)}} \text{ is defined for } t \in [0, h]\}.$$

Recall that the nonholonomic exponential map (see equation (4.3.1)) was defined as

$$\begin{aligned} \exp_h^{\Gamma_{(\widetilde{L},M_L)}} :\ M_{L,h} \subseteq\ M_L &\longrightarrow (Q \times \mathbb{R}) \times (Q \times \mathbb{R}) \\ (q_0, z_0, \dot{q}_0, \dot{z}_0) &\mapsto (q_0, z_0, q_1, z_1), \end{aligned} \tag{8.2.6}$$

where $(q_1, z_1) = \tau_{Q\times\mathbb{R}} \circ \phi_h^{\Gamma_{(\widetilde{L},M_L)}}(q_0, z_0, \dot{q}_0, \dot{z}_0)$, with $\tau_{Q\times\mathbb{R}} : T(Q \times \mathbb{R}) \to Q \times \mathbb{R}$ the tangent bundle projection.

Now according to Theorem 4.3.2, there is an open subset $N_h \subseteq M_{L,h}$ such that the nonholonomic exponential map $\exp_h^{\Gamma_{(\widetilde{L},M_L)}}|_{N_h}$ is a smooth embedding and, hence, a diffeomorphism into its image, which we will denote by $\mathcal{M}_h^{e,nh}$.

Remark 8.2.1. Notice that we are applying Theorem 4.3.2 to a nonholonomic system which has a nonlinear constraint, i.e., it is not determined by a distribution. However, it is not difficult to check that Theorem 4.3.2 still holds when the nonholonomic constraint is a nonlinear submanifold of TQ. Indeed, to prove the theorem in the general case, the only non-trivial step is to extend the vector field $\Gamma_{(\widetilde{L},M_L)}$ to a SODE on $T(Q \times \mathbb{R})$. But, at least at a local level, when we restrict to a coordinate chart this is always possible. (see also [MDM21] for a discussion about exponential maps of SODE vector fields on Lie algebroids).

Theorem 8.2.2. *There exists a sufficiently small $h > 0$ and an open set $V_h \subseteq U_h$ such that the contact exponential map $exp_h^{\xi_L}|_{V_h}$ is a diffeomorphism.*

Proof. Let us consider the non-holonomic system $(\widetilde{L}, M_L)$ defined previously.

According to equation (8.2.5), the vector fields ξ_L and $\Gamma_{(\widetilde{L},M_L)}$ are π-related therefore, its flows satisfy

$$\pi \circ \phi_t^{\Gamma_{(\widetilde{L},M_L)}} = \phi_t^{\xi_L} \circ \pi|_{M_L}.$$

We remark that $\pi|_{M_L}$ is a diffeomorphism, from the definition of M_L. As such, we can also write

$$\phi_t^{\Gamma_{(\widetilde{L},M_L)}} = (\pi|_{M_L})^{-1} \circ \phi_t^{\xi_L} \circ \pi|_{M_L}.$$

Thus, we can write the non-holonomic exponential map in terms of the contact dynamics in the following way

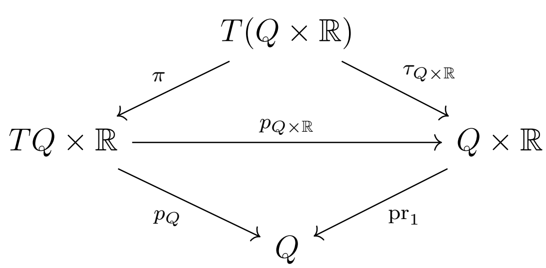
$$\exp_h^{\Gamma_{(\widetilde{L},M_L)}}(q_0, z_0, \dot{q}_0, \dot{z}_0) = (q_0, z_0, q_1, z_1),$$

with $(q_1, z_1) = \tau_{Q\times\mathbb{R}} \circ (\pi|_{M_L})^{-1} \circ \phi_h^{\xi_L} \circ \pi|_{M_L}(q_0, z_0, \dot{q}_0, \dot{z}_0)$ where $\dot{z}_0 = L(q_0, \dot{q}_0, z_0)$.

Also note that $\tau_{Q\times\mathbb{R}} \circ (\pi|_{M_L})^{-1} = p_{Q\times\mathbb{R}}$, where

$$p_{Q\times\mathbb{R}} : TQ \times \mathbb{R} \to Q \times \mathbb{R}, \quad p_{Q\times\mathbb{R}}(v_q, z) = (q, z).$$

In Diagram (8.2.7) we show the different projections we can define on the manifolds involved in this section.



(8.2.7)

With these projections we can also write the contact exponential map as

$$\exp_h^{\xi_L}(q_0, \dot{q}_0, z_0) = (q_0, q_1, z_0),$$

with $q_1 = \mathrm{pr}_1 \circ p_{Q\times\mathbb{R}} \circ \phi_h^{\xi_L}(q_0, \dot{q}_0, z_0)$. Hence, we can write it as

$$\exp_h^{\xi_L} = \widetilde{\mathrm{pr}}_1 \circ \exp_h^{\Gamma_{(\widetilde{L},M_L)}} \circ (\pi|_{M_L})^{-1}, \tag{8.2.8}$$

with

$$\widetilde{\mathrm{pr}}_1 : (Q \times \mathbb{R}) \times (Q \times \mathbb{R}) \longrightarrow Q \times Q \times \mathbb{R}$$
$$(q_0, z_0, q_1, z_1) \mapsto (q_0, \mathrm{pr}_1(q_1, z_1), z_0) = (q_0, q_1, z_0).$$

Therefore, if $\widetilde{\mathrm{pr}}_1|_{\mathcal{M}_h^{e,nh}}$ is a local diffeomorphism then, by equation (8.2.8), the contact exponential map $\exp_h^{\xi_L}|_{V_h}$ is a diffeomorphism if we choose

$$V_h = \pi|_{M_L}(N_h),$$

where N_h is the open subset where $\exp_h^{\Gamma_{(\tilde{L},M_L)}}|_{N_h}$ is an embedding.

We are going to prove in the next Lemma that $\widetilde{\mathrm{pr}}_1|_{\mathcal{M}_h^{e,nh}}$ is a local diffeomorphism.

□

Lemma 8.2.3. *Using the same notation as in the previous theorem, $\widetilde{pr}_1|_{\mathcal{M}_h^{e,nh}}$ is a local diffeomorphism.*

Proof. All we must prove is that $\widetilde{\mathrm{pr}}_1|_{\mathcal{M}_h^{e,nh}}$ is a local submersion (immersion) since, by dimensional reasons, this forces $\widetilde{\mathrm{pr}}_1|_{\mathcal{M}_h^{e,nh}}$ to be also a local immersion (submersion).

Let $x \in \mathcal{M}_h^{e,nh}$. The kernel of $T_x\widetilde{\mathrm{pr}}_1|_{\mathcal{M}_h^{e,nh}}$ is spanned by the velocity vector of curves of the form

$$Z(s) = (q_0, z_0, q_1, w \cdot s) \in \mathcal{M}_h^{e,nh}, \quad w \in \mathbb{R}.$$

Let $\gamma_s(t) = \phi_t^{\Gamma_{(\tilde{L},M_L)}} \circ (\exp_h^{\Gamma_{(\tilde{L},M_L)}})^{-1}(Z(s))$. For each fixed value of s, this is an integral curve of $\Gamma_{(\widetilde{L},M_L)}$ satisfying

$$\tau_{Q\times\mathbb{R}} \circ \gamma_s(0) = (q_0, z_0), \quad \tau_{Q\times\mathbb{R}} \circ \gamma_s(h) = (q_1, w \cdot s).$$

Moreover, note that the projection of $\gamma_s(t)$ to $TQ \times \mathbb{R}$, i.e., the curve $\pi \circ \gamma_s(t)$ is an integral curve of ξ_L with endpoints q_0 and q_1 for each fixed value of s and so $\pi \circ \gamma_0(t)$ must satisfy Herglotz' principle. Note that the action over the curves $\pi \circ \gamma_s(t)$ is given by

$$\mathcal{A}(p_Q \circ \pi \circ \gamma_s(t)) = p_{\mathbb{R}} \circ \pi \circ \gamma_s(h) = w \cdot s,$$

where $p_{\mathbb{R}} : TQ \times \mathbb{R} \to \mathbb{R}$ is the projection onto the second factor.

Therefore, $p_Q \circ \pi \circ \gamma_0(t)$ is a critical value of the action if and only if $w = 0$. Hence, $T_x\widetilde{\mathrm{pr}}_1|_{\mathcal{M}_h^{e,nh}}$ has trivial kernel and $\widetilde{\mathrm{pr}}_1|_{\mathcal{M}_h^{e,nh}}$ must be a local diffeomorphism in a neighbourhood of each point. □

Since the contact exponential map is a local diffeomorphism we can define a local inverse called the *exact retraction* and denote it by $R_h^{e-} : Q\times Q\times\mathbb{R} \to TQ\times\mathbb{R}$. We will also use its translation by the flow

$$R_h^{e+} : Q\times Q\times\mathbb{R} \to TQ\times\mathbb{R}, \quad R_h^{e+} := \phi_h^{\xi_L} \circ R_h^{e-}.$$

8.2.2 The exact discrete Lagrangian function

Let $L_h^e : Q\times Q\times\mathbb{R} \to \mathbb{R}$ and defined by

$$L_h^e(q_0,q_1,z_0) = \int_0^h L\circ\phi_t^{\xi_L}\circ R_h^{e-}(q_0,q_1,z_0)dt \tag{8.2.9}$$

is called the *exact discrete Lagrangian* function.

We will need the following classical result in the proof of the next theorem: the solution of the first order linear equation $\dot{y} = a(t) + \frac{db}{dt}(t)y$ with $b(0)=0$ is

$$y(t) = e^{b(t)}\left(\int_0^t a(s)e^{-b(s)}\ ds + y(0)\right). \tag{8.2.10}$$

Theorem 8.2.4. *The Legendre transformations of a regular Lagrangian $L : TQ\times\mathbb{R}\to\mathbb{R}$ are related to the discrete Legendre transformations of the corresponding exact discrete Lagrangian $L_h^e : Q\times Q\times\mathbb{R}\to\mathbb{R}$ in the following way*

$$\mathbb{F}^+L_h^e = \mathbb{F}L\circ R_h^{e+}, \quad \mathbb{F}^-L_h^e = \mathbb{F}L\circ R_h^{e-}. \tag{8.2.11}$$

Proof. We will prove in local computations that the derivatives of the exact discrete Lagrangian function satisfy

$$\begin{aligned}
D_1L_h^e(q_0,q_1,z_0) &= -\frac{\partial L}{\partial\dot{q}^i}(q_0,\dot{q}_0,z_0)e^{b(h)}dq_0^i;\\
D_2L_h^e(q_0,q_1,z_0) &= \frac{\partial L}{\partial\dot{q}^i}(q_1,\dot{q}_1,z_1)dq_1^i;\\
D_zL_h^e(q_0,q_1,z_0) &= (e^{b(h)}-1)dz_0.
\end{aligned} \tag{8.2.12}$$

where

$$\begin{aligned}
&(q_0,\dot{q}_0,z_0) = R_h^{e-}(q_0,q_1,z_0), \quad (q_1,\dot{q}_1,z_1) = \phi_h^{\xi_L}\circ R_h^{e-}(q_0,q_1,z_0),\\
&\text{and}\quad b(t) = \int_0^t \frac{\partial L}{\partial z}(\phi_s^{\xi_L}\circ R_h^{e-}(q_0,q_1,z_0))\ ds.
\end{aligned} \tag{8.2.13}$$

Then, from the definition of Legendre transform in (3.5.9) and discrete Legendre transformations in (8.1.6), the result follows immediately.

To simplify the notation in the proof we will use the notation $\gamma_{0,1}(t) = (q_{0,1}(t), \dot{q}_{0,1}(t), z_{0,1}(t)) := \phi_t^{\xi_L} \circ R_h^{e-}(q_0, q_1, z_0)$. Under this convention we will have

$$L_h^e(q_0, q_1, z_0) = \int_0^h L(\gamma_{0,1}(t))dt.$$

Note first that any variation of the exact discrete Lagrangian will take the form

$$\begin{aligned} \delta L_h^e(q_0, q_1, z_0) &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} L_h^e(\tilde{q}_0(\varepsilon), \tilde{q}_1(\varepsilon), \tilde{z}_0(\varepsilon)) \\ &= \int_0^h \left[\frac{\partial L}{\partial q}(\gamma_{0,1}(t))\delta q_{0,1} + \frac{\partial L}{\partial \dot{q}}(\gamma_{0,1}(t))\delta \dot{q}_{0,1} + \frac{\partial L}{\partial z}(\gamma_{0,1}(t))\delta z_{0,1}\right] dt. \end{aligned} \tag{8.2.14}$$

Since $\gamma_{0,1}(t)$ is a solution of Herglotz equations, it satisfies

$$\dot{z}_{0,1}(t) = L(q_{0,1}(t), \dot{q}_{0,1}(t), z_{0,1}(t)).$$

Therefore, any variation of $z_{0,1}$ satisfies the variational equation

$$\delta \dot{z}_{0,1} = \frac{\partial L}{\partial q}(\gamma_{0,1}(t))\delta q_{0,1} + \frac{\partial L}{\partial \dot{q}}(\gamma_{0,1}(t))\delta \dot{q}_{0,1} + \frac{\partial L}{\partial z}(\gamma_{0,1}(t))\delta z_{0,1}. \tag{8.2.15}$$

Hence, any variation of the exact discrete Lagrangian reduces to

$$\delta L_h^e(q_0, q_1, z_0) = \delta z_{0,1}(h) - \delta z_{0,1}(0) = \delta z_{0,1}(h). \tag{8.2.16}$$

Moreover, we can solve the function $\delta z_{0,1}$ explicitly, by solving the differential equation (8.2.15) and using (8.2.10). So, we obtain

$$\delta z_{0,1}(h) = e^{b(h)} \left(\int_0^h a(s)e^{-b(s)}\ ds + \delta z_{0,1}(0)\right), \tag{8.2.17}$$

with

$$\begin{aligned} b(t) &= \int_0^t \frac{\partial L}{\partial z}(\gamma_{0,1}(s))\ ds, \\ a(t) &= \frac{\partial L}{\partial q}(\gamma_{0,1}(t))\delta q_{0,1} + \frac{\partial L}{\partial \dot{q}}(\gamma_{0,1}(t))\delta \dot{q}_{0,1}. \end{aligned}$$

Let us compute the integration in the expression of $\delta z_{0,1}$:

$$\begin{aligned}\int_0^h a(s)e^{-b(s)}\,ds &= \int_0^h \left(\frac{\partial L}{\partial q}\delta q_{0,1} + \frac{\partial L}{\partial \dot q}\delta \dot q_{0,1}\right) e^{-b(t)}\,dt \\ &= \int_0^h \left[\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot q} + \frac{\partial L}{\partial \dot q}\frac{\partial L}{\partial z}\right]\delta q_{0,1} e^{-b(t)}\,dt \\ &\quad + \frac{\partial L}{\partial \dot q}(\gamma_{0,1}(h))e^{-b(h)}\delta q_{0,1}(h) - \frac{\partial L}{\partial \dot q}(\gamma_{0,1}(0))\delta q_{0,1}(0),\end{aligned}$$

where we are using integration by parts. Note that the term inside the brackets vanishes, since the exact discrete contact Lagrangian is being evaluated over solutions of Herglotz equations. Therefore,

$$\delta z_{0,1}(h) = \frac{\partial L}{\partial \dot q}(\gamma_{0,1}(h))\delta q_{0,1}(h) - \frac{\partial L}{\partial \dot q}(\gamma_{0,1}(0))e^{b(h)}\delta q_{0,1}(0) + e^{b(h)}\delta z_{0,1}(0). \tag{8.2.18}$$

Note that the differentials of the discrete Lagrangian $D_1 L_h^e$, $D_2 L_h^e$ and $D_z L_h^e$ are instances of particular variations. Therefore, we have that

$$\begin{aligned}D_1 L_h^e(q_0, q_1, z_0) &= \left(\frac{\partial L}{\partial \dot q}(\gamma_{0,1}(h))\frac{\partial q_{0,1}(h)}{\partial q_0^i} - \frac{\partial L}{\partial \dot q}(\gamma_{0,1}(0))e^{b(h)}\frac{\partial q_{0,1}(0)}{\partial q_0^i}\right. \\ &\quad \left. +(e^{b(h)}-1)\frac{\partial z_{0,1}(0)}{\partial q_0^i}\right) dq_0^i \\ &= -\frac{\partial L}{\partial \dot q^i}(\gamma_{0,1}(0))e^{b(h)}dq_0^i,\end{aligned} \tag{8.2.19}$$

since $q_{0,1}(h) \equiv q_1$ and so its derivative with respect to q_0 vanishes, $q_{0,1}(0) \equiv q_0$ and so its derivative with respect to q_0 is the identity and, finally, $z_{0,1}(0) \equiv z_0$ does not depend upon q_0. Likewise, the next derivative follows from applying similar arguments. Indeed, we have that

$$\begin{aligned}D_2 L_h^e(q_0, q_1, z_0) &= \left(\frac{\partial L}{\partial \dot q}(\gamma_{0,1}(h))\frac{\partial q_{0,1}(h)}{\partial q_1^i} - \frac{\partial L}{\partial \dot q}(\gamma_{0,1}(0))e^{b(h)}\frac{\partial q_{0,1}(0)}{\partial q_1^i}\right. \\ &\quad \left. +(e^{b(h)}-1)\frac{\partial z_{0,1}(0)}{\partial q_1^i}\right) dq_1^i \\ &= \frac{\partial L}{\partial \dot q^i}(\gamma_{0,1}(h))dq_1^i.\end{aligned} \tag{8.2.20}$$

Analogously, we also deduce

$$
\begin{aligned}
D_z L_h^e(q_0, q_1, z_0) = &\left(\frac{\partial L}{\partial \dot{q}}(\gamma_{0,1}(h)) \frac{\partial q_{0,1}(h)}{\partial z_0} - \frac{\partial L}{\partial \dot{q}}(\gamma_{0,1}(0)) e^{b(h)} \frac{\partial q_{0,1}(0)}{\partial z_0} \right. \\
&\left. + (e^{b(h)} - 1) \frac{\partial z_{0,1}(0)}{\partial z_0} \right) dz_0 \\
&= (e^{b(h)} - 1) dz_0.
\end{aligned}
\tag{8.2.21}
$$

Now, the result follows by the definition of the discrete Legendre transformations in (8.1.6). □

The commutativity of the following diagram summarizes the statement of the previous theorem

$$
\begin{array}{ccc}
Q \times Q \times \mathbb{R} & \xrightarrow{R_h^{e\pm}} & TQ \times \mathbb{R} \\
& \searrow^{\mathbb{F}^{\pm} L_h^e} & \downarrow \mathbb{F}L \\
& & T^*Q \times \mathbb{R}
\end{array}
\tag{8.2.22}
$$

Now, we are going to relate the continuous contact Lagrangian flow with its discrete counterpart, when we take as discrete Lagrangian the corresponding exact discrete Lagrangian.

Theorem 8.2.5. *Take a regular Lagrangian $L : TQ \to \mathbb{R}$ and fix a time step $h > 0$. Then we have that:*

1. *L_h^e is a regular discrete Lagrangian function;*
2. *If H is the Hamiltonian function corresponding to L introduced at the end of Section 3.5 and $\phi_t^{X_H}$ is its contact Hamiltonian flow, we have that*
$$\mathbb{F}^+ L_h^e = \phi_h^{X_H} \circ \mathbb{F}^- L_h^e. \tag{8.2.23}$$
3. *Let $(q, z) : [0, Nh] \to Q \times \mathbb{R}$ be a solution of the Herglotz equations and let $\{(q_0, z_0), (q_1, z_1), ..., (q_N, z_N)\}$ be a solution of the discrete Herglotz equations for the corresponding exact discrete contact Lagrangian with $(q(0), q(h), z(0))$ as initial conditions. Then they are related in the following way:*
$$q_k = q(kh), \quad z_k = z(kh) \quad \text{for } k = 0, ..., N. \tag{8.2.24}$$

Proof. Item 1. is a consequence of the previous theorem, since $\mathbb{F}^- L_h^e$ is a composition of two local diffeomorphisms it is itself a local diffeomorphism. Item 2. comes from unwinding the definitions:

$$\mathbb{F}^+ L_h^e = \mathbb{F}L \circ R_h^{e+} = \mathbb{F}L \circ \phi_h^{\Gamma_L} \circ R_h^{e-} = \phi_h^{X_H} \circ \mathbb{F}L \circ R_h^{e-} = \phi_h^{X_H} \circ \mathbb{F}^- L_h^e.$$

For item 3., it is not hard to show that

$$\mathbb{F}^+ L_h^e = \mathbb{F}^- L_h^e \circ (\exp_h^{\xi_L} \circ \phi_h^{\xi_L} \circ R_h^{e-}).$$

Moreover, for every $k = 1, ..., N-1$, since the curves q and z are solution of the Herglotz equations, we have that

$$\exp_h^{\xi_L} \circ \phi_h^{\xi_L} \circ R_h^{e-}(q(k-1), q(k), z(k-1)) = (q(k), q(k+1), z(k)).$$

Hence,

$$\mathbb{F}^+ L_h^e(q(k-1), q(k), z(k-1)) = \mathbb{F}^- L_h^e(q(k), q(k+1), z(k))$$

so that $\{(q_0, z_0), (q_1, z_1), ..., (q_N, z_N)\}$ given by (8.2.24) satisfy the discrete Herglotz equations. □

8.3 Numerical examples

Given a mechanical contact Lagrangian with a euclidean metric and a potential function $V : Q \to \mathbb{R}$ of the type

$$L(q, \dot{q}, z) = \frac{1}{2}\dot{q}^2 - V(q) + \gamma z, \quad (q, \dot{q}, z) \in TQ \times \mathbb{R}, \quad \gamma < 0.$$

one usually approximates the exact discrete Lagrangian associated to L by means of a quadrature rule. Note that the restriction of γ to negative values is necessary to model dissipative dynamics, though we could define the integrator for any value of $\gamma \in \mathbb{R}$. If we use the middle point rule to approximate the positions, i.e., $q \approx \frac{q_1+q_0}{2}$, one may define the discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$ in the following way

$$L_d(q_0, q_1, z_0) = \frac{1}{2h}(q_1 - q_0)^2 - hV\left(\frac{q_1 + q_0}{2}\right) + h\gamma z_0.$$

We remark that the value of h should be chosen small enough so that the function σ_d does not vanish anywhere. In this case, the discrete Herglotz equations are of the type

$$\begin{aligned}
&\frac{q_1 - q_0}{h} - \frac{h}{2}\frac{\partial V}{\partial q}\left(\frac{q_1+q_0}{2}\right) = \frac{1}{(1+h\gamma)}\left(\frac{q_2-q_1}{h} + \frac{h}{2}\frac{\partial V}{\partial q}\left(\frac{q_2+q_1}{2}\right)\right)\\
&z_1 = L_d(q_0,q_1,z_0) = \frac{1}{2h}(q_1-q_0)^2 - hV\left(\frac{q_1+q_0}{2}\right) + (h\gamma+1)z_0
\end{aligned}$$

Example 8.3.1. The free single particle contact Lagrangian is

$$L(q,\dot{q},z) = \frac{1}{2}\dot{q}^2 + \gamma z, \quad (q,\dot{q},z) \in TQ \times \mathbb{R}.$$

A simple discretization of this Lagrangian would be

$$L_d(q_0,q_1,z_0) = \frac{1}{2h}(q_1-q_0)^2 + h\gamma z_0. \tag{8.3.1}$$

Then, choosing h small enough so that the function σ_d is non-vanishing, the discrete Herglotz equations for L_d are locally given by

$$\begin{aligned}
&\frac{q_1-q_0}{h} = \frac{q_2-q_1}{h(1+h\gamma)} \quad \Rightarrow \quad q_2 = (h\gamma+2)q_1 - (h\gamma+1)q_0\\
&z_1 = \frac{1}{2h}(q_1-q_0)^2 + (h\gamma+1)z_0
\end{aligned}$$

The discrete flow obtained by solving these equations is plotted in Fig. 8.1.

In this case, one can also compute the exact discrete Lagrangian and solve the exact dynamics.

$$L_h^e(q_0,q_1,z_0) = \frac{\gamma\,(q_1-q_0)^2\, e^{\gamma h}}{2e^{\gamma h}-2} - z_0\left(e^{\gamma h}-1\right). \tag{8.3.2}$$

△

Example 8.3.2. The damped harmonic oscillator is described by the Lagrangian

$$L(q,\dot{q},z) = \frac{1}{2}\dot{q}^2 - \frac{1}{2}q^2 + \gamma z, \quad (q,\dot{q},z) \in TQ \times \mathbb{R}.$$

Using a middle point discretization, i.e., $q \approx \frac{q_1+q_0}{2}$, one may define the discrete Lagrangian

$$L_d(q_0,q_1,z_0) = \frac{1}{2h}(q_1-q_0)^2 - \frac{h}{8}(q_1+q_0)^2 + h\gamma z_0. \tag{8.3.3}$$

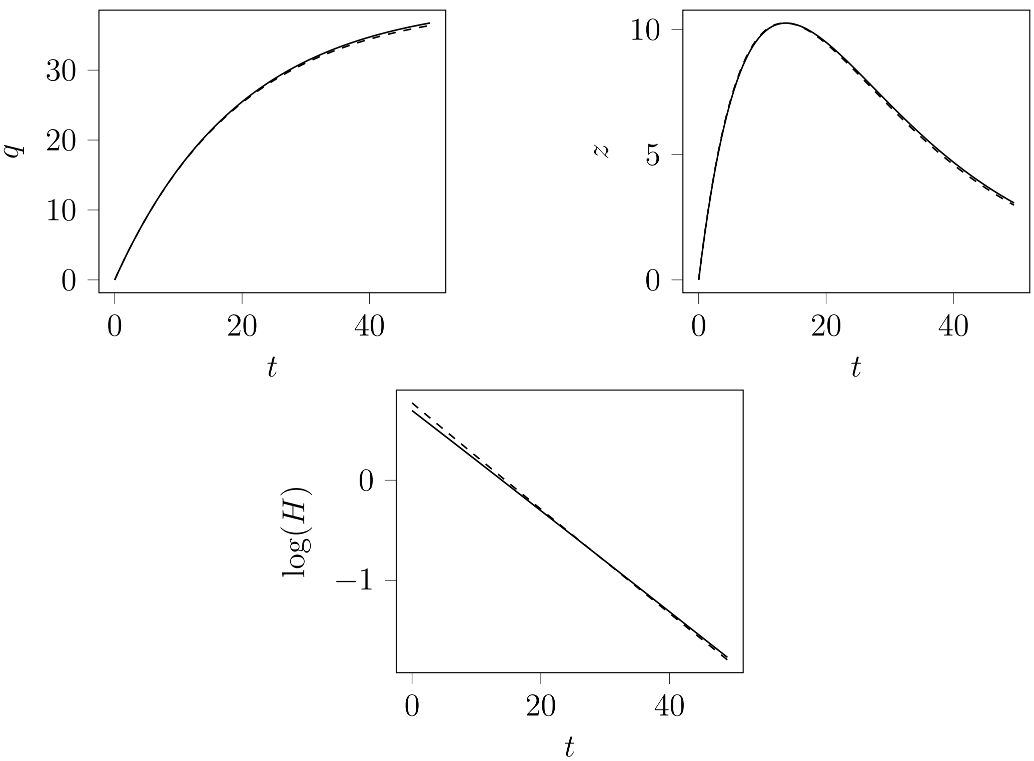


Figure 8.1: Position q and z and logarithm of the discrete Hamiltonian $H \circ \mathbb{F}^{-}L_d$ for a free particle, computed by solving the discrete Herglotz equations for the discrete Lagrangian (8.3.1) (continuous line) and the exact dynamics (dashed line), for $\gamma = -0.05$ and the time-step $h = 0.5$. The initial conditions are $q_0 = 1$, $q_1 = 2$ and $z_0 = 0$.

In this case, after choosing h small enough, the discrete Herglotz equations hold

$$\frac{q_1 - q_0}{h} - \frac{h}{4}(q_1 + q_0) = \frac{1}{(1 + h\gamma)} \left(\frac{q_2 - q_1}{h} + \frac{h}{4}(q_2 + q_1) \right)$$
$$z_1 = \frac{1}{2h}(q_1 - q_0)^2 - \frac{h}{8}(q_1 + q_0)^2 + (h\gamma + 1)z_0,$$

which can be solved explicitly for q_2

$$q_2 = -\frac{(h^3\gamma + 4h\gamma + h^2 + 4)q_0 + (h^3\gamma - 4h\gamma + 2h^2 - 8)q_1}{h^2 + 4}.$$

The discrete flow obtained by solving these equations is plotted in Fig. 8.2.

In this case, the exact discrete Lagrangian and the exact discrete dynamics can be computed with the aid of a Computer Algebra system, but the analytic expressions are complicated, so we only include their graph in Fig. 8.2. △

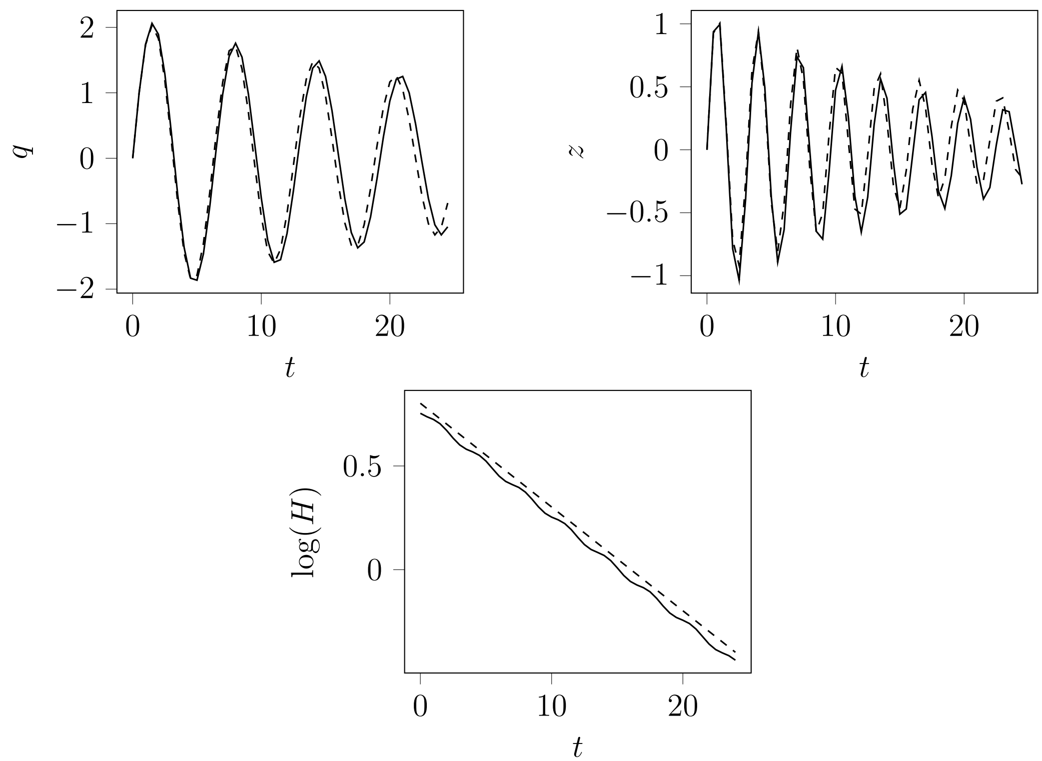


Figure 8.2: Position q and z and logarithm of the discrete Hamiltonian $H \circ \mathbb{F}^{-}L_d$ for a harmonic oscillator, computed by solving the discrete Herglotz equations on the discrete Lagrangian (8.3.3) (continuous line) and the exact dynamics (dashed line), for $\gamma = -0.05$ and the time-step $h = 0.5$. The initial conditions are $q_0 = 1$, $q_1 = 2$ and $z_0 = 0$.

Chapter 9

Conclusions and future research

In this thesis, we have developed four main lines of investigation: radial trajectories of nonholonomic mechanical systems (Chapter 5), nonholonomic Jacobi fields (Chapter 6), discrete nonholonomic mechanics (Chapter 7) and discrete contact mechanics (Chapter 8). Let us summarize our findings and open the door to new results along these lines:

Radial trajectories of nonholonomic mechanical systems

Given a kinetic nonholonomic system, with configuration space Q, we have identified and characterized the family of Riemannian metrics on the image $\mathcal{M}_q^{nh}$ of the nonholonomic exponential map at a fixed point $q \in Q$, satisfying the relevant property that the minimizing geodesics with starting point q and with respect to these metrics are, for sufficiently small times, just the nonholonomic trajectories starting at the same point q.

We have also proved that such metrics on $\mathcal{M}_q^{nh}$ always exist and we illustrate these facts with several examples. We remark that these findings are surprising and unexpected, since until now nonholonomic dynamics was seen to be un-variational and we are now providing a new perspective under which it becomes variational, in the sense that, at least radial trajectories are geodesics with respect to a Riemannian metric.

After these results, a lot of work remains to be done. In fact, our idea is to develop a research program in order to discuss the geometric properties of the nonholonomic trajectories for a kinetic nonholonomic system. Some problems which will be covered in this program are the following ones:

- **Geodesic flows and the kinetic nonholonomic flow.** Let $(g, \mathcal{D})$ be a kinetic nonholonomic system with configuration space Q, $\Gamma_{(g,\mathcal{D})} \in \mathfrak{X}(\mathcal{D})$ the kinetic nonholonomic flow and $\exp^{\Gamma_{(g,\mathcal{D})}} : M^{\Gamma_{(g,\mathcal{D})}} \subseteq \mathcal{D} \to Q \times Q$ the nonholonomic exponential map introduced in Chapter 4 (see Section 4.3). Then, following theorem 4.3.2, one deduces that there exists an open neighbourhood $\mathcal{U}_{0(Q)}$ of the zero section $0(Q)$ in $\mathcal{D}$ such that
$$(\exp^{\Gamma_{(g,\mathcal{D})}})_{|\mathcal{U}_{0(Q)}} : \mathcal{U}_{0(Q)} \subseteq \mathcal{D} \to Q \times Q$$
is an embedding. Denote by $\mathcal{M}^{nh} = \exp^{\Gamma_{(g,\mathcal{D})}}(\mathcal{U}_{0(Q)})$. It is clear that $\mathcal{M}^{nh}$ is an embedded submanifold of $Q \times Q$,
$$\exp^{nh} := (\exp^{\Gamma_{(g,\mathcal{D})}})_{|\mathcal{U}_{0(Q)}} : \mathcal{U}_{0(Q)} \subseteq \mathcal{D} \to \mathcal{M}^{nh} \subseteq Q \times Q$$
is a diffeomorphism and the following diagram
$$\begin{array}{ccc} \mathcal{U}_{0(Q)} \subseteq \mathcal{D} & \xrightarrow{\exp^{nh}} & \mathcal{M}^{nh} \subseteq Q \times Q \\ & {\scriptstyle (\tau_D)_{|\mathcal{U}_{0(Q)}}} \searrow \quad \swarrow {\scriptstyle (\mathrm{pr}_1)_{|\mathcal{M}^{nh}}} & \\ & Q & \end{array}$$
is commutative, where $\tau_{\mathcal{D}} : \mathcal{D} \to Q$ is the vector bundle projection. Note that
$$(\mathrm{pr}_1)^{-1}_{|\mathcal{M}^{nh}}(q) = \exp^{nh}_q(\mathcal{U}_{0(Q)} \cap \mathcal{D}_q) \simeq \mathcal{M}^{nh}_q, \quad \text{for } q \in Q.$$
Now, denote by $\Gamma^{nh}_{(g,\mathcal{D})}$ the nonholonomic flow considered as a vector field on $\mathcal{M}^{nh}$, via the diffeomorphism $\exp^{nh}$. Then, proceeding as in Chapter 5, one may find a family of bundle metrics
$$g^{nh} : V(\mathrm{pr}_1)_{|\mathcal{M}^{nh}} \times_{\mathcal{M}^{nh}} V(\mathrm{pr}_1)_{|\mathcal{M}^{nh}} \to \mathbb{R}$$
on the vertical bundle of the fiber bundle with projection $(\mathrm{pr}_1)_{|\mathcal{M}^{nh}} : \mathcal{M}^{nh} \to Q$ such that if $q \in Q$, the minimizing geodesics on $\mathcal{M}^{nh}_q \subseteq \mathcal{M}^{nh}$ with starting point q are, for sufficiently small times, the nonholonomic trajectories with the same starting point q.

  Moreover, one may consider the geodesic flow $\Gamma^{g^{nh}}$ associated with one of such metrics g^{nh} as a vector field on $V(\mathrm{pr}_1)_{|\mathcal{M}^{nh}}$. Then, it would be interesting to discuss the relation between the vector field $\Gamma^{g^{nh}}$ on $V(\mathrm{pr}_1)_{|\mathcal{M}^{nh}}$ and the nonholonomic flow $\Gamma^{nh}_{(g,\mathcal{D})}$ on $\mathcal{M}^{nh}$.

- **Kinetic Lagrangianization of kinetic nonholonomic systems.** Let $(g, \mathcal{D})$ be a kinetic nonholonomic system with configuration space Q. After the results and examples in Chapter 5, another natural question arises: under what conditions can one get a kinetic Lagrangianization of the system $(g, \mathcal{D})$? In other words, under what conditions does there exist a Riemannian metric g^{nh} on the whole ambient manifold Q, rather than just on $\mathcal{M}_q^{nh}$, such that the kinetic nonholonomic trajectories for the system $(g, \mathcal{D})$ are just the geodesics of the metric g^{nh} with initial velocity in $\mathcal{D}$?

  Note that there are examples of kinetic nonholonomic systems admitting such metrics on the whole configuration space and, despite that, the constraint distribution is still not integrable (see Example 5.2.14 in Chapter 5). On the other hand, Theorem 5.2.4 may be considered as the first step in order to give an answer to the previous hard question. We also remark that if the system admits a kinetic Lagrangianization then, using the Legendre transformation associated with the kinetic Lagrangian system induced by the Riemannian metric g^{nh}, one may produce a Hamiltonian formulation of the original nonholonomic system. So, our question is related with a classical problem in nonholonomic mechanics: the so-called Hamiltonization problem. This problem discusses whether a nonholonomic system admits a Hamiltonian formulation after reduction by symmetries. In this direction, much work has been done in recent years (see, for instance, [BGN12; Ehl+05; FJ04; GNM20; GNM18; Jov10; Koz02; VV88]; see also [BY20] and the references therein).

- **Levi-Civita connections of Gauss Riemannian metrics associated with a kinetic nonholonomic system.** For a kinetic nonholonomic system $(g, \mathcal{D})$ with configuration space Q, in [Lew97] the author describes the set of linear connections on Q

  $$\nabla : \mathfrak{X}(Q) \times \mathfrak{X}(Q) \to \mathfrak{X}(Q)$$

  which satisfy the condition

  $$\nabla_X Y = \mathcal{P}(\nabla^g_X Y), \quad \text{for } X \in \mathfrak{X}(Q) \text{ and } Y \in \Gamma(\mathcal{D}),$$

  where $\Gamma(\mathcal{D})$ is the set of sections of the distribution $\mathcal{D}$, $\mathcal{P} : \mathfrak{X}(Q) \to \Gamma(\mathcal{D})$ is the orthogonal projector and ∇^g is the Levi-Civita connection of g. The nonholonomic connection ∇^{nh}, considered in Section 5.1

(see equation (5.1.2)) is a particular example of such connections. In fact, in [Lew97], the author proves that the geodesics of any of these connections with initial velocity in $\mathcal{D}$ are just the trajectories of the kinetic nonholonomic system $(g, \mathcal{D})$. Another important type of linear connections considered in [Lew97], which are related with the system $(g, \mathcal{D})$, are the so-called energy-preserving connections. A connection ∇ on Q is *energy-preserving* for the system $(g, \mathcal{D})$ if the kinetic energy associated with g is constant along the geodesics of ∇. So, it would be interesting to discuss relations between the previous family of connections and the Levi-Civita connections associated with the Gauss Riemannian metrics g_q^{nh} on the submanifolds $\mathcal{M}_q^{nh}$, with $q \in Q$.

- **Kinetic nonholonomic systems with affine constraints.** It would also be interesting to formulate the analogous results for the special case of kinetic nonholonomic systems with affine constraints with a moving energy (see [FS16; FGNS18]). The argument would be very similar since, in this case, there exists a change of coordinates that transforms the system into a nonholonomic system with linear constraints where the moving energy is precisely the energy of the transformed system.

Some of the previous problems on kinetic nonholonomic systems may be posed for the more general case of nonholonomic Lagrangian systems of mechanical type. The first steps in this direction were given on Section 5.3, where we generalize the results proved before to the general case, using the nonholonomic Maupertuis principle. However, we could ask if there is a more general extension as we explain in the following item:

- **Nonholonomic Lagrangian systems of mechanical type versus unconstrained Lagrangian systems of the same type.** The Lagrangian function $L : TQ \to \mathbb{R}$ of an unconstrained mechanical system is given by

$$L_{(g,V)}(u_q) = \frac{1}{2} g(q)(u_q, u_q) - V(q), \quad \text{for } u_q \in T_qQ,$$

where g is a Riemannian metric on Q and $V : Q \to \mathbb{R}$ is the potential energy. In the presence of a constraint distribution $\mathcal{D}$ on Q, we have a nonholonomic Lagrangian system $(L_{(g,V)}, \mathcal{D})$ of mechanical type. So, a natural question arises: does there exist an unconstrained Lagrangian system of mechanical type such that the nonholonomic trajectories of the system $(L_{(g,V)}, \mathcal{D})$ with a fixed starting point $q \in Q$ are the trajectories of the unconstrained Lagrangian system with the same starting point q and initial velocity in $\mathcal{D}$?

This a slightly different question than the one we answered in Section 5.3 since the goal here is to relate a constrained mechanical system with an unconstrained mechanical system.

Nonholonomic Jacobi fields

In chapter 6, we have introduced a natural definition of nonholonomic Jacobi fields for nonholonomic systems in pure Riemannian geometric terms, we have also characterized them and finally we have given some equivalent versions of the nonholonomic Jacobi equation.

- **Conjugate points and minimizing properties.** In a future research, we will continue this program studying conjugate points, the possible relation with minimizing properties of nonholonomic geodesics where the exponential nonholonomic map will play an important role.

  In fact, in Chapter 6, we introduced in a natural way the notion of a nonholonomic Jacobi field along a nonholonomic trajectory $c : I \to Q$ of a system. In particular, we have defined a nonholonomic Jacobi field Z over $c : I \to Q$ to be the infinitesimal variation of a one-parameter family of nonholonomic trajectories with initial trajectory c. Now, we ask the following question: can we relate the concept of nonholonomic Jacobi fields with the results contained in Chapter 5?

  Indeed, given a kinetic nonholonomic system, if we consider its nonholonomic trajectories with the same starting point $q \in Q$, we may construct a nonholonomic Jacobi field Z associated to this family of trajectories (in particular, Z is the zero vector at the initial point q). Then using Theorem 5.2.4, we deduce that there exists a Riemannian metric on $\mathcal{M}_q^{nh}$ such that the nonholonomic trajectories are geodesics with respect to this Riemannian metric and Z is a Riemannian Jacobi field along the geodesic c. On the other hand, as we know (see, for instance, [Car92; O'N83]), Riemannian Jacobi fields play an important role in the study of the singularities of the Riemannian exponential map and the minimizing properties of the Riemannian geodesics. So, after the previous comments, one may pose the following question: is it possible, using the nonholonomic Jacobi fields, to discuss the singularities of the nonholonomic exponential map and the minimizing properties of the nonholonomic trajectories, as in the case of Riemannian geometry?

Indeed, if $c_{qp} : I = [0,1] \to Q$ is a nonholonomic trajectory joining the points q, p then, as in standard Riemannian geometry, we say that p is *nonholonomic conjugate* to q along c_{qp} if there is a nonholonomic Jacobi field W vanishing at q and p but not identically zero. In terms of the nonholonomic exponential map, this means that if $v_q \in \mathcal{D}_q$ is the initial velocity such that $\exp_q^{nh}(v_q) = p$, then

$$T_{v_q}\exp_q^{nh} : T_{v_q}\mathcal{D}_q \equiv \mathcal{D}_q \longrightarrow T_pQ$$

is no longer injective. Now, suppose that g_q^{nh} is a Gauss metric on $\mathcal{M}_q^{nh}$ and that g^{nh} is an extension of g_q^{nh} to Q such that $\mathcal{M}_q^{nh}$ is totally geodesic at q with respect to g^{nh}. Then, one may deduce that the trajectory c_{qp} stops being length minimizing for g^{nh} past the nonholonomic conjugate point p. In particular, this notion allows us to define the *conjugate locus* of q as the set of points p such that p is the first conjugate point to q along some nonholonomic trajectory. As a consequence, for the qualitative study of nonholonomic dynamics, it is important to study the zeros of the nonholonomic Jacobi fields and/or the zeros of the Jacobi fields for the associated Gauss metrics.

- **Reduction of nonholonomic mechanical systems with symmetries.** Another interesting goal, to be covered in an upcoming publication, is to extend the results of chapter 6 on Jacobi fields to the reduction of nonholonomic mechanical systems with symmetries. This kind of systems have been extensively discussed in the literature (see [Koi92; Blo+96a; Can+98; GG08; Cor+09b; Gra+09; LMD10; Bal14; Bal17; BF15]). On the other hand, attending to the following facts: i) if the group of isometries of a connected Riemannian manifold M of dimension m has maximum dimension $\frac{m(m+1)}{2}$ then M is of constant curvature (see [KN63]); and ii) the expression of the curvature tensor in such spaces is very simple and, thus, the computation of the Jacobi fields also is so (see, for instance, [Lee97]); we suspect that the computation of the nonholonomic Jacobi fields is also probably simple for a kinematic nonholonomic system which admits many symmetries. We will keep investigating the relation between symmetries and solutions of the nonholonomic Jacobi field in the future.

- **Jacobi fields in sub-Riemannian geometry** We remark that many of the results in chapter 6 may be extended for Jacobi fields in sub-Riemannian geometry [Ghe+20].

Discrete Nonholonomic mechanics

In Chapter 7, we have precisely identified the exact discrete equations for a nonholonomic system. The main ingredients were the definition of the exponential map for a constrained second-order differential equation allowing us to define the exact discrete nonholonomic constraint submanifold. Then, we define the main discrete elements that appear on the definition of the exact discrete nonholonomic equations. The special form of these equations allow us to introduce a new family of nonholonomic integrators showing in numerical computations the excellent behaviour of the energy.

In the future, the following topics might be investigated:

- **An intrinsic version of discrete nonholonomic mechanics in $\mathcal{M}_h^{e,nh}$.** In a future research, we will study an intrinsic version of discrete nonholonomic mechanics in $\mathcal{M}_h^{e,nh}$ following the steps given at the end of Section 7.3. Indeed, we may ask if there exists a formulation of the modified Lagrange-d'Alembert principle in a framework similar to that on [MDM06], possibly by replacing a groupoid by an even more general structure.

  Moreover, since typically nonholonomic systems admit symmetries (see [Blo15]), we will study the reduction of the discrete counterparts following the results by [Igl+08].

- **Backward error analysis for nonholonomic mechanics** Since we have a discrete exact version, we could study error analysis and backward error analysis for nonholonomic mechanics (see [MW01; HLW10] and Section 3.7.6, for the case of unconstrained Lagrangian systems).

  In that sense, we aim to find a nonholonomic version of Theorem 3.7.14. Since we already know the exact discrete nonholonomic flow and the objects that we must approximate when discretizing a nonholonomic system (discrete constraint submanifold, discrete Lagrangian and discrete forces), it is natural to wonder if we may estimate the order of the numerical method arising from modified Lagrange-d'Alembert principle just by approximating the previous objects, as in the pure variational case.

- **Geometric integrators for kinetic nonholonomic systems.** In [MV20], the authors show that several constructions of geometric integrators for nonholonomic mechanics that appear in the literature do not behave well for general nonholonomic systems. Therefore, the problem of finding structure preserving integrators for nonholonomic systems is completely open. However, observe that Theorem 5.2.4 opens up the possibility of considering a new class of variational type integrators for nonholonomic mechanics.

  For instance, we may consider a retraction map $R_h : TQ \to Q \times Q$ on a manifold Q (see [AMS08]) and define the following discrete nonholonomic submanifold of $Q \times Q$:

  $$R_h(\mathcal{D}) = \mathcal{M}^{nh,d} .$$

  From the properties of retraction maps, we have that if $q \in Q$, then $(R_{h,q})\big|\mathcal{D}_q$ is a diffeomorphism onto its image $R_{h,q}(\mathcal{D}_q) = \mathcal{M}_q^{nh,d}$ in a neighbourhood of 0_q. In a future paper, we will explore the construction of variational type integrators on $\mathcal{M}^{nh,d}$. One possibility is to induce first, a Riemannian metric $g_q^{nh,d}$ on $\mathcal{M}_q^{nh,d}$ for each $q \in Q$, as in Theorem 5.2.4. This metric is determined by the pullback by $\left(R_{h,q}\big|_{\mathcal{D}_q}\right)^{-1}$ of a Riemannian metric on $\mathcal{D}_q$ verifying Gauss condition.

  Then we can define a discrete Lagrangian $L_d^{nh} : \mathcal{M}^{nh,d} \to \mathbb{R}$ as an approximation of the corresponding action:

  $$L_d^{nh}(q, q') \approx \frac{1}{2} \int_0^h g_q^{nh,d}(c(t))(\dot{c}(t), \dot{c}(t)) \, dt$$

  where $c : [0, h] \to \mathcal{M}_q^{nh,d}$ is the unique geodesic associated with the Riemannian metric $g_q^{nh,d}$ satisfying $c(0) = q$ and $c(h) = q'$.

  Unfortunately, this class of methods typically exhibit a quick deviation from the exact trajectory essentially due to the error in the approximation of the exact discrete constraint space $\mathcal{M}_q^{nh,e}$.

A much more simple and efficient possibility is exploring the idea that the nonholonomic exponential map, transforms lines through the origin into geodesics of $g_q^{nh,d}$. Given a retraction map $R_{h,q} : \mathcal{D}_q \to Q$ we may define the following integrator:

$$
\begin{aligned}
&q_1 = R_{h,q_0}(v_{q_0}) \\
&R_{2h\beta,q_0}(v_{q_0}) = R_{(2\beta-1)h,q_1}(v_{q_1})
\end{aligned}
$$

This integrator will be studied in a future paper and, as by-product of this approach, we will obtain, for a suitable choice of retraction map, a generalization of the Newmark method to the case of nonholonomic systems.

Discrete contact mechanics

In Chapter 8, we have deepen the geometry of discrete contact mechanics fully explaining the discretizations introduced in [VBS19]. We have done a detailed study of the discrete Herglotz principle and its geometric properties, including the discrete Legendre transforms and the associated discrete Lagrangian and Hamiltonian flows. Moreover, we have analysed the existence of dissipated discrete quantities related with discrete symmetries of the system and the construction of the exact discrete Lagrangian function giving the correspondence between the discrete and continuous system.

In future work, we will study some the following problems:

- The **variational error analysis** allowing us to estimate the error order of the proposed methods just from the error of approximation of the exact discrete Lagrangian function, that is, how well the discrete Lagrangian function matches the exact discrete Lagrangian function [MW01; PC09].

- **The extension of the theory of Morse functions to Legendrian submanifolds** (see [LM87; BLn+19; Fer+17]) allowing to introduce higher-order methods for contact Lagrangian systems. For instance, this theory will give a complete geometric explanation of other possible discretizations of the phase space, as for instance, the one used by [VBS19] which is $Q \times Q \times \mathbb{R}^2$ instead of $Q \times Q \times \mathbb{R}$.

- **Simulation of some thermodynamic systems**: Contact geometry has been used to model some thermodynamics systems (see [Bra18]). Recently, it was shown that a modification of the contact Hamiltonian vector field had many advantages in the modelling of thermodynamic systems (see [Sim+20]). It would be interesting to study if our approach could be used to integrate this type of models.

Chapter 10

Conclusiones e investigación futura

En esta tesis, hemos desarrollado cuatro líneas principales de investigación: trayectorias radiales de sistemas mecánicos no holónomos (Capítulo 5), campos Jacobi no holónomos (Capítulo 6), mecánica no holónoma discreta (Capítulo 7) y mecánica de contact discreta (Capítulo 8). Resumiremos a continuación nuestros hallazgos y dejaremos la puerta abierta a nuevos resultados según estas líneas:

Trayectorias radiales de sistemas noholónomos mecánicos

Dado un sistema cinético no holónomo, con espacio de configuración Q, hemos identificado y caracterizado la familia de métricas Riemannianas en la imagen $\mathcal{M}_q^{nh}$ de la aplicación exponencial no holónoma en un punto fijo $q \in Q$ que satisface la siguiente propiedad crucial: que la minimización de geodésicas con punto de partida q y con respecto a estas métricas son, para tiempos suficientemente pequeños, nada más que las trayectorias no holónomas que comienzan en el mismo punto q.

También hemos demostrado que tales métricas en $\mathcal{M}_q^{nh}$ siempre existen e ilustramos estos hechos con varios ejemplos. Observamos que estos resultados son sorprendentes e inesperados, ya que hasta ahora la dinámica no holónoma se consideraba no variacional y ahora proporcionamos una nueva perspectiva bajo la cual se convierte en variacional, en el sentido de que por lo menos las trayectorias radiales son geodésicas repecto a una métrica Riemanniana.

Después de estos resultados, queda mucho trabajo por hacer. De hecho, nuestra idea es desarrollar un programa de investigación para discutir las propiedades geométricas de las trayectorias no holónomas para un sistema cinético no holónomo. Algunos de los problemas que se tratarán en este programa son los siguientes:

- **Flujos geodésicos y el flujo cinético no holónomo.** Sea $(g, \mathcal{D})$ un sistema cinético no holónomo con espacio de configuración Q, $\Gamma_{(g,\mathcal{D})} \in \mathfrak{X}(\mathcal{D})$ el flujo cinético no holónomo y $\exp^{\Gamma_{(g,\mathcal{D})}} : M^{\Gamma_{(g,\mathcal{D})}} \subseteq \mathcal{D} \to Q \times Q$ la aplicación exponencial no holónoma introducida en el capítulo 4 (ver Sección 4.3). Entonces, siguiendo el teorema 4.3.2, se deduce que existe una entorno abierto $\mathcal{U}_{0(Q)}$ de la sección cero $0(Q)$ en $\mathcal{D}$ tal que

$$(\exp^{\Gamma_{(g,\mathcal{D})}})_{|\mathcal{U}_{0(Q)}} : \mathcal{U}_{0(Q)} \subseteq \mathcal{D} \to Q \times Q$$

es un embebimiento. Denotemos por $\mathcal{M}^{nh} = \exp^{\Gamma_{(g,\mathcal{D})}}(\mathcal{U}_{0(Q)})$. Queda claro que $\mathcal{M}^{nh}$ es una subvariedad embebida de $Q \times Q$,

$$\exp^{nh} := (\exp^{\Gamma_{(g,\mathcal{D})}})_{|\mathcal{U}_{0(Q)}} : \mathcal{U}_{0(Q)} \subseteq \mathcal{D} \to \mathcal{M}^{nh} \subseteq Q \times Q$$

es un difeomorfismo y el siguiente diagrama

$$\begin{array}{ccc} \mathcal{U}_{0(Q)} \subseteq \mathcal{D} & \xrightarrow{\exp^{nh}} & \mathcal{M}^{nh} \subseteq Q \times Q \\ & {\scriptstyle (\tau_D)_{|\mathcal{U}_{0(Q)}}} \searrow \quad \swarrow {\scriptstyle (\mathrm{pr}_1)_{|\mathcal{M}^{nh}}} & \\ & Q & \end{array}$$

es conmutativo, donde $\tau_{\mathcal{D}} : \mathcal{D} \to Q$ es la proyección del fibrado vectorial. Nótese que

$$(\mathrm{pr}_1)^{-1}_{|\mathcal{M}^{nh}}(q) = \exp^{nh}_q(\mathcal{U}_{0(Q)} \cap \mathcal{D}_q) \simeq \mathcal{M}^{nh}_q, \quad \text{for } q \in Q.$$

Denotemos ahora por $\Gamma^{nh}_{(g,\mathcal{D})}$ el flujo no holónomo considerado como un campo vectorial sobre $\mathcal{M}^{nh}$, a través del difeomorfismo $\exp^{nh}$. Entonces, procediendo como en el capítulo 5, se puede encontrar una familia de métricas de fibrado

$$g^{nh} : V(\mathrm{pr}_1)_{|\mathcal{M}^{nh}} \times_{\mathcal{M}^{nh}} V(\mathrm{pr}_1)_{|\mathcal{M}^{nh}} \to \mathbb{R}$$

en el fibrado vertical del fibrado con proyección $(\mathrm{pr}_1)_{\mathcal{M}^{nh}} : \mathcal{M}^{nh} \to Q$ tal que si $q \in Q$, las geodésicas minimizantes sobre $\mathcal{M}^{nh}_q \subseteq \mathcal{M}^{nh}$ con punto de partida q son, para tiempos suficientemente pequeños, las trayectorias no holónomas con el mismo punto de partida q.

Además, se puede considerar el flujo geodésico Γ^{nh} asociado a una de estas métricas g^{nh} como un campo vectorial sobre $V(\mathrm{pr}_1)_{\mathcal{M}^{nh}}$. Entonces, sería interesante discutir la relación entre el campo vectorial $\Gamma^{g^{nh}}$ en $V(\mathrm{pr}_1)_{\mathcal{M}^{nh}}$ y el flujo no holónomo $\Gamma^{nh}_{(g,\mathcal{D})}$ en $\mathcal{M}^{nh}$.

- **Lagrangianización de sistemas cinéticos no holónomos**. Sea $(g, \mathcal{D})$ un sistema cinético no holónomo con espacio de configuración Q. Después de los resultados y ejemplos del capítulo 5, surge otra pregunta natural: ¿bajo qué condiciones se puede obtener una Lagrangianización cinética del sistema $(g, \mathcal{D})$? En otras palabras, ¿bajo qué condiciones existe una métrica riemanniana g^{nh} en toda la variedad ambiente Q, en lugar de sólo en $\mathcal{M}^{nh}_q$, tal que las trayectorias cinéticas no holónomas para el sistema $(g, \mathcal{D})$ son justo las geodésicas de la métrica g^{nh} con velocidad inicial en $\mathcal{D}$?

  Obsérvese que hay ejemplos de sistemas cinéticos no holónomos que admiten tales métricas en todo el espacio de configuración y, a pesar de ello, la distribución de ligaduras sigue sin ser integrable (véase el ejemplo 5.2.14 en el capítulo 5). Por otro lado, el Teorema 5.2.4 puede considerarse como el primer paso para dar respuesta a la difícil pregunta anterior. También observamos que si el sistema admite una Lagrangianización cinética entonces, utilizando la transformación de Legendre asociada al sistema lagrangiano cinético inducido por la métrica riemanniana g^{nh}, se puede producir una formulación hamiltoniana del sistema no holónomo original. Así pues, nuestra pregunta está relacionada con un problema clásico de la mecánica no holónoma: el llamado problema de Hamiltonización. Este problema discute si un sistema no holónomo admite una formulación hamiltoniana después de la reducción por simetrías. En esta dirección, se ha trabajado mucho en los últimos años (véase, por ejemplo, [BGN12; Ehl+05; FJ04; GNM20; GNM18; Jov10; Koz02; VV88]; véase también [BY20] y las referencias en ella).

- **Conexiones Levi-Civita de las métricas de Gauss Riemannianas asociadas a un sistema cinético no holónomo.** Para un sistema cinético no holónomo $(g, \mathcal{D})$ con espacio de configuración Q, en [Lew97] el autor describe el conjunto de conexiones lineales sobre Q

$$\nabla : \mathfrak{X}(Q) \times \mathfrak{X}(Q) \to \mathfrak{X}(Q)$$

que satisfacen la condición

$$\nabla_X Y = \mathcal{P}(\nabla^g_X Y), \quad \text{for } X \in \mathfrak{X}(Q) \text{ and } Y \in \Gamma(\mathcal{D}),$$

donde $\Gamma(\mathcal{D})$ es el conjunto de secciones de la distribución $\mathcal{D}$, $\mathcal{P} : \mathfrak{X}(Q) \to \Gamma(\mathcal{D})$ es el proyector ortogonal y ∇^g es la conexión Levi-Civita de g. La conexión no holónoma ∇^{nh}, considerada en la sección 5.1 (véase la ecuación 5.1.2) es un ejemplo particular de tales conexiones. De hecho, en [Lew97], el autor demuestra que las geodésicas de cualquiera de estas conexiones con velocidad inicial en $\mathcal{D}$ son justo las trayectorias del sistema cinético no holónomo $(g, \mathcal{D})$. Otro tipo importante de conexiones lineales consideradas en [Lew97], que están relacionadas con el sistema $(g, \mathcal{D})$, son las llamadas conexiones que preservan energía. Una conexión ∇ sobre Q se dice que *preserva energía* para el sistema $(g, \mathcal{D})$ si la energía cinética asociada a g es constante a lo largo de las geodésicas de ∇. Así pues, sería interesante discutir las relaciones entre la familia de conexiones anterior y las conexiones de Levi-Civita asociadas a las métricas riemannianas de Gauss g_q^{nh} en las subvariedades $\mathcal{M}_q^{nh}$, con $q \in Q$.

- **Sistemas cinéticos no holónomos con restricciones afines.** También sería interesante formular los resultados análogos para el caso especial de los sistemas cinéticos no holónomos con restricciones afines con una energía en movimiento (ver [FS16; FGNS18]). El argumento sería muy similar ya que, en este caso, existe un cambio de coordenadas que transforma el sistema en un sistema no holónomo con restricciones lineales donde la energía móvil es precisamente la energía del sistema transformado.

Algunos de los problemas anteriores sobre sistemas cinéticos no holónomos pueden plantearse para el caso más general de sistemas lagrangianos no holónomos de tipo mecánico. Los primeros pasos en esta dirección se han dado en la sección 5.3, donde generalizamos los resultados demostrados anteriormente al caso general, utilizando el principio no holónomo de Maupertuis. Sin embargo, podríamos preguntarnos si existe una extensión más general, como explicamos en el siguiente punto:

- **Sistemas lagrangianos no holónomos de tipo mecánico frente a sistemas lagrangianos sin restricciones del mismo tipo.** La función Lagrangiana $L : TQ \to \mathbb{R}$ de un sistema mecánico no restringido viene dada por

  $$L_{(g,V)}(u_q) = \frac{1}{2} g(q)(u_q, u_q) - V(q), \quad \text{for } u_q \in T_qQ,$$

  donde g es una métrica riemanniana sobre Q y $V : Q \to \mathbb{R}$ es la energía potencial. En presencia de una distribución de ligaduras $\mathcal{D}$ sobre Q, tenemos un sistema lagrangiano no holónomo $(L_{(g,V)}, \mathcal{D})$ de tipo mecánico. Entonces, surge una pregunta natural: ¿existe un sistema lagrangiano no restringido de tipo mecánico tal que las trayectorias no holónomas del sistema $(L_{(g,V)}, \mathcal{D})$ con un punto de partida fijo $q \in Q$ son las trayectorias del sistema lagrangiano no restringido con el mismo punto de partida q y velocidad inicial en $\mathcal{D}$?

  Esta es una pregunta ligeramente diferente a la que respondimos en la sección 5.3 ya que el objetivo aquí es relacionar un sistema mecánico restringido con un sistema mecánico no restringido.

Campos de Jacobi no holónomos

En el capítulo 6 hemos introducido una definición natural de campos de Jacobi no holónomos para sistemas no holónomos en términos geométricos riemannianos puros, también los hemos caracterizado y finalmente hemos dado algunas versiones equivalentes de la ecuación de Jacobi no holónoma.

- **Puntos conjugados y propiedades minimizantes.** En investigación futura, continuaremos este programa estudiando los puntos conjugados, la posible relación con las propiedades minimizantes de las geodésicas no holónomas donde la aplicación exponencial no holónoma jugará un papel importante.

  De hecho, en el capítulo 6, introducimos de forma natural la noción de campo de Jacobi no holónomo a lo largo de una trayectoria no holónoma $c : I \to Q$ de un sistema. En particular, hemos definido un campo de Jacobi no holónomo Z sobre $c : I \to Q$ como la variación infinitesimal de una familia uniparamétrica de trayectorias no holónomas con trayectoria inicial c. Ahora, nos planteamos la siguiente pregunta: ¿podemos relacionar el concepto de campos de Jacobi no holónomos con los resultados contenidos en el capítulo 5?

En efecto, dado un sistema cinético no holónomo, si consideramos sus trayectorias no holónomas con el mismo punto inicial $q \in Q$, podemos construir un campo de Jacobi no holónomo Z asociado a esta familia de trayectorias (en particular, Z es el vector cero en el punto inicial q). Entonces, utilizando el Teorema 5.2.4, deducimos que existe una métrica riemanniana sobre $\mathcal{M}_q^{nh}$ tal que las trayectorias no holónomas son geodésicas con respecto a esta métrica riemanniana y Z es un campo de Jacobi riemanniano a lo largo de la geodésica c. Por otra parte, como sabemos (véase, por ejemplo, [Car92; O'N83]), los campos de Jacobi riemannianos juegan un papel importante en el estudio de las singularidades de la aplicación exponencial riemanniana y las propiedades minimizantes de las geodésicas riemannianas. Así, tras los comentarios anteriores, cabe plantear la siguiente pregunta: ¿es posible, utilizando los campos de Jacobi no holónomos, discutir las singularidades de la aplicación exponencial no holónoma y las propiedades minimizantes de las trayectorias no holónomas, como en el caso de la geometría riemanniana?

En efecto, si $c_{qp} : I = [0, 1] \to Q$ es una trayectoria no holonómica que une los puntos q, p entonces, como en la geometría riemanniana estándar, decimos que p es *nonholonómico conjugado* a q a lo largo de c_{qp} si existe un campo de Jacobi no holonómico W que desaparece en q y p pero no es idéntico a cero. En términos del mapa exponencial no holonómico, esto significa que si $v_q \in \mathcal{D}_q$ es la velocidad inicial tal que $\exp_q^{nh}(v_q) = p$, entonces

$$T_{v_q}\exp_q^{nh} : T_{v_q}\mathcal{D}_q \equiv \mathcal{D}_q \longrightarrow T_pQ$$

ya no es inyectiva. Supongamos ahora que g_q^{nh} es una métrica de Gauss sobre $\mathcal{M}_q^{nh}$ y que g^{nh} es una extensión de g_q^{nh} a Q tal que $\mathcal{M}_q^{nh}$ es totalmente geodésica en q con respecto a g^{nh}. Entonces, se puede deducir que la trayectoria c_{qp} deja de ser minimizadora de longitud para g^{nh} pasado el punto conjugado no holonómico p. En particular, esta noción nos permite definir el *locus conjugado* de q como el conjunto de puntos

p tales que p es el primer punto conjugado a q a lo largo de alguna trayectoria no holonómica. En consecuencia, para el estudio cualitativo de la dinámica no holonómica, es importante estudiar los ceros de los campos de Jacobi no holonómicos y/o los ceros de los campos de Jacobi para las métricas de Gauss asociadas.

- **Reducción de sistemas mecánicos no holónomos con simetrías.** Otro objetivo interesante, que se tratará en una próxima publicación, es extender los resultados del capítulo 6 sobre campos de Jacobi a la reducción de sistemas mecánicos no holónomos con simetrías. Este tipo de sistemas ha sido ampliamente discutido en la literatura (ver [Koi92; Blo+96a; Can+98; GG08; Cor+09b; Gra+09; LMD10; Bal14; Bal17; BF15]). Por otro lado, atendiendo a los siguientes hechos: i) si el grupo de isometrías de una variedad riemanniana conexa M de dimensión m tiene dimensión máxima $\frac{m(m+1)}{2}$ entonces M es de curvatura constante (ver [KN63]); y ii) la expresión del tensor de curvatura en tales espacios es muy sencilla y, por tanto, el cálculo de los campos de Jacobi también lo es (véase, por ejemplo, [Lee97]); sospechamos que probablemente el cálculo de los campos de Jacobi no-holónomos es también sencillo para un sistema cinemático no-holónomo que admite muchas simetrías. Seguiremos investigando la relación entre las simetrías y las soluciones del campo de Jacobi no-holónomo en el futuro.

- **Campos de Jacobi en geometría sub-riemanniana** Observamos que muchos de los resultados del capítulo 6 pueden extenderse para campos de Jacobi en geometría subreimana [Ghe+20].

Mecánica discreta no holónoma

En el capítulo 7, hemos identificado con precisión las ecuaciones discretas exactas para un sistema no holónomo. Los ingredientes principales fueron la definición de la aplicación exponencial para una ecuación diferencial de segundo orden restringida que nos permite definir la subvariedad discreta exacta de ligaduras no holónoma. A continuación, definimos los principales elementos discretos que aparecen en la definición de las ecuaciones discretas exactas no holónomas. La forma especial de estas ecuaciones nos permite introducir una nueva familia de integradores no holónomos que muestran en los cálculos numéricos el excelente comportamiento de la energía.

- **Una versión intrínseca de la mecánica discreta no holónoma en $\mathcal{M}_h^{e,nh}$.** En investigación futura, estudiaremos una versión intrínseca de la mecánica discreta no holónoma en $\mathcal{M}_h^{e,nh}$ siguiendo los pasos dados al final de la Sección 7.3. De hecho, podemos preguntarnos si existe una formulación del principio de Lagrange-d'Alembert modificado en un marco similar al de [MDM06], posiblemente sustituyendo un groupoide por una estructura aún más general.

  Además, dado que los sistemas típicamente no holónomos admiten simetrías (véase [Blo15]), estudiaremos la reducción del análogo discreto, siguiendo los resultados de [Igl+08].

- **Análisis regresivo del error para la mecánica no holónoma.** Dado que tenemos una versión exacta discreta, podríamos estudiar el análisis de errores y el análisis regresivo del error para la mecánica no holónoma (véase [MW01; HLW10] y la sección 3.7.6, para el caso de sistemas lagrangianos no restringidos).

  En este sentido, nuestro objetivo es encontrar una versión no holónoma del Teorema 3.7.14. Dado que ya conocemos el flujo discreto no holónomo exacto y los objetos que debemos aproximar al discretizar un sistema no holónomo (subvariedad de ligaduras discreta, lagrangiano discreto y fuerzas discretas), es natural preguntarse si podemos estimar el orden del método numérico que surge del principio de Lagrange-d'Alembert modificado, simplemente aproximando los objetos anteriores, como en el caso variacional puro.

- **Integradores geométricos para sistemas cinéticos no holónomos.** En [MV20], los autores muestran que varias construcciones de integradores geométricos para la mecánica no holónoma que aparecen en la literatura no se comportan bien para sistemas generales no holónomos. Por lo tanto, el problema de encontrar integradores que preserven la estructura para sistemas no holónomos está completamente abierto. Sin embargo, obsérvese que el Teorema 5.2.4 abre la posibilidad de considerar una nueva clase de integradores de tipo variacional para la mecánica no holónoma.

  Por ejemplo, podemos considerar un mapa de retracción $R : TQ \to Q \times Q$ sobre una variedad Q (ver [AMS08]) y definir la siguiente subvariedad discreta no holónomo de $Q \times Q$:

$$R(\mathcal{D}) = \mathcal{M}^{nh,d} \; .$$

A partir de las propiedades de los mapas de retracción, tenemos que si $q \in Q$, entonces $(R_q)\big|\mathcal{D}_q$ es un difeomorfismo hacia su imagen $R(\mathcal{D}_q) = \mathcal{M}^{nh,d}_q$ en un entorno de 0_q. En un futuro trabajo, exploraremos la construcción de integradores de tipo variacional en $\mathcal{M}^{nh,d}$. Una posibilidad es inducir, en primer lugar, una métrica riemanniana $g^{nh,d}_q$ sobre $\mathcal{M}^{nh,d}_q$ para cada $q \in Q$, como en el Teorema 5.2.4. Esta métrica viene determinada por el pullback por $\left(R\big|_{D_q}\right)^{-1}$ de una métrica riemanniana sobre $\mathcal{D}_q$ verificando la condición de Gauss.

Entonces podemos definir un Lagrangiano discreto $L^{nh}_d : \mathcal{M}^{nh,d} \to \mathbb{R}$ como una aproximación de la acción correspondiente:

$$L^{nh}_d(q, q') \approx \frac{1}{2} \int_0^h g^{nh,d}_q(c(t))(\dot{c}(t), \dot{c}(t)) \; dt$$

donde $c : [0, h] \to \mathcal{M}^{nh,d}_q$ es la única curva geodésica para $g^{nh,d}_q$ que satisface $c(0) = q$ y $c(h) = q'$.

Desgraciadamente, esta clase de métodos suele presentar una rápida desviación de la trayectoria exacta debido esencialmente al error en la aproximación del espacio de ligaduras discreto exacto $\mathcal{M}^{nh,e}_q$.

Una posibilidad mucho más sencilla y eficiente es explorar la idea de que la aplicación exponencial no holónoma, transforma las rectas que pasan por el origen en geodésicas de $g^{nh,d}_q$. Dada una retracción $R_{h,q} : \mathcal{D}_q \to Q$ podemos definir el siguiente integrador::

$$\begin{aligned} &q_1 = R_{h,q_0}(v_{q_0}) \\ &R_{2h\beta,q_0}(v_{q_0}) = R_{(2\beta-1)h,q_1}(v_{q_1}) \end{aligned}$$

Este integrador será estudiado en un futuro trabajo y, como subproducto de este enfoque, obtendremos, para una elección adecuada de una retracción, una generalización del método de Newmark para el caso de sistemas no holónomos.

Mecánica discreta de contacto

En el capítulo 8, hemos profundizado en la geometría de la mecánica de contacto discreta explicando completamente las discretizaciones introducidas en [VBS19]. Hemos realizado un estudio detallado del principio discreto de Herglotz y sus propiedades geométricas, incluyendo las transformadas discretas de Legendre y los flujos discretos Lagrangianos y Hamiltonianos asociados. Además, hemos analizado la existencia de cantidades discretas disipadas relacionadas con simetrías discretas del sistema y la construcción de la función Lagrangiana discreta exacta dando la correspondencia entre el sistema discreto y el continuo.

En futuros trabajos, estudiaremos algunos de los siguientes problemas:

- El **análisis del error variacional** nos permite estimar el orden de error de los métodos propuestos sólo a partir del error de aproximación de la función Lagrangiana discreta exacta, es decir, lo bien que la función Lagrangiana discreta coincide con la función Lagrangiana discreta exacta [MW01; PC09].

- **La extensión de la teoría de las funciones de Morse a las subvariedades Legendrianas** (veáse [LM87; BLn+19; Fer+17]) que permite introducir métodos de orden superior para sistemas Lagrangianos de contacto. Por ejemplo, esta teoría dará una explicación geométrica completa de otras posibles discretizaciones del espacio de fase, como por ejemplo, la utilizada por [VBS19] que es $Q \times Q \times \mathbb{R}^2$ en lugar de $Q \times Q \times \mathbb{R}$.

- **Simulación de algunos sistemas termodinámicos**: La geometría de contacto se ha utilizado para modelar algunos sistemas termodinámicos (véase [Bra18]). Recientemente, se demostró que una modificación del campo vectorial hamiltoniano de contacto tenía muchas ventajas en el modelado de sistemas termodinámicos (ver [Sim+20]). Sería interesante estudiar si nuestro método se podría utilizar para integrar este tipo de modelos.

Appendix A

Complete and vertical lifts in mechanics

The main result of this section is to prove how a pseudo-Riemannian metric h is related to the Poincaré-Cartan two-form induced by this metric. In particular, the action of the later on complete and vertical lifts may be expressed with objects that depend only on the metric structure.

Recall that when we have a manifold Q and a Lagrangian L on its tangent bundle, the Poincaré-Cartan one-form is defined to be $\theta_L = S^* dL$.

In agreement with the previous notation, whenever we are given a $(0, 2)$-tensor h on Q, we will denote by $L_h : TQ \to \mathbb{R}$ the Lagrangian function associated with h defined by

$$L_h(v) = \frac{1}{2} h(v, v), \quad v \in TQ.$$

Also we will denote by

$$\flat_h : TQ \to T^*Q,$$

the musical isomorphism associated with h by $\flat_h(X)(Y) = h(X, Y)$ for all $X, Y \in \mathfrak{X}(Q)$.

During the remaining of this section, we will denote just by ∇ the Levi-Civita connection with respect to a symmetric non-degenerate $(0, 2)$-tensor, whenever it is clear from the context to which tensor it is associated.

Now we will see how complete and vertical lifts act on metric Lagrangians.

Lemma A.0.1. *Let h be a pseudo-Riemannian metric and L_h its associate Lagrangian. Given $X \in \mathfrak{X}(Q)$ we have that*

$$X^c(L_h) = L_{\mathcal{L}_X h}, \quad X^v(L_h) = \widehat{\flat_h(X)}, \tag{A.0.1}$$

where $L_{\mathcal{L}_X h} : TQ \to \mathbb{R}$ denotes the Lagrangian function associated to the $(0,2)$-tensor $\mathcal{L}_X h$.

Proof. Let us prove the result on natural coordinates. Let (q^i) be coordinates on Q and (q^i, v^i) be the natural coordinates on TQ. Let $X = X^i \frac{\partial}{\partial q^i}$ and $L_h = \frac{1}{2} h_{ij} v^i v^j$. Then

$$X^v(L_h) = X^i h_{ij} v^j = \widehat{\flat_h(X)}$$

and

$$X^c(L_h) = \frac{1}{2} X^k \frac{\partial h_{ij}}{\partial q^k} v^i v^j + v^i \frac{\partial X^k}{\partial q^i} h_{kj} v^j.$$

On the other hand

$$L_{\mathcal{L}_X h} = \frac{1}{2}\left(X^k \frac{\partial h_{ij}}{\partial q^k} + h_{kj} \frac{\partial X^k}{\partial q^i} + h_{ik} \frac{\partial X^k}{\partial q^j} \right) v^i v^j.$$

Since h is symmetric and changing indices $i \leftrightarrow j$ in the last term above, we get

$$L_{\mathcal{L}_X h} = \frac{1}{2} X^k \frac{\partial h_{ij}}{\partial q^k} v^i v^j + h_{kj} \frac{\partial X^k}{\partial q^i} v^i v^j,$$

which equals $X^c(L_h)$. □

The fundamental formula of Riemannian geometry given by (2.1.3) allows us to express the Lagrangian function $L_{\mathcal{L}_X h}$ introduced before in terms of a new Lagrangian function associated with the $(0,2)$-tensor

$$(\nabla^h X)(Y, Z) = h(\nabla^h_Y X, Z), \tag{A.0.2}$$

where ∇^h is the Levi-Civita connection with respect to h.

Lemma A.0.2. *The Lagrangian function $L_{\mathcal{L}_X h}$ coincides with the Lagrangian function $2L_{(\nabla^h X)}$.*

Proof. Given any $Y \in \mathfrak{X}(Q)$, by skew-symmetry of the exterior derivative one has that

$$\begin{aligned} L_{\mathcal{L}_X h} \circ Y &= \frac{1}{2}\mathcal{L}_X h(Y,Y) = h(\nabla^h_Y X, Y) - \frac{1}{2} d(\flat_h(X))(Y,Y) \\ &= (\nabla^h X)(Y,Y) = 2L_{(\nabla^h X)} \circ Y. \end{aligned}$$

□

Lemma A.0.3. *Let h be a pseudo-Riemannian metric on Q, L_h is its associated Lagrangian function and $\omega_{L_h} = -d\theta_{L_h}$ the corresponding Poincaré-Cartan 2-form. If $\flat_{\omega_{L_h}}$ and $\flat_h$ denote the musical isomorphisms associated to the symplectic form and to the tensor h, respectively, then for every $X \in \mathfrak{X}(Q)$ the Poincaré-Cartan 1-form acts on vertical and complete lifts of X according to*

$$\theta_{L_h}(X^v) = 0, \quad \theta_{L_h}(X^c) = \widehat{\flat_h(X)} \tag{A.0.3}$$

and the Poincaré-Cartan 2-form acts according to

$$\begin{aligned} &\omega_{L_h}(X^c, Y^c) = d(\widehat{\flat_h(X)})(Y^c) - 2\theta_{L_{(\nabla^h X)}}(Y^c), \\ &\omega_{L_h}(X^c, Y^v) = d(\widehat{\flat_h(X)})(Y^v), \quad \omega_{L_h}(X^v, Y^v) = 0, \end{aligned} \tag{A.0.4}$$

where $(\nabla^h X)$ is the (0,2)-tensor defined in (A.0.2). Hence, we may also write

$$\flat_{\omega_{L_h}}(X^v) = -(\flat_h(X))^v, \quad \flat_{\omega_{L_h}}(X^c) = d(\widehat{\flat_h(X)}) - 2\theta_{L_{(\nabla^h X)}}. \tag{A.0.5}$$

Proof. Recalling the definition of θ_{L_h}, the formulas in the statement are rewritten as

$$\theta_{L_h}(X^v) = S^*(dL_h)(X^v), \quad \theta_{L_h}(X^c) = S^*(dL_h)(X^c),$$

respectively. Now applying (2.4.19) we immediately prove that $\theta_{L_h}(X^v) = 0$ and

$$\theta_{L_h}(X^c) = X^v(L_h).$$

By the previous Lemma we conclude $\theta_{L_h}(X^c) = \widehat{\flat_h(X)}$.

Choosing an arbitrary $Y \in \mathfrak{X}(Q)$, we will now evaluate the symplectic form over complete and vertical lifts in order to find the desired formulas for $\flat_{\omega_{L_h}}(X^c)$ and $\flat_{\omega_{L_h}}(X^v)$.

Using that ω_{L_h} is an exact symplectic form and the characterization of the exterior derivative of a 1-form we get

$$\omega_{L_h}(X^c, Y^c) = -X^c\left(\theta_{L_h}(Y^c)\right) + Y^c\left(\theta_{L_h}(X^c)\right) + \theta_{L_h}([X^c, Y^c]).$$

Using equations (A.0.3) we have just proved and the formulas in (2.4.13) we get

$$\omega_{L_h}(X^c, Y^c) = -X^c\left(\widehat{\flat_h(Y)}\right) + Y^c\left(\widehat{\flat_h(X)}\right) + \flat_h\widehat{([X, Y]}).$$

Applying now the definition of complete lift over fiberwise linear functions

$$\omega_{L_h}(X^c, Y^c) = -\widehat{\mathcal{L}_X\flat_h(Y)} + \widehat{\mathcal{L}_Y\flat_h(X)} + \flat_h\widehat{([X, Y]}).$$

Note that

$$\flat_h([X, Y])(Z) - \mathcal{L}_X(\flat_h(Y))(Z) = -\mathcal{L}_X h(Y, Z).$$

Hence, the right-hand side of the above equation may be rewritten using the musical isomorphism associated to the $(0, 2)$-tensor $\mathcal{L}_X h$ which we denote by $\flat_{\mathcal{L}_X h}$. So we deduce that

$$\omega_{L_h}(X^c, Y^c) = -\widehat{\flat_{\mathcal{L}_X h}(Y)} + \widehat{\mathcal{L}_Y\flat_h(X)}.$$

Now using again the relations in (A.0.3) and the definition of complete lift

$$\omega_{L_h}(X^c, Y^c) = -\theta_{L_{\mathcal{L}_X h}}(Y^c) + Y^c(\widehat{\flat_h(X)}).$$

Using that $L_{\mathcal{L}_X h} = 2L_{(\nabla^h X)}$ and rewriting the last term above, we finally get

$$\omega_{L_h}(X^c, Y^c) = -2\theta_{L_{(\nabla^h X)}}(Y^c) + d(\widehat{\flat_h(X)})(Y^c).$$

Proceeding analogously in the other cases we find that

$$\omega_{L_h}(X^c, Y^v) = -X^c\left(\theta_{L_h}(Y^v)\right) + Y^v\left(\theta_{L_h}(X^c)\right) + \theta_{L_h}([X^c, Y^v]).$$

Using (A.0.3) and (2.4.13)

$$\omega_{L_h}(X^c, Y^v) = Y^v\left(\widehat{\flat_h(X)}\right) + \theta_{L_h}([X, Y]^v).$$

Again using (A.0.3) we conclude

$$\omega_{L_h}(X^c, Y^v) = d\left(\widehat{\flat_h(X)}\right)(Y^v).$$

Note also that $\theta_{L_{(\nabla^h X)}}(Y^v) = 0$ is also implied by (A.0.3). Therefore we have concluded the proof of the expression for $\flat_{\omega_{L_h}}(X^c)$.

Let us now prove the expression for $\flat_{\omega_{L_h}}(X^v)$. Let us use the same strategy and compute

$$\omega_{L_h}(X^v, Y^c) = -X^v\left(\theta_{L_h}(Y^c)\right) + Y^c\left(\theta_{L_h}(X^v)\right) + \theta_{L_h}([X^v, Y^c]).$$

Analogously,

$$\omega_{L_h}(X^v, Y^c) = -X^v\left(\widehat{\flat_h(Y)}\right).$$

Now, using the definition of vertical lift

$$\omega_{L_h}(X^v, Y^c) = -\flat_h(Y)(X) \circ \tau_Q.$$

By symmetry of h, we may rewrite the last line as

$$\omega_{L_h}(X^v, Y^c) = -\flat_h(X)(Y) \circ \tau_Q.$$

Notice that the right-hand side of the previous equation is nothing more than the vertical lift of the function $\flat_h(X)(Y)$. Using (2.4.12), we finally get

$$\omega_{L_h}(X^v, Y^c) = -(\flat_h(X))^v(Y^c).$$

At last, we need to check how the symplectic form acts on vertical lifts. However, note that using (A.0.3) and (2.4.13) then the expression

$$\omega_{L_h}(X^v, Y^v) = -X^v\left(\theta_{L_h}(Y^v)\right) + Y^v\left(\theta_{L_h}(X^v)\right) + \theta_{L_h}([X^v, Y^v])$$

vanishes for all X and Y, as it is the case of $(\flat_h(X))^v(Y^v)$ again by (2.4.12), which finishes the proof. □

A.1 The complete lift of a Lagrangian system of kinetic type

In this appendix, we will review some results related with the complete lift of a regular Lagrangian system of kinetic type.

Let Q be a smooth manifold of dimension n, $\tau_Q : TQ \to Q$ the canonical projection and TTQ the double tangent bundle to Q.

Now suppose that g is a Riemannian metric on Q and that $L_g : TQ \to \mathbb{R}$ is the Lagrangian function of kinetic type induced by g (see Appendix B). Then, we may consider the complete lift g^c of g [YI73]. It is not a Riemannian metric on TQ but a pseudo-Riemannian metric of signature (n, n). In fact g^c is characterized by the following conditions

$$\begin{aligned} &g^c(X^c, Y^c) = (g(X, Y))^c, \\ &g^c(X^c, Y^v) = g^c(X^v, Y^c) = (g(X, Y))^v, \\ &g^c(X^v, Y^v) = 0, \end{aligned} \tag{A.1.1}$$

for $X, Y \in \mathfrak{X}(Q)$.

If $(q^i, \dot{q}^i)$ are local coordinates on TQ and the local expression of the Riemannian metric g is $g = g_{ij} dq^i \otimes dq^j$ then the local expression of its complete lift is

$$g^c(q^i, \dot{q}^i) = \dot{q}^k \frac{\partial g_{ij}}{\partial q^k} dq^i \otimes dq^j + g_{ij} dq^i \otimes d\dot{q}^j + g_{ij} d\dot{q}^i \otimes dq^j.$$

Anyway, we may consider the Lagrangian function $L_{g^c} : TTQ \to \mathbb{R}$ on TTQ induced by the pseudo-Riemannian metric g^c on TQ. Then, the relation of the previous construction with the canonical involution is given by the following result:

Lemma A.1.1. *We have that the Lagrangian function* $L_{g^c} : TTQ \to \mathbb{R}$ *is regular and satisfies the following equation*

$$L_{g^c} = L_g^c \circ \kappa_Q. \tag{A.1.2}$$

Proof. The Lagrangian function L_{g^c} is regular since its Hessian matrix is the tensor g^c which is a non-degenerate tensor. In fact, it is a pseudo-Riemannian metric.

If $Z \in T_u(TQ)$, with $u \in T_qQ$, then it is easy to see that there exist vector fields $X, Y \in \mathfrak{X}(Q)$ such that

$$Z = X^c(u) + Y^v(u).$$

So, it is sufficient to prove that

$$L_{g^c}(X^c(u) + Y^v(u)) = L_g^c \circ \kappa_Q(X^c(u) + Y^v(u)).$$

Now, since κ_Q is a vector bundle isomorphism between the vector bundles τ_{TQ} and $T\tau_Q$, it follows that

$$\kappa_Q(X^c(u) + Y^v(u)) = (T_{(u,0(q))}(+))(\kappa_Q(X^c(u)), \kappa_Q(Y^v(u))),$$

and then from (2.4.16) and the definition of complete lift of a function we deduce

$$L_g^c \circ \kappa_Q(X^c(u) + Y^v(u)) = (T_{(u,0(q))}(+))(TX(u), \widetilde{Y}^v(u))(L_g).$$

where $(+) : TQ \times_Q TQ \to TQ$ in the right-hand side of the equality is the addition on the fibers of the vector bundle $\tau_Q : TQ \to Q$. So we have that

$$L_g^c \circ \kappa_Q(X^c(u) + Y^v(u)) = (TX(u), \widetilde{Y}^v(u))(L_g \circ (+)). \tag{A.1.3}$$

Next, let $\sigma : (-\varepsilon, \varepsilon) \to Q$ be a curve on Q such that

$$\sigma(0) = q, \quad \dot{\sigma}(0) = u$$

and $Z : (-\varepsilon, \varepsilon) \to TQ$ a curve over σ satisfying

$$Z(0) = 0(q), \quad \dot{Z}(0) = \widetilde{Y}^v(u) = (T_q 0)(u) + Y^v(0(q)). \tag{A.1.4}$$

Then, from (A.1.3), it follows that

$$L_g^c \circ \kappa_Q(X^c(u) + Y^v(u)) = \left.\frac{d}{dt}\right|_{t=0} (L_g \circ (+))(X(\sigma(t)), Z(t)),$$

hence

$$\begin{aligned} L_g^c \circ \kappa_Q(X^c(u) + Y^v(u)) = \left.\frac{d}{dt}\right|_{t=0} \Big(&\frac{1}{2} g(X(\sigma(t)), X(\sigma(t))) \\ &+ g(X(\sigma(t)), Z(t)) + \frac{1}{2} g(Z(t), Z(t))\Big), \end{aligned}$$

Thus, using (A.1.4) and the following equalities

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} g(X(\sigma(t)), X(\sigma(t))) &= \dot{\sigma}(0)(g(X, X)), \\ \left.\frac{d}{dt}\right|_{t=0} g(X(\sigma(t)), Z(t)) &= (\mathcal{L}_{\dot{\sigma}} g)|_{t=0}(X(q), Z(0)) \\ &\quad + g(\mathcal{L}_{\dot{\sigma}} X(\sigma(t))|_{t=0}, Z(0)) + g(X(q), \mathcal{L}_{\dot{\sigma}} Z(t)|_{t=0}), \\ \mathcal{L}_{\dot{\sigma}} Z(t)|_{t=0} &= Y(q), \end{aligned}$$

we deduce that

$$L_g^c \circ \kappa_Q(X^c(u) + Y^v(u)) = \frac{1}{2}u(g(X,X)) + g(X(q), Y(q)).$$

On the other hand, from (A.1.1), we deduce that

$$\begin{aligned} L_{g^c}(X^c(u) + Y^v(u)) &= \frac{1}{2}g^c(u)(X^c(u) + Y^v(u), X^c(u) + Y^v(u)) \\ &= \frac{1}{2}(g(X,X))^c(u) + g(X(q), Y(q)) \\ &= \frac{1}{2}u(g(X,X)) + g(X(q), Y(q)), \end{aligned}$$

which proves the Lemma. □

In local coordinates, the Lagrangian function associated to g^c has the local expression

$$L_{g^c}(q, \dot{q}, v, \dot{v}) = \frac{1}{2}\dot{q}^k \frac{\partial g_{ij}}{\partial q^k} v^i v^j + g_{ij} v^i \dot{v}^j. \tag{A.1.5}$$

Proposition A.1.2. *Let g be a Riemannian metric on Q, g^c the complete lift of g to TQ, $L_g : TQ \to \mathbb{R}$ and $L_{g^c} : TTQ \to \mathbb{R}$ the corresponding Lagrangian functions. Then*

$$\theta_{L_{g^c}} = \kappa_Q^*(\theta_{L_g}^c), \;\; \omega_{L_{g^c}} = \kappa_Q^*(\omega_{L_g}^c), \;\; \textit{and } E_{L_{g^c}} = (E_{L_g})^c \circ \kappa_Q,$$

where θ_{L_g} (respectively $\theta_{L_{g^c}}$), ω_{L_g} (respectively $\omega_{L_{g^c}}$) and E_{L_g} (respectively $E_{L_{g^c}}$) are the Poincaré-Cartan 1-form, the Poincaré-Cartan 2-form and the Lagrangian energy associated with L_g (respectively with L_{g^c}).

Proof. Notice that once we establish that $\theta_{L_{g^c}} = \kappa_Q^*(\theta_{L_g}^c)$, then the corresponding formula for the Poincaré-Cartan 2-forms follows since the pullback commutes with the differential and the differential of the complete lift behaves according to (2.4.8).

It is sufficient to prove $\theta_{L_{g^c}} = \kappa_Q^*(\theta_{L_g}^c)$ over the vector fields $(X^c)^c$, $(X^c)^v$, $(X^v)^c$ and $(X^v)^v$, with $X \in \mathfrak{X}(Q)$. Now,

$$\theta_{L_{g^c}}((X^c)^c) = dL_{g^c}((X^c)^v) = d(L_g^c \circ \kappa_Q)((X^c)^v) = (\kappa_Q)^* dL_g^c((X^c)^v),$$

where we used the definition of the Poincaré-Cartan 1-form and the properties in (2.4.19) in the first equality, while we used Lemma A.1.1 in the second equality. Then using the definition of pullback together with (2.4.17) we get

$$\theta_{L_{g^c}}((X^c)^c) = (dL_g^c \circ \kappa_Q)((X^v)^c \circ \kappa_Q) = (dL_g^c((X^v)^c)) \circ \kappa_Q = (dL_g(X^v))^c \circ \kappa_Q,$$

where we used (2.4.7) and (2.4.8) in the last equality. Now, using the definition of θ_{L_g} followed by (2.4.7) again we obtain

$$\theta_{L_{g^c}}((X^c)^c) = (\theta_{L_g}(X^c))^c \circ \kappa_Q = \theta_{L_g}^c((X^c)^c) \circ \kappa_Q.$$

Finally, using (2.4.17) we obtain

$$\theta_{L_{g^c}}((X^c)^c) = (\kappa_Q^*(\theta_{L_g}^c))((X^c)^c).$$

By applying the same arguments we may deduce the remaining expressions. Indeed

$$\theta_{L_{g^c}}((X^c)^v) = 0 = (\theta_{L_g}(X^v))^c \circ \kappa_Q = \theta_{L_g}^c((X^v)^c) \circ \kappa_Q = (\kappa_Q^*(\theta_{L_g}^c))((X^c)^v),$$

where we used the definition of the Poincaré-Cartan 1-forms and (2.4.7), (2.4.19) and (2.4.17). Also using these arguments and Lemma A.1.1 we deduce

$$\begin{aligned}\theta_{L_{g^c}}((X^v)^c) &= dL_{g^c}((X^v)^v) = d(L_g^c \circ \kappa_Q)((X^v)^v) = (\kappa_Q)^* dL_g^c((X^v)^v) \\ &= (dL_g^c((X^v)^v)) \circ \kappa_Q = (dL_g(X^v))^v \circ \kappa_Q,\end{aligned}$$

where the last equality follows from (2.4.12). Then

$$\theta_{L_{g^c}}((X^v)^c) = (\theta_{L_g}(X^c))^v \circ \kappa_Q = \theta_{L_g}^c((X^c)^v) \circ \kappa_Q = (\kappa_Q^*(\theta_{L_g}^c))((X^v)^c).$$

The last expression follows directly from (2.4.12) and (2.4.19).

On the other hand,

$$\theta_{L_{g^c}}((X^v)^v) = 0 = (\theta_{L_g}(X^v))^v \circ \kappa_Q = \theta_{L_g}^c((X^v)^v) \circ \kappa_Q = (\kappa_Q^*(\theta_{L_g}^c))((X^v)^v).$$

To prove the expression relating the Lagrangian energies, denote first by $\Delta_{TQ} \in \mathfrak{X}(TQ)$ the Liouville vector field on the tangent bundle and by $\Delta_{TTQ} \in \mathfrak{X}(TTQ)$ the Liouville vector field on the double tangent bundle. Recall the definition of Lagrangian energy

$$E_{L_{g^c}} = \Delta_{TTQ}(L_{g^c}) - L_{g^c} \text{ and } E_{L_g} = \Delta_{TQ}(L_g) - L_g.$$

Moreover, note that since the Lagrangian functions are of kinetic type, then

$$\Delta_{TTQ}(L_{g^c}) = 2L_g^c \text{ and } \Delta_{TQ}(L_g) = 2L_g.$$

Then, using Lemma A.1.1 we have that

$$E_{L_{g^c}} = L_{g^c} = L_g^c \circ \kappa_Q = (E_{L_g})^c \circ \kappa_Q.$$

□

Finally, we will describe the dynamics associated with the Lagrangian function $L_{g^c} : TTQ \to \mathbb{R}$ in terms of the complete lift of the geodesic flow associated with g. In addition, we will prove that the trajectories of the Lagrangian system (TTQ, L_{g^c}) are just the Jacobi fields of the Riemannian metric g.

Proposition A.1.3. *Let g be a Riemannian metric on Q, g^c the complete lift of g to TQ, $L_g : TQ \to \mathbb{R}$ and $L_{g^c} : TTQ \to \mathbb{R}$ the corresponding Lagrangian functions, Γ_{L_g} and $\Gamma_{L_{g^c}}$ the corresponding Lagrangian vector fields on TQ and TTQ, respectively. Then*

1. $\Gamma_{L_{g^c}} = T\kappa_Q \circ \Gamma^c_{L_g} \circ \kappa_Q$;
2. *If $Z : I \to TQ$ is a trajectory of the SODE $\Gamma_{L_{g^c}}$, then Z is a Jacobi field for g over a geodesic $c_v : I \to Q$ of g.*

Proof. Let $w \in TTQ$ and $X \in T_w(TTQ)$. Then we have that

$$\left(i_{(T_w\kappa_Q)(\Gamma^c_{L_g}(w))}\omega_{L_{g^c}}(\kappa_Q(w))\right)(T_w\kappa_Q(X)) = \left(\kappa_Q^*\omega_{L_{g^c}}(w)\right)(\Gamma^c_{L_g}(w), X).$$

Using Proposition A.1.2, the last expression reduces to

$$\left(\omega^c_{L_g}(w)\right)(\Gamma^c_{L_g}(w), X) = \left(i_{\Gamma^c_{L_g}(w)}\omega^c_{L_g}(w)\right)(X) = \left(i_{\Gamma_{L_g}}\omega_{L_g}\right)^c(w)(X),$$

where we used (2.4.9) in the last equality. Using now the geometric equation of motion the last line becomes

$$\left(dE_{L_g}\right)^c(w)(X) = \left(dE^c_{L_g}\right)(w)(X) = \left(d(E_{L_{g^c}} \circ \kappa_Q)\right)(w)(X),$$

where we used (2.4.8) and Proposition A.1.2. Using now the definition of pullback and the geometric equations of motion we get

$$\left(dE_{L_{g^c}}(\kappa_Q(w))\right)(T_w\kappa_Q(X)) = \left(i_{\Gamma_{L_{g^c}}\kappa_Q(w)}\omega_{L_{g^c}}(\kappa_Q(w))\right)(T_w\kappa_Q(X)).$$

Therefore, since $\omega_{L_{g^c}}$ is non-degenerate we deduce

$$\Gamma_{L_{g^c}}(\kappa_Q(w)) = (T_w\kappa_Q)(\Gamma^c_{L_g})(w),$$

from where the first item follows.

Next we prove the second item. By Proposition (A.1.2), if $Z : I \to TQ$ is a curve on the tangent bundle such that $\kappa_Q \circ \dot{Z} : I \to TTQ$ is an integral curve of $\Gamma^c_{L_g}$, then

$$\kappa_Q \circ \dot{Z}(t) = \left(T_{Z(0)}\phi_t^{\Gamma_{L_g}}\right)(\kappa_Q \circ \dot{Z}(0)),$$

where $\phi_t^{\Gamma_{L_g}}$ is the flow of the vector field Γ_{L_g}. Projecting the previous equation using $T\tau_Q$ we get that

$$\tau_{TQ} \circ \dot{Z}(t) = \left(T_{Z(0)}(\tau_Q \circ \phi_t^{\Gamma_{L_g}})\right)(\kappa_Q \circ \dot{Z}(0)),$$

One one hand, note that $\tau_{TQ} \circ \dot{Z}(t)$ is just the curve $Z(t)$, by definition of tangent vector field to a curve. On the other hand, let $V : I \to TQ$ be a curve with initial velocity such that $V'(0) = \kappa_Q \circ \dot{Z}(0)$. Then, the expression above may be rewritten as

$$Z(t) = \left.\frac{d}{ds}\right|_{s=0} \left((\tau_Q \circ \phi_t^{\Gamma_{L_g}})(V(s))\right).$$

Therefore, Z is an infinitesimal variation vector field for a family of trajectories of Γ_{L_g}. Hence, it is a Jacobi field for the Riemannian metric g. □

Appendix B

Auxiliary analytical results in ODE

Let

$$\frac{d^2q^i}{dt^2} = \xi^i(t, q^j, \frac{dq^j}{dt}), \ \ \forall i \in \{1, \dots, n\}$$

be a system of second order differential equations on $\mathbb{R}^n$, with ξ^i a real C^∞-function on a compact subset of $\mathbb{R} \times \mathbb{R}^{2n}$ which contains the origin.

We will consider the problem of the existence of solutions satisfying the boundary conditions

$$q^i(0) = 0, \quad q^i(h) = 0, \ \ \forall i, \ \text{with} \ h > 0.$$

In this direction, if we take $x_0 = 0$ in Corollary 4.1 of Chapter XII in [Har02], we deduce the following result.

Theorem B.0.1. *Let $\xi^i(t, q, \dot{q})$ be continuous for $1 \leq i \leq n$, $0 \leq t \leq h$, $\|q\| \leq r$, $\|\dot{q}\| \leq r'$ such that ξ satisfies a Lipschitz condition with respect to $q, \dot{q}$ of the form*

$$\|\xi(t, q_1^j, \dot{q}_1^j) - \xi(t, q_2^j, \dot{q}_2^j)\| \leq K\|q_2 - q_1\| + K'\|\dot{q}_2 - \dot{q}_1\|$$

with Lipschitz constants K, K', so small that

$$\frac{Kh^2}{8} + \frac{K'h}{2} < 1 \quad \textit{and } h > 0.$$

In addition, suppose that $\|\xi(t, q^j, \dot{q}^j)\| \leq M$ and that

$$\frac{Mh^2}{8} \leq r, \quad \frac{Mh}{2} \leq r'.$$

Then, the system of second order differential equations

$$\frac{d^2q^j}{dt^2} = \xi^j(t, q^i, \dot{q}^i), \quad \textit{for all } j$$

has a unique solution satisfying

$$\|q(t)\| \leq r, \;\; \|\dot{q}(t)\| \leq r', \;\; q^i(0) = 0, \;\; q^i(h) = 0,$$

for all $t \in [0, h]$ *and* $1 \leq i \leq n$*.*

If we have a second order linear system of differential equations of the type

$$\ddot{x} = A(t)\dot{x} + B(t)x \tag{B.0.1}$$

and satisfying the boundary conditions

$$x(0) = 0, \quad x(h) = 0, \tag{B.0.2}$$

Lemma 3.1. in Chapter XII, also in [Har02], provides a necessary and sufficient condition for the existence of non-trivial solutions of the previous problem in terms of the corresponding matrix solution. Then, the lemma states the following:

Lemma B.0.2. *Let* $A(t)$ *and* $B(t)$ *be continuous* $d \times d$ *matrices on* $t \in [0, h]$*. If* $U(t)$ *is the matrix solution of the initial value problem*

$$\ddot{U} = A(t)\dot{U} + B(t)U, \quad U(0) = 0, \quad \dot{U}(0) = Id,$$

then (B.0.1) *has a non-trivial solution satisfying* (B.0.2) *if and only if* $U(h)$ *is singular.*

At last, we need the following classical result:

Proposition B.0.3. *Let* $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ *be a* C^∞*-smooth map, with* U *a convex open subset of* $\mathbb{R}^n$ *and suppose that there exists a positive constant* $C > 0$ *such that*

$$\|df(z)\| \leq C, \;\; \forall z \in U.$$

Then, we have that

$$\|f(x) - f(y)\| \leq C\|x - y\|, \quad \textit{for } x, y \in U.$$

Proof. Suppose that $x, y \in U$ and denote by $f_1, \ldots, f_n$ the components of f and by

$$g_i : [0,1] \to \mathbb{R}$$

the smooth real function on the interval $[0,1]$ given by

$$g_i(t) = f_i(x + t(y - x)).$$

Then, we have that

$$f_i(y) - f_i(x) = g_i(1) - g_i(0) = \int_0^1 g_i'(t)dt = \int_0^1 \left(\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x + t(y-x))(y_j - x_j) \right) dt.$$

So, we deduce that

$$f(y) - f(x) = \int_0^1 df(x + t(y - x))(y - x)dt$$

which implies that

$$\|f(y) - f(x)\| \leq \int_0^1 \|df(x + t(y - x))\| \|y - x\| dt = \|y - x\| \int_0^1 \|df(x + t(y - x))\| dt.$$

Thus, using that $\|df(z)\| \leq C$ for every $z \in U$, we conclude that

$$\|f(y) - f(x)\| \leq C\|y - x\|.$$

□

Bibliography

[ABR18] A. Agrachev, D. Barilari, and L. Rizzi. "Curvature: a variational approach". In: *Mem. Amer. Math. Soc.* 256.1225 (2018), pp. v+142. ISSN: 0065-9266. DOI: `10.1090/memo/1225`. URL: `https://doi.org/10.1090/memo/1225`.

[AM78] R. Abraham and J.E. Marsden. *Foundations of Mechanics*. Benjamin/Cummings Publishing Co.Inc. Advanced Book Program, Reading, Mass, 1978.

[AMM20a] Alexandre Anahory Simoes, Juan Carlos Marrero, and David Martín de Diego. "Exact discrete Lagrangian mechanics for nonholonomic mechanics". In: *arXiv e-prints*, arXiv:2003.11362 (Mar. 2020), arXiv:2003.11362. arXiv: `2003.11362 [math-ph]`.

[AMM20b] Alexandre Anahory Simoes, Juan Carlos Marrero, and David Martín de Diego. "Jacobi fields in nonholonomic mechanics". In: *arXiv e-prints*, arXiv:2004.10457 (Apr. 2020), arXiv:2004.10457. arXiv: `2004.10457 [math.DG]`.

[AMM20c] Alexandre Anahory Simoes, Juan Carlos Marrero, and David Martín de Diego. "Radial kinetic nonholonomic trajectories are Riemannian geodesics!" In: *arXiv e-prints*, arXiv:2010.12444 (Oct. 2020), arXiv:2010.12444. arXiv: `2010.12444 [math-ph]`.

[AMM21] Alexandre Anahory Simoes, Juan Carlos Marrero, and David Martín de Diego. *Contact bundle formulation of nonholonomic Maupertuis-Jacobi principle and a length minimizing property of nonholonomic dynamics*. 2021. arXiv: `2104.13178 [math-ph]`.

[AMS08] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. With a foreword by Paul Van Dooren. Princeton University Press, Princeton, NJ, 2008, pp. xvi+224.

ISBN: 978-0-691-13298-3. DOI: 10.1515/9781400830244. URL: https://doi.org/10.1515/9781400830244.

[Ana+21] A. Anahory Simoes et al. "On the Geometry of Discrete Contact Mechanics". In: *Journal of Nonlinear Science* 31 (2021). DOI: 10.1007/s00332-021-09708-2.

[Bal14] Paula Balseiro. "The Jacobiator of nonholonomic systems and the geometry of reduced nonholonomic brackets". In: *Arch. Ration. Mech. Anal.* 214.2 (2014), pp. 453–501. ISSN: 0003-9527. DOI: 10.1007/s00205-014-0759-4. URL: https://doi.org/10.1007/s00205-014-0759-4.

[Bal17] Paula Balseiro. "Hamiltonization of solids of revolution through reduction". In: *J. Nonlinear Sci.* 27.6 (2017), pp. 2001–2035. ISSN: 0938-8974. DOI: 10.1007/s00332-017-9394-1. URL: https://doi.org/10.1007/s00332-017-9394-1.

[BBM10] A. Bolsinov, A. Borisov, and Ivan Mamaev. "Hamiltonization of Non-Holonomic Systems in the Neighborhood of Invariant Manifolds". In: *Regular and Chaotic Dynamics* 16 (2010), pp. 443–464. DOI: 10.1134/S1560354711050030.

[BBM15] A V Bolsinov, A V Borisov, and I S Mamaev. "Geometrisation of Chaplygins reducing multiplier theorem". In: *Nonlinearity* 28.7 (2015), pp. 2307–2318. DOI: 10.1088/0951-7715/28/7/2307. URL: https://doi.org/10.1088/0951-7715/28/7/2307.

[BC16] Sergio Blanes and Fernando Casas. *A concise introduction to geometric numerical integration*. Monographs and Research Notes in Mathematics. CRC Press, Boca Raton, FL, 2016, pp. xiv+218. ISBN: 978-1-4822-6342-8.

[BC95] A. M. Bloch and P. E. Crouch. "Nonholonomic Control Systems on Riemannian Manifolds". In: *SIAM Journal on Control and Optimization* 33.1 (1995), pp. 126–148. DOI: 10.1137/S036301299223533X. eprint: https://doi.org/10.1137/S036301299223533X. URL: https://doi.org/10.1137/S036301299223533X.

[BC98] A. M. Bloch and P. E. Crouch. "Newton's Law and Integrability of Nonholonomic Systems". In: *SIAM Journal on Control and Optimization* 36.6 (1998), pp. 2020–2039. DOI: 10.1137/S0363012995291634. eprint: https://doi.org/10.

1137/S0363012995291634. URL: https://doi.org/10.1137/S0363012995291634.

[BF15] Paula Balseiro and Oscar E. Fernandez. "Reduction of nonholonomic systems in two stages and Hamiltonization". In: *Nonlinearity* 28.8 (2015), pp. 2873–2912. ISSN: 0951-7715. DOI: 10.1088/0951-7715/28/8/2873. URL: https://doi.org/10.1088/0951-7715/28/8/2873.

[BFM09] A.M. Bloch, O.E. Fernandez, and T. Mestdag. "Hamiltonization of nonholonomic systems and the inverse problem of the calculus of variations". In: *Reports on Mathematical Physics* 63.2 (2009), pp. 225 –249. ISSN: 0034-4877. DOI: https://doi.org/10.1016/S0034-4877(09)90001-5.

[BGN12] Paula Balseiro and Luis C. García-Naranjo. "Gauge transformations, twisted Poisson brackets and Hamiltonization of nonholonomic systems". In: *Arch. Ration. Mech. Anal.* 205.1 (2012), pp. 267–310. ISSN: 0003-9527. DOI: 10.1007/s00205-012-0512-9. URL: https://doi.org/10.1007/s00205-012-0512-9.

[BLn+19] María Barbero Liñán et al. "Morse families and Dirac systems". In: *J. Geom. Mech.* 11.4 (2019), pp. 487–510. ISSN: 1941-4889. DOI: 10.3934/jgm.2019024. URL: https://doi.org/10.3934/jgm.2019024.

[Blo11] Anthony M. Bloch. "Nonholonomic Mechanics, Dissipation and Quantization". In: *Advances in the Theory of Control, Signals and Systems with Physical Modeling*. Ed. by Jean Lévine and Philippe Müllhaupt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 141–152. ISBN: 978-3-642-16135-3. DOI: 10.1007/978-3-642-16135-3_12. URL: https://doi.org/10.1007/978-3-642-16135-3_12.

[Blo15] A. Bloch. *Nonholonomic Mechanics and Control*. Springer, Interdisciplinary Applied Mathematics 24, 2015.

[Blo+96a] A. Bloch et al. "Nonholonomic mechanical systems with symmetry". In: *Arch. Rational Mech. Anal.* 136 (1996), pp. 21–99.

[Blo+96b] Anthony M. Bloch et al. “Nonholonomic Mechanical Systems with Symmetry”. In: *Archive for Rational Mechanics and Analysis* 136.1 (1996), pp. 21–99. DOI: 10.1007/BF02199365.

[BM01] A.V. Borisov and Ivan Mamaev. “Chaplygin’s Ball Rolling Problem Is Hamiltonian”. In: *Mathematical Notes* 70 (Nov. 2001), pp. 720–723. DOI: 10.1023/A:1012995330780.

[BM03] A. Borisov and Ivan Mamaev. “The Rolling Body Motion Of a Rigid Body on a Plane and a Sphere. Hierarchy of Dynamics”. In: *Regular and Chaotic Dynamics* 7 (July 2003). DOI: 10.1070/RD2002v007n02ABEH000204.

[BM05] A. Borisov and Ivan Mamaev. “On the History of the Development of the Nonholonomic Dynamics”. In: *Regular and Chaotic Dynamics* 7 (2005). DOI: 10.1070/RD2002v007n01ABEH000194.

[BM07] A. V. Borisov and I. S. Mamaev. “Isomorphism and Hamilton representation of some nonholonomic systems”. In: *Sib Math J* 48 (2007), 26–36. DOI: https://doi.org/10.1007/s11202-007-0004-6.

[BMZ05] Anthony Bloch, Jerrold Marsden, and Dmitry Zenkov. “Nonholonomic Dynamics”. In: *Notices of the AMS* 52 (2005).

[BMZ09] Anthony M. Bloch, Jerrold E. Marsden, and Dmitry V. Zenkov. “Quasivelocities and symmetries in non-holonomic systems”. In: *Dynamical Systems* 24.2 (2009), pp. 187–222. DOI: 10.1080/14689360802609344. eprint: https://doi.org/10.1080/14689360802609344. URL: https://doi.org/10.1080/14689360802609344.

[BR17] D. Barilari and L. Rizzi. “On Jacobi fields and a canonical connection in sub-Riemannian geometry”. In: *Arch. Math. (Brno)* 53.2 (2017), pp. 77–92. ISSN: 0044-8753. DOI: 10.5817/AM2017-2-77. URL: https://doi.org/10.5817/AM2017-2-77.

[Bra17] Alessandro Bravetti. “Contact Hamiltonian Dynamics: The Concept and Its Use”. en. In: *Entropy* 19.12 (Oct. 2017), p. 535. ISSN: 1099-4300. DOI: 10.3390/e19100535.

[Bra18] Alessandro Bravetti. “Contact Geometry and Thermodynamics”. In: *Int. J. Geom. Methods Mod. Phys.* 16.supp01 (Oct. 2018), p. 1940003. ISSN: 0219-8878. DOI: 10.1142/S0219887819400036.

[Bra+20] Alessandro Bravetti et al. "Numerical integration in Celestial Mechanics: a case for contact geometry". In: *Celestial Mech. Dynam. Astronom.* 132.1 (2020), Paper No. 7. ISSN: 0923-2958. DOI: 10.1007/s10569-019-9946-9. URL: https://doi.org/10.1007/s10569-019-9946-9.

[BS08] P. Balseiro and J.E. Solomin. "On Generalized Non-holonomic Systems". In: *Lett Math Phys* 84 (2008), 15–30. DOI: 10.1007/s11005-008-0236-9. URL: https://doi.org/10.1007/s11005-008-0236-9.

[BS15] Paula Balseiro and Nicola Sansonetto. "A Geometric Characterization of Certain First Integrals for Nonholonomic Systems with Symmetries". In: *Symmetry, Integrability and Geometry: Methods and Applications* 12 (2015). DOI: 10.3842/SIGMA.2016.018.

[BS93] Larry Bates and Jedrzej Śniatycki. "Nonholonomic reduction". In: *Rep. Math. Phys.* 32.1 (1993), pp. 99–115. ISSN: 0034-4877. DOI: 10.1016/0034-4877(93)90073-N. URL: https://doi.org/10.1016/0034-4877(93)90073-N.

[BY20] Paula Balseiro and Luis P Yapu. "Conserved quantities and Hamiltonization of nonholonomic systems". In: *Annales de L' Institut Henri Poincare C, Analyse non lineaire* 33.3 (2020), pp. 1297–1341. ISSN: 0305-4470. URL: https://doi.org/10.1016/j.anihpc.2020.05.003.

[BZ15] K. R. Ball and D. V. Zenkov. "Hamel's formalism and variational integrators". In: *Geometry, mechanics, and dynamics.* Vol. 73. Fields Inst. Commun. Springer, New York, 2015, pp. 477–506. DOI: 10.1007/978-1-4939-2441-720.

[Can+98] F. Cantrijn et al. "Reduction of nonholonomic mechanical systems with symmetries". In: vol. 42. 1-2. Pacific Institute of Mathematical Sciences Workshop on Nonholonomic Constraints in Dynamics (Calgary, AB, 1997). 1998, pp. 25–45. DOI: 10.1016/S0034-4877(98)80003-7. URL: https://doi.org/10.1016/S0034-4877(98)80003-7.

[Car92] Manfredo do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Translated from the second Portuguese edition by Francis Flaherty. Birkhäuser Boston Inc. Boston, MA, 1992, pp. xiv+300. ISBN: 0-8176-3490-8. URL: https://doi.org/10.1007/978-1-4757-2201-7.

[CCM18] Florio Maria Ciaglia, Hans Cruz, and Giuseppe Marmo. "Contact manifolds and dissipation, classical and quantum". In: *Annals of Physics* 398 (2018). DOI: 10.1016/j.aop.2018.09.012.

[CCS87] M. Crampin, F. Cantrijn, and W. Sarlet. "Lifting geometric objects to a cotangent bundle, and the geometry of the cotangent bundle of a tangent bundle". In: *Journal of Geometry and Physics* 4.4 (1987), pp. 469–492. ISSN: 0393-0440. DOI: https://doi.org/10.1016/0393-0440(87)90003-9. URL: https://www.sciencedirect.com/science/article/pii/0393044087900039.

[Cel+19] Elena Celledoni et al. "Energy-preserving integrators applied to nonholonomic systems". In: *J. Nonlinear Sci.* 29.4 (2019), pp. 1523–1562. ISSN: 0938-8974. DOI: 10.1007/s00332-018-9524-4. URL: https://doi.org/10.1007/s00332-018-9524-4.

[CGM15] J. Cariñena, I. Gheorghiu, and E. Martínez. "Jacobi field equations for second-order differential equations on Lie algebroids". In: Dynamical Systems, Differential Equations and Applications. AIMS Proceddings, 2015, pp. 213–222.

[Cha54] S. A. Chaplygin. *Selected Works on Mechanics and Mathematics*. Techincal-Theoretical Literature, Moscow. State Publ. House, 1954.

[CM01] J. Cortés and S. Martínez. "Non-holonomic integrators". In: *Nonlinearity* 14.5 (2001), pp. 1365–1392. ISSN: 0951-7715. DOI: 10.1088/0951-7715/14/5/322.

[CM92] J. Cariñena and E. Martínez. "Generalized Jacobi equation and inverse problem in classical mechanics". In: Integral Systems, Solid State Physics, Theory of Phase Transitions. Part 2: Symmetries, and Algebraic Structures in Physics. Nova Science., 1992.

[CMR01] H. Cendra, J. Marsden, and T. Ratiu. "Geometric Mechanics, Lagrangian Reduction, and Nonholonomic Systems". In: *Mathematics Unlimited-2001 and Beyond. Springer-Verlag* (2001), pp. 221–273.

[Cor02] J. Cortés. *Geometric, control and numerical aspects of nonholonomic systems*. Vol. 1793. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2002, pp. xvi+219. ISBN: 3-540-44154-9. DOI: 10.1007/b84020. URL: https://doi.org/10.1007/b84020.

[Cor+03] J. Cortés et al. "Geometric description of vakonomic and nonholonomic dynamics. Comparison of solutions". In: *SIAM J. Control Optim.* 41.5 (2003), pp. 1389–1412.

[Cor+09a] Jorge Cortés et al. "Nonholonomic Lagrangian systems on Lie algebroids". In: *Discrete Contin. Dyn. Syst.* 24.2 (2009), pp. 213–271. ISSN: 1078-0947. DOI: 10.3934/dcds.2009.24.213. URL: https://doi.org/10.3934/dcds.2009.24.213.

[Cor+09b] J. Cortés et al. "Nonholonomic lagrangian systems on Lie algebroids". In: *Discrete and continuous dynamical systems* 24.2 (2009), pp. 213–271.

[Cra83] M Crampin. "Tangent bundle geometry Lagrangian dynamics". In: *Journal of Physics A: Mathematical and General* 16.16 (1983), pp. 3755–3772. DOI: 10.1088/0305-4470/16/16/014. URL: https://doi.org/10.1088/0305-4470/16/16/014.

[CS98] R. Cushman and J. Sniatycki, eds. *Pacific Institute of Mathematical Sciences Workshop on Nonholonomic Constraints in Dynamics*. Rep. Math. Phys. **42** (1998), no. 1-2. Pergamon Press, Oxford, 1998, i–ii and 1–271.

[DA18] D. Martín de Diego and R. Sato Martín de Almagro. "Variational order for forced Lagrangian systems". In: *Nonlinearity* 31.8 (2018), pp. 3814–3846. ISSN: 0951-7715. DOI: 10.1088/1361-6544/aac5a6.

[Ehl+05] Kurt Ehlers et al. *Nonholonomic systems via moving frames: Cartan equivalence and Chaplygin Hamiltonization*. Vol. 232. Progr. Math. Birkhäuser Boston, MA, 2005, pp. 75–120. DOI:

10.1007/0-8176-4419-9-4. URL: https://doi.org/10.1007/0-8176-4419-9-4.

[FB08] Oscar E Fernandez and Anthony M Bloch. "Equivalence of the dynamics of nonholonomic and variational nonholonomic systems for certain initial data". In: *Journal of Physics A: Mathematical and Theoretical* 41.34 (2008), p. 344005. DOI: 10.1088/1751-8113/41/34/344005. URL: https://doi.org/10.1088/1751-8113/41/34/344005.

[FBO12] O. E. Fernandez, A. M. Bloch, and P. J. Olver. "Variational integrators from Hamiltonizable nonholonomic systems". In: *J. Geom. Mech.* 4.2 (2012), pp. 137–163. ISSN: 1941-4889. DOI: 10.3934/jgm.2012.4.137.

[FBZ14] Oscar E. Fernandez, Anthony M. Bloch, and Dmitry V. Zenkov. "The geometry and integrability of the Suslov problem". In: *Journal of Mathematical Physics* 55.11 (2014), p. 112704. DOI: 10.1063/1.4901754. eprint: https://doi.org/10.1063/1.4901754. URL: https://doi.org/10.1063/1.4901754.

[Fer+17] Sebastián Ferraro et al. "On the geometry of the Hamilton-Jacobi equation and generating functions". In: *Arch. Ration. Mech. Anal.* 226.1 (2017), pp. 243–302. ISSN: 0003-9527. DOI: 10.1007/s00205-017-1133-0. URL: https://doi.org/10.1007/s00205-017-1133-0.

[FGN10] Yuri N Fedorov and Luis C García-Naranjo. "The hydrodynamic Chaplygin sleigh". In: *Journal of Physics A: Mathematical and Theoretical* 43.43 (2010), p. 434013. DOI: 10.1088/1751-8113/43/43/434013. URL: https://doi.org/10.1088/1751-8113/43/43/434013.

[FGNM15] Y. N. Fedorov, L. C. García-Naranjo, and J. C. Marrero. "Unimodularity and Preservation of Volumes in Nonholonomic Mechanics". In: *J Nonlinear Sci* 25 (2015), 203–246. DOI: https://doi.org/10.1007/s00332-014-9227-4.

[FGNM19] Francesco Fassò, Luis García-Naranjo, and James Montaldi. "Integrability and Dynamics of the n-Dimensional Symmetric Veselova Top". In: *Journal of Nonlinear Science* 29 (2019). DOI: 10.1007/s00332-018-9515-5.

[FGNS18] Francesco Fassò, Luis C. García-Naranjo, and Nicola Sansonetto. "Moving energies as first integrals of nonholonomic systems with affine constraints". In: *Nonlinearity* 31.3 (2018), pp. 755–782. ISSN: 0951-7715. DOI: 10.1088/1361-6544/aa9837. URL: https://doi.org/10.1088/1361-6544/aa9837.

[FID08] S. Ferraro, D. Iglesias, and D. Martín de Diego. "Momentum and energy preserving integrators for nonholonomic dynamics". In: *Nonlinearity* 21.8 (2008), pp. 1911–1928.

[FJ04] Y. N. Fedorov and B. Jovanović. "Nonholonomic LR systems as generalized Chaplygin systems with an invariant measure and flows on homogeneous spaces". In: *J. Nonlinear Sci.* 14.4 (2004), pp. 341–381. ISSN: 0938-8974. DOI: 10.1007/s00332-004-0603-3. URL: https://doi.org/10.1007/s00332-004-0603-3.

[FMB09] Oscar Fernandez, Tom Mestdag, and Anthony Bloch. "A Generalization of Chaplygin's Reducibility Theorem". In: *Regular and Chaotic Dynamics* 14 (2009), pp. 635–655. DOI: 10.1134/S1560354709060033.

[FS16] Francesco Fassò and Nicola Sansonetto. "Conservation of 'moving' energy in nonholonomic systems with affine constraints and integrability of spheres on rotating surfaces". In: *J. Nonlinear Sci.* 26.2 (2016), pp. 519–544. ISSN: 0938-8974. DOI: 10.1007/s00332-015-9283-4. URL: https://doi.org/10.1007/s00332-015-9283-4.

[FZ05] Yuri N. Fedorov and D. Zenkov. "Discrete Nonholonomic LL Systems on Lie Groups". In: *Nonlinearity* 18 (2005), pp. 2211–2241.

[FZG21] Javier Fernández, Sebastián Graiff Zurita, and Sergio Grillo. *Error analysis of forced discrete mechanical systems.* 2021. arXiv: 2103.11060 [math.NA].

[Gas+20] Jordi Gaset et al. "New Contributions to the Hamiltonian and Lagrangian Contact Formalisms for Dissipative Mechanical Systems and Their Symmetries". In: *Int. J. Geom. Methods Mod. Phys.* 17.6 (2020), p. 2050090.

[GG08] K. Grabowska and J. Grabowski. “Variational calculus with constraints on general algebroids”. In: *J. Phys. A: Mathematical and Theoretical* 41.17 (2008), pp. 175–204.

[Ghe+20] R. Ghezzi et al. “Jacobi fields in sub-riemannian geometry”. In: (2020). URL: `workinprogress`.

[GN10] Luis García-Naranjo. “Reduction of almost Poisson brackets and Hamiltonization of the Chaplygin sphere”. In: *Discrete and Continuous Dynamical Systems S* 3.1 (2010), pp. 37–60.

[GN19] Luis C García-Naranjo. “Generalisation of Chaplygin’s reducing multiplier theorem with an application to multi-dimensional nonholonomic dynamics”. In: *Journal of Physics A: Mathematical and Theoretical* 52.20 (2019), p. 205203. DOI: `10.1088/1751-8121/ab15f8`. URL: `https://doi.org/10.1088/1751-8121/ab15f8`.

[GNJ17] Luis C. García-Naranjo and Fernando Jiménez. “The geometric discretisation of the Suslov problem: A case study of consistency for nonholonomic integrators”. In: *Discrete & Continuous Dynamical Systems* 37.8 (2017), pp. 4249–4275.

[GNM18] Luis García-Naranjo and James Montaldi. “Gauge momenta as Casimir functions of nonholonomic systems”. In: *Arch. Ration. Mech. Anal.* 228.2 (2018), pp. 563–602. ISSN: 0003-9527. DOI: `10.1007/s00205-017-1200-6`. URL: `https://doi.org/10.1007/s00205-017-1200-6`.

[GNM20] Luís Garcia-Naranjo and Juan Carlos Marrero. “The geometry of the non-holonomic Chaplygin systems revisited”. In: *Nonlinearity* 33.3 (2020), pp. 1297–1341. ISSN: 0305-4470. URL: `https://doi.org/10.1088/1361-6544/ab5c0a`.

[GNV21] Luis C. García-Naranjo and Mats Vermeeren. “Structure preserving discretization of time-reparametrized Hamiltonian systems with application to nonholonomic mechanics”. In: *Journal of Computational Dynamics* 0 (2021), pp. –.

[God69] Claude Godbillon. *Géométrie différentielle et mécanique analytique*. French. OCLC: 1038025757. Paris: Hermann, 1969. ISBN: 978-2-7056-5658-4.

[Gra+09] J. Grabowski et al. “Nonholonomic constraints: a new viewpoint”. In: *J. Math. Phys.* 50.1 (2009), pp. 013520, 17. ISSN: 0022-2488. DOI: 10.1063/1.3049752. URL: https://doi.org/10.1063/1.3049752.

[Har02] P. Hartman. *Ordinary differential equations*. Vol. 38. Classics in Applied Mathematics. Corrected reprint of the second (1982) edition [Birkhäuser Boston, MA; MR0658490 (83e:34002)], With a foreword by Peter Bates. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002, pp. xx+612. ISBN: 0-89871-510-5. DOI: 10.1137/1.9780898719222.

[Her30] Gustav Herglotz. “Beruhrungstransformationen”. In: *Lectures at the University of Gottingen*. Gottingen, 1930.

[Her56] Heinrich Hertz. *The principles of mechanics*. Preface by H. von Helmholtz, Translation by D. E. Jones and J. T. Walley, Introduction by Robert S. Cohen. Dover Publications, Inc., New York, 1956, pp. xlii+274.

[HLW10] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*. Vol. 31. Springer Series in Computational Mathematics. Structure-preserving algorithms for ordinary differential equations, Reprint of the second (2006) edition. Springer, Heidelberg, 2010, pp. xviii+644. ISBN: 978-3-642-05157-9.

[HM20] S. Hajdú and T. Mestdag. “Conjugate points for systems of second-order ordinary differential equations”. In: *International Journal of Geometric Methods in Modern Physics* 17.1, 2050012-2139 (Jan. 2020), pp. 2050012–2139. DOI: 10.1142/S0219887820500127. arXiv: 1909.05012 [math.DG].

[Hol+09] D.D. Holm et al. *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*. Oxford Texts in Applied and En. OUP Oxford, 2009. ISBN: 9780199212903. URL: https://books.google.pt/books?id=e5MSDAAAQBAJ.

[Igl+08] D. Iglesias et al. “Discrete nonholonomic Lagrangian systems on Lie Groupoids”. In: *J. Nonlinear Sci.* 18 (2008), pp. 221–276.

[IP+13] David Iglesias-Ponte et al. “Discrete dynamics in implicit form”. In: *Discrete Contin. Dyn. Syst.* 33.3 (2013), pp. 1117–1135. ISSN: 1078-0947. DOI: 10.3934/dcds.2013.33.1117.

[JÍ3] M. Jóźwikowski. "Jacobi vector fields for Lagrangian systems on algebroids". In: *Int. J. Geom. Methods Mod. Phys.* 10.5 (2013), pp. 1350011, 35. ISSN: 0219-8878. DOI: 10.1142/S0219887813500114. URL: https://doi.org/10.1142/S0219887813500114.

[JM70] Souriau Jean-Marie. *Structure des systèmes dynamiques : maîtrises de mathématiques / J.-M. Souriau,...* fre. Dunod université. Paris: Dunod, 1970.

[Jov10] Božidar Jovanović. "Hamiltonization and integrability of the Chaplygin sphere in $\mathbb{R}^n$". In: *J. Nonlinear Sci.* 20.5 (2010), pp. 569–593. ISSN: 0938-8974. DOI: 10.1007/s00332-010-9067-9. URL: https://doi.org/10.1007/s00332-010-9067-9.

[KN63] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry. Vol I.* Interscience Publishers, a division of John Wiley & Sons, New York-London, 1963, pp. xi+329.

[Koi92] J. Koiller. "Reduction of some classical nonholonomic systems with symmetry". In: *Arch. Rational Mech. Anal.* 118.2 (1992), pp. 113–148. ISSN: 0003-9527. DOI: 10.1007/BF00375092. URL: https://doi.org/10.1007/BF00375092.

[Koz02] V. V. Kozlov. "On the integration theory of equations of nonholonomic mechanics". In: *Regul. Chaotic Dyn.* 7.2 (2002), pp. 161–176. ISSN: 1560-3547. DOI: 10.1070/RD2002v007n02ABEH000203. URL: https://doi.org/10.1070/RD2002v007n02ABEH000203.

[LD96] M. de León and D. Martín de Diego. "On the geometry of nonholonomic Lagrangian systems". In: *Journal of Mathematical Physics* 37 (1996), pp. 3389–3414. DOI: 10.1063/1.531571.

[LDSM04] Manuel de León, David Martín de Diego, and A. Santamaría-Merino. "Geometric numerical integration of nonholonomic systems and optimal control problems". In: *Eur. J. Control* 10.5 (2004), pp. 515–521. ISSN: 0947-3580. DOI: 10.3166/ejc.10.515-521.

[Lee13] J.M. Lee. *Introduction to smooth manifolds.* Springer, Graduate Texts in Mathematics 218, 2013.

[Lee97] J.M. Lee. *Riemannian manifolds, an introduction to curvature.* Springer, Graduate Texts in Mathematics 176, 1997.

[Lew97] Andrew D. Lewis. "Energy-preserving affine connections". In: (1997). URL: https://mast.queensu.ca/~andrew/notes/pdf/1997b.pdf.

[Lew98] A.D. Lewis. "Affine connections and distributions with applications to nonholonomic mechanics". In: *Reports on Math. Physics* 42.1-2 (1998), pp. 135–164. DOI: 10.1016/S0034-4877(98)80008-6.

[Lin95] E. Lindelof. In: *Acta Societatis Scientiarum Fennicae* XX.10 (1895).

[LL21] M. de León and M. Lainz. *A review on contact Hamiltonian and Lagrangian systems*. 2021. arXiv: 2011.05579 [math-ph].

[LL66] Josiane Lehmann-Lejeune. "Intégrabilité des G-structures définies par une 1-forme 0-déformable à valeurs dans le fibré tangent". fr. In: *Annales de l'Institut Fourier* 16.2 (1966), pp. 329–387. DOI: 10.5802/aif.246. URL: http://www.numdam.org/item/AIF_1966__16_2_329_0/.

[LLV19a] Manuel de León and Manuel Lainz Valcázar. "Contact Hamiltonian systems". In: *Journal of Mathematical Physics* 60.10 (2019), p. 102902. DOI: 10.1063/1.5096475. eprint: https://doi.org/10.1063/1.5096475. URL: https://doi.org/10.1063/1.5096475.

[LLV19b] Manuel de León and Manuel Lainz Valcázar. "Singular Lagrangians and precontact Hamiltonian systems". In: *International Journal of Geometric Methods in Modern Physics* 16.10 (2019), p. 1950158. DOI: 10.1142/S0219887819501585. eprint: https://doi.org/10.1142/S0219887819501585. URL: https://doi.org/10.1142/S0219887819501585.

[LM87] Paulette Libermann and Charles-Michel Marle. *Symplectic geometry and analytical mechanics*. Vol. 35. Mathematics and its Applications. Translated from the French by Bertram Eugene Schwarzbach. D. Reidel Publishing Co., Dordrecht, 1987, pp. xvi+526. ISBN: 90-277-2438-5. DOI: 10.1007/978-94-009-3807-6. URL: https://doi.org/10.1007/978-94-009-3807-6.

[LM95] A.D. Lewis and R.M. Murray. “Variational principles for constrained systems: theory and experiment”. In: *Int. J. Non-linear Mechanics* 30.6 (1995), pp. 793–815.

[LMD10] M. de León, J.C. Marrero, and D. Martín de Diego. “Linear almost Poisson structures and Hamilton-Jacobi equation. Applications to nonholonomic mechanics”. In: *J. Geom. Mech.* 2.2 (2010), pp. 159–198. DOI: `10.3934/jgm.2010.2.159`.

[LR89] M. de León and P.R. Rodrigues. *Methods of differential geometry in analytical mechanics.* North Holland Mathematics Studies 158, 1989.

[LV20] Manuel de León and Manuel Lainz Valcázar. “Infinitesimal symmetries in Contact Hamiltonian systems”. In: *Journal of Geometry and Physics* 153 (2020), p. 103651. arXiv: `1909.07892`.

[Mar96] C.-M. Marle. “Kinematic and geometric constraints, servomechanism and control of mechanical systems”. In: vol. 54. 4. Geometrical structures for physical theories, II (Vietri, 1996). 1996, pp. 353–364.

[MCS93] E. Martínez, J. F. Cariñena, and W. Sarlet. “Derivations of differential forms along the tangent bundle projection. II”. In: *Differential Geom. Appl.* 3.1 (1993), pp. 1–29. ISSN: 0926-2245. DOI: `10.1016/0926-2245(93)90020-2`. URL: `https://doi.org/10.1016/0926-2245(93)90020-2`.

[MDM06] J.C. Marrero, D. Martín de Diego, and E. Martínez. “Discrete Lagrangian and Hamiltonian mechanics on Lie groupoids”. In: *Nonlinearity* 19 (2006), pp. 1313–1348.

[MDM16] J.C. Marrero, D. Martín de Diego, and E. Martínez. *On the exact discrete lagrangian function for variational integrators: theory and applications. arXiv:1608.01586v1 [math.DG]*. 2016. URL: `https://arxiv.org/abs/1608.01586`.

[MDM21] J.C. Marrero, D. Martín de Diego, and E. Martínez. *Local convexity for second order differential equations on a Lie algebroid.* To appear in *Journal of Geometric Mechanics*, 2021. 2021. arXiv: `2103.14418 [math.DG]`.

[Mes05] Tom Mestdag. “Lagrangian reduction by stages for non-holonomic systems in a Lie algebroid framework”. In: *Journal of Physics A: Mathematical and General* 38.47 (2005), pp. 10157–10179. DOI: 10.1088/0305-4470/38/47/008. URL: https://doi.org/10.1088/0305-4470/38/47/008.

[ML05] Tom Mestdag and Bavo Langerock. “A Lie algebroid framework for non-holonomic systems”. In: *Journal of Physics A: Mathematical and General* 38.5 (2005), pp. 1097–1111. DOI: 10.1088/0305-4470/38/5/011. URL: https://doi.org/10.1088/0305-4470/38/5/011.

[MP06] R. McLachlan and M. Perlmutter. “Integrators for Nonholonomic Mechanical Systems”. In: *J. Nonlinear Sci.* 16 (2006), pp. 283–328.

[MR13] J.E. Marsden and T.S. Ratiu. *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*. Texts in Applied Mathematics. Springer New York, 2013. ISBN: 9780387217925. URL: https://books.google.pt/books?id=k-7kBwAAQBAJ.

[MS17] D. McDuff and D. Salamon. *Introduction to Symplectic Topology*. Oxford: Oxford Graduate Texts in Mathematics 27, Third Edition, 2017.

[MS96] R. I. McLachlan and C. Scovel. “A survey of open problems in symplectic integration”. In: *Integration algorithms and classical mechanics (Toronto, ON, 1993)*. Vol. 10. Fields Inst. Commun. Amer. Math. Soc., Providence, RI, 1996, pp. 151–180.

[MV19] K Modin and O. Verdier. “What makes nonholonomic integrators work?” In: *arXiv:1707.04514v2* (2019).

[MV20] Klas Modin and Olivier Verdier. “What makes nonholonomic integrators work?” In: *Numer. Math.* 145.2 (2020), pp. 405–435. ISSN: 0029-599X. DOI: 10.1007/s00211-020-01126-y. URL: https://doi.org/10.1007/s00211-020-01126-y.

[MV91] J. Moser and A.P. Veselov. “Discrete versions of some classical integrable systems and factorization of matrix polynomials”. In: *Commun.Math. Phys* 139 (1991), 217–243. DOI: https://doi.org/10.1007/BF02352494.

[MW01] J. E. Marsden and M. West. “Discrete mechanics and variational integrators”. In: *Acta Numer.* 10 (2001), pp. 357–514. ISSN: 0962-4929. DOI: 10.1017/S096249290100006X.

[NF72] Ju. I. Neĭmark and N. A. Fufaev. *Dynamics of nonholonomic systems*. Vol. 33. Translations of Mathematical Monographs. Translated from the 1967 Russian original by J. R. Barbour. American Mathematical Society, Providence, RI, 1972, pp. ix+518. ISBN: [0-8218-1583-0].

[Ohs+11] Tomoki Ohsawa et al. “Nonholonomic Hamilton–Jacobi theory via Chaplygin Hamiltonization”. In: *Journal of Geometry and Physics* 61.8 (2011), pp. 1263–1291. ISSN: 0393-0440. DOI: https://doi.org/10.1016/j.geomphys.2011.02.015. URL: https://www.sciencedirect.com/science/article/pii/S0393044011000428.

[O'N83] B. O'Neill. *Semi-Riemannian Geometry With Applications to Relativity, 103, Volume 103 (Pure and Applied Mathematics)*. Academic Press, 1983. URL: http://www.amazon.com/Semi-Riemannian-Geometry-Applications-Relativity-Mathematics/dp/0125267401.

[PC09] G. W. Patrick and C. Cuell. “Error analysis of variational integrators of unconstrained Lagrangian systems”. In: *Numer. Math.* 113.2 (2009), pp. 243–264. ISSN: 0029-599X. DOI: 10.1007/s00211-009-0245-3.

[PL19] Helen Parks and Melvin Leok. “Constructing equivalence-preserving Dirac variational integrators with forces”. In: *IMA J. Numer. Anal.* 39.4 (2019), pp. 1706–1726. ISSN: 0272-4979. DOI: 10.1093/imanum/dry053.

[SCS95] W Sarlet, F Cantrijn, and D J Saunders. “A geometrical framework for the study of non-holonomic Lagrangian systems”. In: *Journal of Physics A: Mathematical and General* 28.11 (1995), pp. 3253–3268. DOI: 10.1088/0305-4470/28/11/022. URL: https://doi.org/10.1088/0305-4470/28/11/022.

[SCS99] D J Saunders, F Cantrijn, and W Sarlet. “Regularity aspects and Hamiltonization of non-holonomic systems”. In: *Journal of Physics A: Mathematical and General* 32.39 (1999), pp. 6869–

6890. DOI: 10.1088/0305-4470/32/39/311. URL: https://doi.org/10.1088/0305-4470/32/39/311.

[Shi+17] Donghua Shi et al. "Hamel's Formalism for Infinite-Dimensional Mechanical Systems". In: *Journal of Nonlinear Science* 27 (2017), pp. 1–43. DOI: 10.1007/s00332-016-9332-7.

[Sim+20] A. A. Simoes et al. "Contact geometry for simple thermodynamical systems with friction". In: *Proc. R. Soc. A* 476.20200244 (2020). DOI: http://dx.doi.org/10.1098/rspa.2020.0244.

[SSC94] J. M. Sanz-Serna and M. P. Calvo. *Numerical Hamiltonian problems*. Vol. 7. Applied Mathematics and Mathematical Computation. Chapman & Hall, London, 1994, pp. xii+207. ISBN: 0-412-54290-0.

[SSC96] D J Saunders, W Sarlet, and F Cantrijn. "A geometrical framework for the study of non-holonomic Lagrangian systems: II". In: *Journal of Physics A: Mathematical and General* 29.14 (1996), pp. 4265–4274. DOI: 10.1088/0305-4470/29/14/042. URL: https://doi.org/10.1088/0305-4470/29/14/042.

[Syn28] J.L. Synge. "Geodesics in non-holonomic geometry". In: *Mathematische Annalen* 99 (1928), pp. 738–751.

[SZB20] Donghua Shi, Dmitry Zenkov, and Anthony Bloch. "Hamel's Formalism for Classical Field Theories". In: *Journal of Nonlinear Science* 30 (2020). DOI: 10.1007/s00332-020-09609-w.

[Vai94] I. Vaisman. *Lectures on the geometry of Poisson manifolds*. Vol. 118. Progress in Mathematics. Birkhäuser Verlag (Basel), 1994, p. 205. ISBN: 3764350164.

[VBS19] Mats Vermeeren, Alessandro Bravetti, and Marcello Seri. "Contact variational integrators". In: *J. Phys. A* 52.44 (2019), pp. 445206, 28. ISSN: 1751-8113. DOI: 10.1088/1751-8121/ab4767. URL: https://doi.org/10.1088/1751-8121/ab4767.

[Ves88] A.P. Veselov. "Integrable discrete-time systems and difference operators". In: *Functional Analysis Its Applications* 22 (1988), 83–93. DOI: https://doi.org/10.1007/BF01077598.

[VF72] A. M. Vershik and L. D. Faddev. "Differential Geometry and Lagrangian Mechanics with Constraints." In: *Soviet Physics Doklady* 17 (1972), p. 34.

[VM94] A.J. Van Der Schaft and B.M. Maschke. "On the Hamiltonian formulation of nonholonomic mechanical systems". In: *Reports on Mathematical Physics* 34.2 (1994), pp. 225–233. ISSN: 0034-4877. DOI: https://doi.org/10.1016/0034-4877(94)90038-8. URL: https://www.sciencedirect.com/science/article/pii/0034487794900388.

[Vra28] G. Vranceanu. "Parallélisme dans une varieté non holonome". In: *Congresso Internazionale dei Matematici, Bologna* V (1928), pp. 61–67. URL: https://www.mathunion.org/fileadmin/ICM/Proceedings/ICM1928.5/ICM1928.5.ocr.pdf.

[VV88] A. P. Veselov and L. E. Veselova. "Integrable nonholonomic systems on Lie groups". In: *Mat. Zametki* 44.5 (1988), pp. 604–619, 701. ISSN: 0025-567X. DOI: 10.1007/BF01158420. URL: https://doi.org/10.1007/BF01158420.

[YI73] K. Yano and Sh. Ishihara. *Tangent and cotangent bundles: differential geometry*. Pure and Applied Mathematics, No. 16. Marcel Dekker, Inc., New York, 1973, pp. ix+423.

[ZB03] Dmitry V Zenkov and Anthony M Bloch. "Invariant measures of nonholonomic flows with internal degrees of freedom". In: *Nonlinearity* 16.5 (2003), pp. 1793–1807. DOI: 10.1088/0951-7715/16/5/313. URL: https://doi.org/10.1088/0951-7715/16/5/313.

[ZBM98] D. V. Zenkov, A. M. Bloch, and J. E. Marsden. "The energy-momentum method for the stability of non-holonomic systems". In: *Dynamics and Stability of Systems* 13.2 (1998), pp. 123–165. DOI: 10.1080/02681119808806257. eprint: https://doi.org/10.1080/02681119808806257. URL: https://doi.org/10.1080/02681119808806257.