Some geometric aspects of implicit dynamical systems

Ph. D. Thesis

Rubén Martín Grillo
Advisor: Dr. Xavier Gràcia Sabaté

Departament de Matemàtica Aplicada IV
Universitat Politècnica de Catalunya

Barcelona, November 2006

A mis padres
y a toda mi familia

Agraïments / Agradecimientos

Vull expressar la meva gratitud envers les persones que m'han ajudat a realitzar aquest treball i m'han acompanyat i recolzat durant aquest temps.

Em primer lloc al meu tutor i director Xavier Gràcia per rebre'm amb els braços oberts, introduir-me en el món de la investigació, transmetre'm coneixements, donar-me idees, guiar-les i indicar els meus errors, això i més coses que em deixo, sempre amb esperit positivament crític, paciència i bon humor.

També vull donar les gràcies als membres del Departament de Matemática Aplicada IV de la UPC per la seva bona acollida, especialment als membres del grup de recerca en Geometria Diferencial, Sistemes Dinàmics i Aplicacions. I molt especialment a Miguel Muñoz i Narciso Román pels seus consells i recolzament constant. Vull mencionar també a Jaume Franch per fer que la vida en un despatx sigui més agradable.

Quiero dar gracias a mi familia, que, entiendan o no el contenido de mi trabajo, siempre han confiado en mí. Especialmente a mis padres por dármelo todo y contagiarme desde pequeño el deseo de aprender.

Y a los amigos con los cuales he compartido momentos de diversos colores, sean Chemas, Davides, Elisas, Joanes, Lauras o Marías.

Contents

Chapter 1

Introduction

The concept of dynamical system arises from the need to represent real physical systems with a mathematical language, so that mathematical tools can be used to deduce some properties of the physical system, maybe even to solve it completely (that is, to know exactly its behaviour at any time). Although there are more general definitions for the concept of dynamical system (which may include, for example, discrete systems) we will work with the one which is more “geometric”, in the sense that the elements of differential geometry are used. In fact, many of the structures and tools of differential geometry were developed during and for the study of dynamical systems.

A dynamical system is constituted by a manifold M, whose points represent the possible states of the physical system, and a differential equation on M, which represents the rule that governs the behaviour of the physical system. Thus, in every local chart $(U, \mathbf{x})$ of M, there is a differential equation

$$F\left(t, \mathbf{x}, \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}^k\mathbf{x}}{\mathrm{d}t^k}\right) = 0. \tag{1.1}$$

It is known that such a k-th order differential equation is equivalent to a first-order differential equation on another “extended” manifold (namely, the $(k-1)$-th order tangent bundle $\mathrm{T}^{k-1}M$). To work with a higher-order differential equation or with a first-order differential equation on a manifold of higher dimension is a question of the approach one takes to the problem. For instance, one of the pillars of Newtonian mechanics is that the equation of motion is a second-order differential equation in the configuration space, whereas the equations of the Hamiltonian formulation of mechanics are first-order differential equations in the phase space.

In this dissertation we will not consider the general case represented by equation (1.1), but the particular case represented by differential equations that are affine in the highest-order derivative:

$$\mathbf{A}\left(t, \mathbf{x}, \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}^{k-1}\mathbf{x}}{\mathrm{d}t^{k-1}}\right) \cdot \frac{\mathrm{d}^k\mathbf{x}}{\mathrm{d}t^k} = \mathbf{b}\left(t, \mathbf{x}, \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}^{k-1}\mathbf{x}}{\mathrm{d}t^{k-1}}\right),$$

where $\mathbf{A}$ is a (generically singular) matrix and $\mathbf{b}$ is a vector. Of course, the main case is the first-order case

$$\mathbf{A}(t, \mathbf{x}) \cdot \dot{\mathbf{x}} = \mathbf{b}(t, \mathbf{x}), \tag{1.2}$$

since a higher-order equation can be transformed into it.

If $\mathbf{A}$ is a regular matrix, then the highest-order derivative can be isolated, obtaining the equation

$$\dot{\mathbf{x}} = \mathbf{c}(t, \mathbf{x}),$$

where $\mathbf{c} = \mathbf{A}^{-1} \cdot \mathbf{b}$. Then it is said that the differential equation is written in *normal form*. Equations in normal form are easier to solve: at least, by the theorem of existence and uniqueness of solutions of ordinary differential equations we know that local solutions of initial value problems exist and are unique.

From the geometrical point of view, a dynamical system with differential equation in normal form is associated with a (time-dependent) vector field, and the integral curves of this vector field are the solutions of the dynamical system. Therefore, it makes sense to study equation (1.2) with the aim of finding an equivalent differential equation in normal form, though perhaps not defined on the whole manifold. Thus, **one of the goals of this dissertation is to give a geometrical formulation for the dynamical systems whose differential equation is of the kind of equation (1.2) and to find equivalent systems with equation in normal form (or, in other words, to find vector fields whose integral curves are solutions of the dynamical system).**

It may seem too restrictive to consider only affine differential equations, but there are strong motivations for the particular study of them. Our main reason is that several formalisms of mechanics lead to equations of this kind, amongst them, the most important two: the Lagrangian and Hamiltonian formalisms. In fact, both the Lagrangian and Hamiltonian formalisms have presymplectic formulations, and the equation of the presymplectic systems are affine. Some constrained mechanical systems also fall in this category, notably the so-called nonholonomic systems. So, **a second objective of this thesis is to show how various formulations of mechanics fit into our geometrical framework, with particular attention to nonholonomic systems.** Although not discussed in this work, this kind of equations also arise in other fields, such as electric o chemical engineering, control theory or economics —see some references in [BCP 96, GMR 04].

We will see that the geometric formulation of the differential equation (1.2) is different whether $\mathbf{A}$ and $\mathbf{b}$ are autonomous (independent of the evolution variable t) or not. It is well-known that a time-dependent differential equation may be transformed into an autonomous one by means of the following "trick": consider the independent variable t as dependent of a new variable s, satisfying the differential equation $\frac{\mathrm{d}t}{\mathrm{d}s} = 1$.

Since clearly $t = s + c$ for some constant c, equation (1.1) is equivalent to

$$\begin{cases} F\left(t, \mathbf{x}, \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}, \dots, \frac{\mathrm{d}^k\mathbf{x}}{\mathrm{d}s^k}\right) = 0 \\ \frac{\mathrm{d}t}{\mathrm{d}s} = 1 \end{cases} .$$

Hence, at the cost of extending the set of dependent variables to $(t, \mathbf{x})$ we now have an autonomous differential equation. In some cases, this may be advantageous. For example, the autonomous differential equation may be easier to solve than the original time-dependent differential equation, or we may know some method to solve the autonomous differential equation that can not be directly applied to solve time-dependent differential equations. The "trick" described above deals with differential equations and we have given geometric formulations for theses differential equations. Therefore, **a third objective of this work is to develop a geometrical equivalent of the process of converting a time-dependent differential equation into an autonomous one, that fits well with the geometrical formulations proposed for the autonomous and time-independent cases.** It turns out that the concept of **vector hull** is the appropriate geometrical tool, and a chapter of this dissertation will be devoted to its discussion.

Now let us give an outline of this dissertation. We will explain the contents of the different chapters, indicating the original contributions as well as the most relevant references.

Chapter 2. Review of results on differential geometry

In this chapter we give the basic mathematical tools from differential geometry that are used throughout the dissertation, and we fix the notation.

We review some notions like tangent bundle, submanifold, tensor field, riemannian metric or (pre)symplectic form. These concepts can be found with more detail in some reference books as [AM 78, Car 92, KN 63, Lan 85, LM 87].

We study more deeply the theory of bundles. Roughly speaking, a bundle consists of a surjective submersion $\pi: A \to B$ such that all the fibres $\pi^{-1}(b)$ are isomorphic to a "typical fibre". We put special emphasis on vector bundles and affine bundles, which have a vector space and an affine space respectively as typical fibre. We also review the theory of jet bundles, which is the framework where time-dependent systems and derivatives are naturally described. Basically, we have followed the book by Saunders [Sau 89] for the exposition of the topics on bundles.

Finally, we study the geometry of the tangent and cotangent bundles, as well as higher order tangent bundles. We will show later that these concepts have a fundamental role in the geometric description of Lagrangian and Hamiltonian systems; this is explained in detail in [LR 89].

Chapter 3. Time independent systems

Here we review a geometrical model for the differential equation

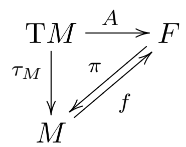
$$\mathbf{A}(\mathbf{x})\dot{\mathbf{x}} = \mathbf{b}(\mathbf{x}), \tag{1.3}$$

that is, equation (1.2) when the matrix **A** and the vector **b** are independent of time. We give the name of *(autonomous) linearly singular system* to this structure because the matrix **A** is singular in general. Autonomous linearly singular systems were first presented by Gràcia and Pons [GP 91] and subsequently developed in [GP 92a, GP 02]. The first part of this chapter follows these works.

An autonomous linearly singular system consists of these elements:

- a vector bundle $F \to M$ over the configuration manifold M
- a vector bundle morphism $A\colon \mathrm{T}M \to F$ between the tangent bundle $\mathrm{T}M$ and a vector bundle F
- a section $b\colon M \to F$ of the vector bundle F

and it is denoted by $(A\colon \mathrm{T}M \to F, b)$. We show all these elements in the following diagram:



Then, for a curve γ in M, the equation

$$A \circ \dot{\gamma} = b \circ \gamma$$

has equation (1.3) as local expression. We can also try to find the vector fields on M such that its integral curves are solutions of the system. In this case, the condition is, for a vector field X,

$$A \circ X = b. \tag{1.4}$$

Since A is singular, we do not have a theorem of existence and uniqueness of solutions, so we can not expect a global solution. In order to identify the subset of the configuration manifold where the system has solutions one has to apply a recursive procedure. Procedures like the presented here are generically known as constraint algorithms. The first one was given, in a non-geometrical language, by Bergmann [BG 55] and Dirac [Dir 64], and it was conceived to deal with the Hamiltonian formulation of system with singular lagrangian. A geometric version was developed by Gotay and Nester [GNH 78, GN 79] for presymplectic systems. Since a presymplectic system is a time-independent linearly singular system, the algorithm presented here is a generalization of the Gotay–Nester one. Geometric constraint algorithms for first-order implicit differential equations of the form $F(\mathbf{x}, \dot{\mathbf{x}}) = 0$ were given in [RR 94, MMT 95].

We also give some notions on regularity and symmetries of autonomous linearly singular systems.

The aim of the second part of this chapter is to show that different formulations of (time-independent) Lagrangian systems can be written as autonomous linearly singular systems.

The basic notions and the usual symplectic formulation of Lagrangian systems can be found, for instance, in [AM 78]. The differences between regular and singular systems are considered, notably the need to add, in the singular case, a "second–order condition" that is directly fulfilled in the regular case.

We present two less-known formulations of Lagrangian systems, which overcome this problem. The first one makes use of the so-called time-evolution operator K, which was introduced in coordinates in [BGPR 86] and was defined geometrically as a vector field along the Legendre transformation in [GP 89]. The other formulation is the first-order formulation of Skinner and Rusk [Ski 83, SR 83].

We also review the Hamiltonian formalism of Lagrangian systems, following [Car 90].

Finally we show how these different formulations of Lagrangian systems can be considered as autonomous linearly singular systems.

Chapter 4. Generalized nonholonomic systems

Except for the introduction to nonholonomic Lagrangian systems, the contents of this chapter are original, they were partially presented in [GM 04a] and can be found in [GM 05a].

It is not unusual that a mechanical system has constraints, that is, some states (position and velocities) of the system are forbidden. In other words, the derivative of any solution of the system must lie on a submanifold, determined by the constraints, of the tangent bundle of the configuration space.

If the constraints of the states of the system are determined by constraints of the configuration space, it is said that the constraints are holonomic. For example, a pendulum is constrained to move on a sphere, so the velocities must be tangent to this sphere. Another possibility is that the constraints are integrable (also called semi-holonomic constraints), so the configuration space is foliated by integral submanifolds and the motion evolves separately in each integral submanifold. In both cases the system can be reduced to subsystems that can be considered unconstrained.

Here we will consider the remaining case. Nonholonomic mechanical systems are mechanical systems with non-integrable constraints. An easy example is a disk rolling without sliding on a horizontal plane, and many wheeled vehicles has been modelled with nonholonomic systems. More examples can be found in [NF 72].

Nonholonomic systems have been discussed since the last years of nineteenth century, however, the geometric foundations for the theory were given in the 1970s with the work by Vershik and Faddeev [VF 72]. Since then, different geometric approaches have been taken to deal with the subject, for example, a hamiltonian approach in [BŚ 93], a lagrangian approach in [LM 96], a more general Poisson framework in [Mar 98], an approach based on a gauge independent formulation of lagrangian and hamiltonian mechanics in [MVB 02], or a model based on ideals of differential forms and distributions [KM 01, Kru 02]. Symmetries of these systems, as well as reduction schemes derived from them, have also been considered in the literature, see [Koi 92, BKMM 96, KM 98, CL 99, Mar 03]. A very comprehensive work on nonholonomic systems and specifically its geometric control can be found in [Cor 02].

In the first section we give the basic notions on nonholonomic systems and we present the so-called Chetaev's rule that provides the dynamics (although we note that other principles has been proposed). We find that the resulting equation of motion, in coordinates, has the form

$$\begin{cases} \mathbf{A}(\mathbf{x})\dot{\mathbf{x}} - \mathbf{b}(\mathbf{x}) = \sum_{\alpha=1}^{m} \lambda_\alpha \mathbf{f}^\alpha(\mathbf{x}) \\ \phi^\alpha(\mathbf{x}) = 0, \quad 1 \le \alpha \le m \end{cases} ,$$

where ϕ^α are m functions that define the constraints and λ_α are multipliers to be determined. In terms of vector fields, the equation of motion has the following general form:

$$\begin{cases} (A|_C) \circ X - b|_C \in F' \\ X \in \mathfrak{X}(C) \end{cases} . \tag{1.5}$$

This is nearly the equation of a linearly singular system (see equation (1.4)), but we notice two modifications: the restriction of the vector bundle morphism A and the section b to some submanifold C and the appearance of some subbundle F' of the vector bundle F (the range of A and b). We are not going to specify here which precise objects are the base manifold M, A, b, C, F and F' in the nonholonomic case, but let us say that $(A\colon \mathrm{T}M \to F, b)$ is the linearly singular system associated with the unconstrained dynamics.

In the second section we see that equation (1.5) is indeed the equation of a linearly singular system, and that this linearly singular system can be constructed from the linearly singular system $(A\colon \mathrm{T}M \to F, b)$. The key is that we can perform two particular operations with linearly singular systems: restriction to a subsystem and projection to a quotient. With each operation we obtain a linearly singular system from another linearly singular system, and combining both of them we can obtain a system with equation (1.5) from the system $(A\colon \mathrm{T}M \to F, b)$.

We call *generalized nonholonomic systems* to the systems obtained from another by restriction and projection to a quotient. The reason for this name is clear, since nonholonomic systems can be seen as a particular case. The detailed proof of this fact is given in section 4.

We study the regularity, consistency and equations of motion of the generalized nonholonomic systems in relation to the "original system" in the second section. In the third section we study the circumstances under which symmetries and constants of motion of a linearly singular system induce symmetries and constants of motion of a generalized nonholonomic system derived from it.

In section 5 we deal with another class of dynamical systems that has been a topic of research for the last decade: the implicit Hamiltonian systems. We review the basic notions on these systems and we show that they fall into the class of generalized nonholonomic systems.

The notion of implicit hamiltonian system was introduced by van der Schaft and Maschke [SM 95a, SM 95b], and it can be regarded as a generalization of the notion of symplectic and Poisson systems. Implicit hamiltonian systems can be used to model a wide range of physical systems. For example, physical systems with a singular Lagrangian or mechanical systems with linear nonholonomic constraints naturally give rise to implicit hamiltonian systems.

Furthermore, external variables can be easily added to obtain the so-called implicit port-controlled Hamiltonian systems, which are suitable to model control systems as for instance electrical LC-circuits. As well, by means of the external variables, these systems can be interconnected, allowing a modular approach to the modeling process.

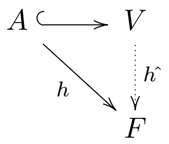
In section 6 we apply the theory to the example of a relativistic particle. We consider two lagrangian functions, the usual singular lagrangian and a regular one, studied in [KM 01], where the authors study the relativistic particle as a mechanical system with nonholonomic constraints. We study both systems without constraints and with the nonholonomic constraint $v^2 = c^2$, and we find that the singular system becomes regular with the addition of the constraint.

We finish the chapter with two simple examples to illustrate the relations between the symmetries and constants of motion of the unconstrained and constrained systems. One of the examples, the particle in $\mathbf{R}^3$ under the nonholonomic constraint $\dot{z} - y\dot{x} = 0$ was proposed in [Ros 77] and it has been studied in various papers about reduction of nonholonomic systems [BGM 96, BKMM 96, BŚ 93, CL 99].

Chapter 5. Vector hulls

In this chapter we leave aside the differential equations and we undertake the study of the vector hull, a topic more related to linear algebra. Firstly, we are interested in this subject because vector hulls can be applied to convert, in a neat geometrical way, a time-dependent differential equation into an autonomous one. We will see this application in the next chapter, devoted to time-dependent systems. Nevertheless, vector hulls have other applications and are interesting by themselves, so we want to have a closer look at them. The contents of this chapter are presented in [GM 06] and have been partially published in [GM 05c].

Roughly speaking, a vector hull of an affine space A (over a field K) is a vector space V that contains A as a proper hyperplane (that is, an hyperplane that does not contain the zero vector). Such a vector space satisfies the following universal property: for every vector space F and affine function $h: A \to F$, there exists a unique linear function $\hat{h}: V \to F$ such that the following diagram is commutative

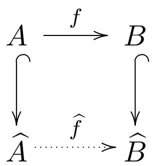


As it is expected, all the vector hulls of A are isomorphic, and we denote by $\widehat{A}$ any of them. One obvious choice for the vector hull of A is $K \oplus \vec{A}$, where $\vec{A}$ is the vector space associated with A. The drawback of this construction is that the inclusion of A into $K \oplus \vec{A}$ depends on the choice of a privileged point in A, so it is not a canonical construction.

There are various canonical constructions of the vector hull in the bibliography. We show with detail that the set of affine maps $X: A \to \vec{A}$ such that its associated linear map $X: \vec{A} \to \vec{A}$ is a homothety is a vector hull of A. This construction is inspired by the construction given in [BB 75] (included in the book [Sch 75]), where the homothetic vector fields are characterized by another property called equiprojectivity.

The first canonical construction of the vector hull appears as an exercise of Bourbaki's *Algèbre*[Bou 70], where it is constructed as a quotient of the vector space $K^{(A)}$, that has all the points of A as a basis, by a suitable subspace. A similar construction appears in [GGU 03]. Another usual choice, only valid when A is of finite dimension, is the dual space of affine functions from A to K: $\mathscr{A}(A, K)^*$. This construction is used, for instance, in [MMS 02]. Other constructions appear in [Ber 77, Fre 73, Gal 01, IREM 75].

Along with vector hulls of affine spaces come vector prolongations of affine maps. We see that for every affine map map $f: A \to B$ there is a unique linear map $\widehat{f}: \widehat{A} \to \widehat{B}$, called the vector prolongation of f, such that the following diagram is commutative:



It turns out that the assignment $A \rightsquigarrow \widehat{A}$, $f \rightsquigarrow \widehat{f}$, is a covariant functor from the category of K-affine spaces to the category of K-vector spaces.

The first half of the chapter is a review of the vector hull functor, presenting also new results about it. In the first section we review the basic notions on affine spaces and give some results on hyperplanes that will be used later. In particular, it is crucial the fact that the set of proper hyperplanes of a vector space is in bijection with the set of non-vanishing linear forms.

In the second section we define the vector hull and the vector prolongation, and we state some properties of these objects. It will be of special relevance that if there exists an exact sequence of vector spaces

$$0 \longrightarrow \vec{A} \stackrel{i}{\longrightarrow} W \stackrel{w}{\longrightarrow} K \longrightarrow 0 \tag{1.6}$$

then W is a vector hull of A. We also study the case when the affine space A is indeed a vector space (then $\widehat{A} = K \times A$ canonically) and finally we take a look on the coordinate expressions of the different objects defined.

The notion of vector hull has various applications in linear algebra. For example, in the framework of affine geometry, expressions like $a + \mathbf{u} = b$ and $\mathbf{u} = b - a$, and barycentric combinations like $\sum \lambda_i a_i$ (where $\sum \lambda_i = 1$), have a neat interpretation under the vector space structure of the vector hull $\widehat{A}$. Another application is a pretty interpretation of the well-known linear representation of the affine group $\mathbf{GA}(n, K)$ as a subgroup of the linear group $\mathbf{GL}(n+1, K)$. There is also benefit for projective geometry: there is a canonical inclusion $A \hookrightarrow \mathbf{P}(\widehat{A})$, so that $\mathbf{P}(\widehat{A})$ is the projective completion of A, and affine transformations yield projective transformations in the projective completion [Fre 73] [Ber 77]. These applications of the vector hull are explained in section 3.

There are another examples of the use of the vector hull in the literature. For instance, they have been applied to the study of mechanics of rigid bodies (see for instance [BB 75]), or to devise algorithms to draw curves and surfaces for computer-aided geometric design [Ram 89] [Gal 01].

In sections 4 and 5 we explain our construction of the vector hull and discuss the other constructions mentioned above.

The vector hull functor can be extended to the categories of affine bundles and vector bundles, that is, every affine bundle $\pi: A \to M$ has a vector hull, denoted by $\widehat{\pi}: \widehat{A} \to M$. Without going into details, each fibre $(\widehat{A})_m$ of the vector hull is the vector hull $\widehat{(A_m)}$ of the fibre A_m. All the properties of the vector hull functor of affine spaces extend well to the vector hull of affine bundles. We discuss these subjects in section 6.

In the field of differential geometry, jet manifolds provide with examples of affine bundles. For instance, the first-order jet space $\mathrm{J}^1 M$ of a bundle $M \to \mathbf{R}$ is an affine bundle over M. There is a canonical affine immersion of $\mathrm{J}^1 M$ in the tangent bundle $\mathrm{T}M$, and it turns out that this bundle is a model of the vector hull of $\mathrm{J}^1 M$. Although in this chapter we do not talk about differential equations, here we already glimpse a relation between vector hulls and differential equations, since (first-order) time-dependent differential equations are geometrically modelled on (first-order) jet spaces and (first-order) autonomous differential equations on (first-order) tangent bundles. In section 7 it is shown that the vector hull of the k-th order jet bundle $\mathrm{J}^k M \to \mathrm{J}^{k-1} M$ of a bundle $M \to \mathbf{R}$ can be identified with the so-called Cartan distribution on $\mathrm{J}^{k-1} M$.

Jet bundles over an arbitrary base are discussed in section 8. This may be useful in the study of partial differential equations, for instance, these jet bundles are the appropriate geometrical framework for the description of classical field theories [CCI 91]. Indeed, the vector hull of the first-order jet bundle is the dual of the so-called multimomentum bundle, which arises in the multisymplectic formalism of field theories—see, for instance [EMR 00]. We identify the vector hull of the k-th order jet bundle $\mathrm{J}^k M \to \mathrm{J}^{k-1} M$ of a bundle $M \to B$ with a subspace of the space of homomorphisms $\mathrm{Hom}(\pi_{k-1}^* \mathrm{T}B, \mathrm{TJ}^{k-1}\pi)$. This result includes all the cases studied in sections 6 and 7. We point out that the proof of all the results of these two sections are based on the use of an exact sequence (1.6).

Finally, in section 9 we study the second order tangent bundle $\mathrm{T}^2 M \to \mathrm{T}M$, which is another affine bundle. This case is different from the previous cases because $\mathrm{T}^2 M$ is naturally included into the vector bundle $\mathrm{T}(\mathrm{T}M)$, but its vector hull can not be embedded into $\mathrm{T}(\mathrm{T}M)$.

Chapter 6. Time-dependent systems

The main objective of this chapter is to extend the theory of autonomous linearly singular systems exposed in chapter 3 to the time-independent case. Most of the contents of this chapter are exposed in [GM 04b, GM 05b].

We study the geometric framework of time-dependent first-order implicit differential equations,

$$F(t, \mathbf{x}, \dot{\mathbf{x}}) = 0.$$

Time-dependent systems in general are studied in many books, as for instance [AM 78, Olv 93]. We are interested in the case when F is affine in the velocities, the case that we call linearly singular case, represented by equation (1.2)

$$\mathbf{A}(t, \mathbf{x}) \cdot \dot{\mathbf{x}} = \mathbf{b}(t, \mathbf{x}),$$

where $\mathbf{A}$ is a matrix that is generically singular.

First we should point out that our model for the time-dependent configuration space, rather than a trivial product $M = \mathbf{R} \times Q$, is a fibre bundle $\rho \colon M \to \mathbf{R}$, where the base $\mathbf{R}$ contains the time variable. Such an M is isomorphic to a product $\mathbf{R} \times Q$, but in practical applications there may not be a privileged trivialization, and a possible extension to deal with field theory of course should not be based on a trivial bundle.

The basic difference between the formulation of the autonomous and the non-autonomous case is the use of tangent bundles and jet bundles respectively. To describe an autonomous differential equation on a configuration space Q we use the tangent bundle $\mathrm{T}Q$, which is a vector bundle. On the other hand, to describe a non-autonomous differential equation on a time-dependent configuration space M, we use its jet bundle $\mathrm{J}^1\rho$, which is an affine bundle over M. Naturally, we will use affine morphisms defined on this affine bundle to describe a linearly singular equation on M.

In section 1 we present a geometrical model for equation (1.2), which we call time-dependent linearly singular system. It consists of the following elements:

- a vector bundle $E \to M$ over the configuration manifold M

- an affine bundle morphism $\mathcal{A}\colon \mathrm{J}^1 M \to E$ between the first-order jet bundle $\mathrm{J}^1 M$ and a vector bundle E.

$$\begin{array}{ccc} \mathrm{J}^1\rho & \xrightarrow{\mathcal{A}} & E \\ {\scriptstyle \rho_{1,0}}\big\downarrow & \swarrow{\scriptstyle \pi} & \\ M & & \end{array}$$

Then, equation

$$\mathcal{A} \circ \mathrm{j}^1 \xi = 0,$$

for a section $\xi\colon I \to M$ has equation (1.2) as local expression.

Now we have the same problem as in the autonomous case. If $\mathcal{A}$ is singular, there is no existence and/or uniqueness of solutions guaranteed. In section 2 we propose a constraint algorithm for time-dependent singular systems. This algorithm is the natural generalization of the algorithm for the autonomous case, both in the general implicit case [RR 94, MMT 95] and the linearly singular case [GP 91, GP 92a], that was presented in chapter 3. The case of an implicit equation in a product $M = \mathbf{R} \times Q$ has already been discussed in [Del 04]. It is worth noting that constraint algorithms for some particular time-dependent systems (arising from mechanics) have been described in several recent works, as for instance [CF 93, CLM 94, ILMM 99, LMM 96, LMMMR 02, Vig 00]. Since all these systems are of linearly singular type, they are included within our framework. Their various algorithms are also particular instances of the general constraint algorithms that we will study here. So, there is a general procedure that can be applied to these several systems, and their particular details are secondary with respect to the algorithm followed to obtain their solutions.

When studying a time-dependent differential equation, sometimes it is useful to convert it into an equivalent time-independent one. This is even more interesting for implicit equations; for instance, the constraint algorithm for the autonomous case is easier to implement than for the non-autonomous case, because of the fact that vector fields instead of jet fields are used to obtain the constraint functions. In the beginning of the introduction we have sketched the usual method (using coordinates) to perform this transformation.

Here in section 3 we examine the possibility of associating an autonomous linearly singular system with a time-dependent one, so that the solutions of both systems will be in correspondence. Essentially, we use the canonical inclusion of $\mathrm{J}^1 M$ into $\mathrm{T}M$. In order to perform this association, we propose two different strategies. One possibility is to choose a connection on the jet bundle to induce a splitting of the tangent bundle. The other possibility, which does not make use of any choice, is based on the notion of vector hull that we studied in the previous chapter. The time-dependent linearly singular system $\mathcal{A}\colon \mathrm{J}^1 M \to E$ is converted into the autonomous linearly singular system $(\widehat{\mathcal{A}}\colon \mathrm{T}M \to \widehat{E}, \widehat{0})$.

The process of converting a time-dependent linearly singular system into an autonomous one is shown in section 4 with a concrete example, a pendulum of variable length.

Since our main motivation for studying the linearly singular systems comes from Euler–Lagrange equations and mechanical systems, where equations of motion are of second order, it is natural to extend the preceding study to second-order implicit and linearly singular equations:

$$F(t, \mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}) = 0, \qquad \mathbf{A}(t, \mathbf{x}, \dot{\mathbf{x}})\ddot{\mathbf{x}} = \mathbf{b}(t, \mathbf{x}, \dot{\mathbf{x}}).$$

This is done in section 5, also exploring the two constructions to pass from a time-independent system to an autonomous one.

Section 6 is devoted to the study of time-dependent Lagrangian systems. We review its formulation with jet bundles, that can be found in [CPT 84, LR 89], for the trivial product case ($M = \mathbf{R} \times Q$), and [LMM 96] for the general fibre bundle case. The equations of motion are found and it is shown that these equations are those of a time-dependent linearly singular system.

In chapter 2, we studied the first-order formulation of autonomous mechanics, geometrically developed by Skinner and Rusk [Ski 83, SR 83]. There is a similar formulation for time-dependent systems, proposed by Cortés, Martínez and Cantrijn [CMC 02]. We review it and see that it is also equivalent to a time-dependent linearly singular system.

Other formulations of time-dependent lagrangian systems can be found in [EMR 91], for the regular case, and [Kru 97, MPL 00, MS 98, LMMMR 02] for the general case; see also references therein.

We saw in chapter 4 that (time-independent) nonholonomic mechanical systems can be modeled in a natural way with (time-independent) linearly singular systems by means of the construction of generalized nonholonomic systems. Here in section 7, we introduce the concept of time-dependent generalized nonholonomic system, which is the analogous construction for the time-dependent case.

Chapter 2

Review of results on differential geometry

Here we are going to review the notions of differential geometry that will be relevant to us, and we will also fix the notation. The reader can find all the details in reference books as [AM 78, KN 63, Lee 03]. We assume that basic concepts of differential geometry such as local chart and differential manifold are known. Unless stated otherwise, we will suppose that the manifolds and mappings we deal with are smooth (that is, of class C^∞). We will also assume that manifolds are finite-dimensional, paracompact and Hausdorff.

2.1 Manifolds and maps

We will denote the linear space of tangent vectors to a manifold M at a point $p \in M$ by T_pM, and the tangent bundle by TM, with projection to the base denoted by $\tau_M\colon TM \to M$. The dual space of a tangent space T_pM is the cotangent space T_p^*M and the cotangent bundle will be denoted by T^*M, with projection $\pi_M\colon T^*M \to M$. Every local chart (U, x^i) of M induces basis $\left(\frac{\partial}{\partial x^i}\big|_p\right)$ of T_pM and $(\mathrm{d}x_p^i)$ of T_p^*M. It also induces local systems of coordinates $(x^i, \dot{x}^i)$ (defined by $\dot{x}^i\left(\frac{\partial}{\partial x^j}\big|_p\right) = \delta_j^i$)) on TM and (x^i, p_i) (defined by $p_i(\mathrm{d}x_p^j) = \delta_i^j$) on T^*M. These systems of coordinates are called *natural coordinates*

The set of maps $f\colon M \to N$ between two manifolds M and N is denoted by $C^\infty(M, N)$. We use the notation $Tf\colon TM \to TN$ for the tangent map of f and $T_pf\colon T_pM \to T_{f(p)}N$ its restriction to a tangent space. If T_pf is surjective (respectively, injective) we say that f is a submersion (respectively, immersion) at p. The map f is called a *submersion* (or *immersion*) if f it is a submersion (or immersion) at every point $p \in M$.

If the target manifold is $N = \mathbf{R}$, a map $f\colon M \to \mathbf{R}$ is called a function on M, and we denote the set of functions on M simply by $\mathrm{C}^\infty(M)$. On the other hand, when the source manifold is an open interval $M = I \subset \mathbf{R}$ (maybe the whole $\mathbf{R}$), we say that a map $\gamma\colon I \to N$ is a path (or a curve) in N. The *derivative* (or *velocity*) of a path $\gamma\colon I \to N$ is the path $\dot\gamma\colon I \to \mathrm{T}N$ in the tangent bundle defined by

$$\dot\gamma(s) = \mathrm{T}_s\gamma\left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_s\right),$$

where t denotes the natural coordinate of I, the identity. If, in local coordinates, the path γ is expressed by $\gamma(t) = (\gamma^i(t))$, then its derivative has local expression $\dot\gamma(t) = (\gamma^i(t), \dot\gamma^i(t))$, where here the dot on the right side of the equation denotes the known derivative of real-valued functions of real variable.

2.2 Tensor fields

A section of the tangent bundle (that is, a mapping $X\colon M \to \mathrm{T}M$ such that $\tau_M \circ X = \mathrm{Id}_M$) is a *vector field* on M. The set of all vector fields over M is denoted by $\mathfrak{X}(M)$. The action of a vector field X on a function $f \in \mathrm{C}^\infty(M)$ will we written $X \cdot f$ (sometimes $\mathcal{L}_X F$). Let $f\colon M \to N$ be a map. We say that two vector fields $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ are f-related if $\mathrm{T}f \circ X = Y \circ f$, and we write $X \sim_f Y$.

An *integral curve* of a vector field $X \in \mathfrak{X}(M)$ is a path γ on M such that

$$\dot\gamma = X \circ \gamma.$$

In local coordinates, this equations reads as $\dot\gamma^i(t) = X^i(\gamma^1(t), \ldots, \gamma^m(t))$, that is, a system of first-order ordinary differential equations. For each $p \in M$, denote by I_p the biggest interval for which there exist an integral curve of X with $\gamma(0) = p$ (we say that γ is a *maximal integral curve*). We can define the set $\mathcal{D}_X = \{(t,p) \in \mathbf{R} \times M \mid t \in I_p\}$. The map $F_X\colon \mathcal{D}_X \to M$, defined as $F_X(t,p) = \gamma(t)$, where γ is the maximal integral curve of X with $\gamma(0) = p$, is called the *flow* of the vector field X. We also write $F_X^t(p) = F_X(t,p)$. Flows of vector fields have the following properties: for every t the maps F_X^t are diffeomorphisms between its domain and image, $F_X^0 = \mathrm{Id}_M$ and $F_X^t \circ F_X^s = F_X^{t+s}$ where the equation is well-defined.

The dual concept of vector field is the *one-form*, a section $\alpha\colon M \to \mathrm{T}^*M$ of the cotangent bundle. The set of all one-forms over M is denoted by $\Omega^1(M)$. We have a pairing $\langle\cdot,\cdot\rangle\colon \mathfrak{X}(M) \times \Omega^1(M) \to \mathrm{C}^\infty(M)$ defined by $\langle X, \alpha\rangle(p) = \alpha(p) \cdot X(p)$.

In general, we can consider *tensor fields* of contravariant order r and covariant order s (or (r,s)-tensor fields), which are sections of the bundle $\bigotimes_s^r \mathrm{T}M \to M$. At a point $p \in M$, the fibre of this bundle is $\mathrm{T}_p^*M \otimes \overset{s)}{\cdots} \otimes \mathrm{T}_p^*M \otimes \mathrm{T}_pM \otimes \overset{r)}{\cdots} \otimes \mathrm{T}_pM$. The set of skew-symmetric $(0,k)$-tensor fields is denoted by $\Omega^k(M)$ (note that this is consistent with the definition of $\Omega^1(M)$). The elements of $\Omega^k(M)$ are called *k-forms*. For a detailed explanation of the tensor fields and forms, as well as the tensor product $\otimes$, the wedge product $\wedge$, the exterior derivative d, the pullback $f^*\omega$ of differential forms by maps and the Lie derivative $\mathcal{L}_X$ with respect to a vector field X, we refer to [AMR 83].

There are two distinguished types of tensor fields that are the central objects of two branches of differential geometry: the Riemannian geometry and the symplectic geometry. We give here the basic definitions, more detailed introductions to these two vast topics can be found in [Lee 97, Car 92] for Riemannian geometry and [LM 87] for symplectic geometry.

Definition 2.1 *A* Riemannian metric g *on a manifold* M *is a symmetric and definite positive* $(0,2)$*-tensor field on* M*.*

Roughly speaking, a Riemannian metric assigns an inner product to each tangent space T_pM. The pair (M,g) is called a *Riemmanian manifold.* A Riemannian manifold (M,g) is provided with the so-called musical isomorphisms between $\mathrm{T}M$ and T^*M,

$$^\flat\colon \mathrm{T}M \to \mathrm{T}^*M \quad \text{and} \quad ^\sharp\colon \mathrm{T}^*M \to \mathrm{T}M,$$

defined as $v^\flat = g(v,\cdot)$ and $^\sharp$ the inverse of $^\flat$.

Another geometrical object associated with a Riemannian metric g is the Levi–Civita connection ∇^g. It is the unique affine connection (see [KN 63] for an explanation of affine connections and its related concepts, such as covariant derivative and torsion) that is Riemannian ($\nabla^g g = 0$) an torsion-free.

More generally, a symmetric $(0,2)$-tensor field which is nondegenerate is called a *pseudo-Riemmanian metric.* The signature of a pseudo-Riemannian metric is the signature (p,q) of the bilinear form that the metric induces on any tangent space (it can be seen that it is fixed on every connected component), that is, the number of positive and negative eigenvalues. Thus, a Riemannian metric has signature $(m,0)$. A pseudo-Riemannian metric with signature $(1,n)$ is called *Lorentzian metric.*

Definition 2.2 *A* symplectic form ω *on a manifold* M *is a nondegenerate closed* 2*-form on* M*.*

The pair (M,ω) is called a *symplectic manifold.* It can be seen that, since a symplectic form is nondegenerate, the dimension of every symplectic manifold is even.

Like Riemannian manifolds, a symplectic manifold has an isomorphism $\widehat{\omega}\colon \mathrm{T}M \to \mathrm{T}^*M$, defined by $\widehat{\omega}(v) = \omega(v,\cdot)$. We also use the notations $X^\flat = \widehat{\omega}\circ X$ and $\alpha^\sharp = \widehat{\omega}^{-1}\circ\alpha$, for $X \in \mathfrak{X}(M)$ and $\alpha \in \Omega^1(M)$.

Darboux's theorem provides us a nice local description of a symplectic manifold (M,ω) of dimension $2r$. It states that around every point $m \in M$, there exist local coordinates (q^i, p_i), $1 \le i \le r$ such that ω has the local expression

$$\omega = \sum_{i=1}^{r} \mathrm{d}q^i \wedge \mathrm{d}p_i.$$

These coordinate are called *symplectic coordinates*.

Given a function $H \in \mathrm{C}^\infty(M)$, the corresponding *Hamiltonian vector field* is $X_H = \mathrm{d}H^\sharp$. It is the unique vector field on M such that

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

The triple (M, ω, H) is called a *Hamiltonian system* and H is the *hamiltonian* function of the system. The solutions of the Hamiltonian system are the integral curves of the Hamiltonian vector field X_H. In symplectic coordinates (q^i, p_i), a curve $\gamma(t) = (q^i(t), p_i(t))$ is a solution of the Hamiltonian system if and only if satisfy the equations

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q^i},$$

which are called *Hamilton equations*.

If a 2-form $\omega \in \Omega^2(M)$ is closed but degenerate is called *presymplectic*, and the pair (M, ω) is a *presymplectic manifold*. Given a one-form α in M, the triple (M, ω, α) is said to be a *presymplectic (dynamical) system* and has the associated equation

$$i_X\omega = \alpha,$$

for $X \in \mathfrak{X}(M)$. Thus, a Hamiltonian system is a special case of presymplectic system.

A generalization of the concept of symplectic manifold is that of Poisson manifold. Here we only give the definition, we refer to [Vai 94] for a detailed discussion.

Definition 2.3 *A* Poisson bracket *on a manifold M is a mapping $\{\cdot,\cdot\}\colon \mathrm{C}^\infty(M) \times \mathrm{C}^\infty(M) \to \mathrm{C}^\infty(M)$ with the following properties:*

1. *bilinearity over* $\mathbf{R}$*,*
2. *skew-symmetry,*
3. *the Leibniz rule: $\{f, gh\} = \{f, g\}h + g\{f, h\}$, for all f, g and $h \in \mathrm{C}^\infty(M)$, and*
4. *the Jacobi identity: $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$, for all f, g and $h \in \mathrm{C}^\infty(M)$.*

A *Poisson manifold* is s manifold endowed with a Poisson bracket. For example, any symplectic manifold (M, ω) is a Poisson manifold with Poisson bracket $\{f, g\} = \omega(X_f, X_g)$. More generally, a mapping $\{\cdot,\cdot\}\colon \mathrm{C}^\infty(M) \times \mathrm{C}^\infty(M) \to \mathrm{C}^\infty(M)$ that satisfies the first three conditions of definition 2.3 is called an *almost-Poisson bracket* and M an *almost-Poisson manifold*.

An (almost-)Poisson manifold carries an antisymmetric $(2,0)$-tensor field (or bivector field) Λ, characterized by the equation $\{f, g\} = \Lambda(\mathrm{d}f, \mathrm{d}g)$. If the manifold is a Poisson manifold, then the bivector field Λ is nondegenerate if and only if M can be endowed with a symplectic form ω that induces the Poisson bracket $\{\cdot,\cdot\}$.

2.3 Submanifolds

A *submanifold* is a subset of a manifold which is a manifold itself, with a certain compatibility between both manifold structures. The precise definition is as follows.

Let N be a manifold of dimension n and M a subset of N. We say that M is a *regular submanifold* (or simply *submanifold*) of N of dimension m if for every $p \in M$ there exists a local chart (U, Φ) of N around p such that $M \cap U = \Phi^{-1}(\mathbf{R}^m \times \{0\})$. This charts are said to be *adapted to* M. Furthermore, local charts on M of the form $(M \cap U, \Phi_m)$ defined from local charts $(U, \Phi = (\Phi_m, \Phi_{n-m}))$ of N adapted to M constitute an atlas that provides M with a differentiable structure.

An *immersed submanifold* is the image $F(M)$ of an injective immersion $F: M \to N$, with the quotient topology determined by the bijection $F|_M: M \to F(M)$. However, the image $F(M)$, as a subset of N, also has an induced topology. If the two topologies coincide, that is, M is homeomorphic to its image $F(M)$, then F is called an *embedding*. It turns out that the image $F(M)$ of an embedding is a regular submanifold and, conversely, if $M \subset N$ is a regular submanifold, the inclusion $M \hookrightarrow N$ is an embedding.

Now, the following theorem gives alternative definitions of regular submanifolds:

Theorem 2.4 *Let N be a manifold of dimension n, M a subset of N and m a natural number with $0 \leq m \leq n$. The following properties are equivalent:*

i) M is a regular submanifold of N of dimension m.

ii) For every $p \in M$ there exists an open neighborhood $U \subset N$ of p and a submersion $g: U \to \mathbf{R}^{n-m}$ such that $M \cap U = g^{-1}(0)$.

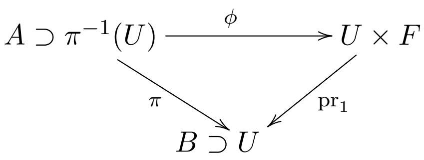
iii) For every $p \in M$ there exist an open neighborhood $U \subset N$ of p and an injective immersion $f: V \to N$, where V is an open subset of $\mathbf{R}^m$, such that $M \cap U = f(V)$.

Note that condition *ii*) of the theorem can be written in terms of the components of g, that is: there exist $n-m$ functions $g_i: U \to \mathbf{R}$ such that $\mathrm{d}g_i(p)$ are linearly independent and $M \cap U = \{x \in U \mid g_1(x) = \cdots = g_{n-m}(x) = 0\}$. Then, the annihilator of T_pM (as a linear subspace of T_pN) is the subspace of T^*_pN generated by the covectors $(\mathrm{d}g_i)_p$, with $1 \leq i \leq n-m$.

2.4 Bundles

The tangent and cotangent bundles are two examples of *bundles*, a geometric structure which generalizes the concept of pair of manifolds and maps. We will mainly follow here the exposition by Saunders [Sau 89].

A *bundle* is a quadruple (A, π, B, F), where A, B, and F are manifolds and $\pi: A \to B$ is a surjective submersion such that for every $b \in B$ there exist an open neighborhood U of b and a diffeomorphism $\phi: \pi^{-1}(U) \to U \times F$ satisfying the condition $\mathrm{pr}_1 \circ \phi = \pi|_{\pi^{-1}(U)}$, where pr_1 is the projection to the first factor. This condition says that the following diagram is commutative:



A is called the *total space*, π the *projection*, B the *base space*, F the *typical fibre* and the diffeomorphisms ϕ are called *local trivializations of* π. For each point $b \in B$, the subset $\pi^{-1}(b)$ is called the *fibre over* b and is denoted by A_b.

Note that all the fibres A_b are diffeomorphic to the typical fibre. We usually refer to a bundle (A, π, B, F) simply as π (or A if the projection is understood) and we will sometimes use a subscript in the notation of an element of the total space A, as for instance a_b, to indicate that it belongs to the fibre A_b.

In the case that the total space is the cartesian product of the base space and the typical fibre, and the projection is exactly $\mathrm{pr}_1 \colon B \times F \to B$, we say that the bundle is a *trivial bundle*.

Suppose that $\dim B = m$ and $\dim F = n$ (so $\dim A = m + n$). A coordinate system $\varphi \colon U \to \mathbf{R}^{m+n}$ on an open set $U \subset A$ is called an *adapted coordinate system* or *fibred chart* if, whenever two points a_1, a_2 belong to the same fibre, then $\mathrm{pr}_1 \circ \varphi(a_1) = \mathrm{pr}_1 \circ \varphi(a_2)$ (where $\mathrm{pr}_1 \colon \mathbf{R}^{m+n} \to \mathbf{R}^m$). We always can construct adapted coordinates systems in the following way. Let $\phi \colon \pi^{-1}(U) \to U \times F$ a local trivialization of π, with U the domain of a local chart $(U, \psi = (x^i))$ of B, and let $(W, \chi = (u^\alpha))$ be a local chart of F. Then $(\psi \times \chi) \circ \phi|_{\phi^{-1}(U \times W)}$ is a fibred chart and we adopt the notation (x^i, u^α) for its components. The local expression of π in adapted coordinates is simply $\pi(x, y) = x$.

A *section* of a bundle (A, π, B, F) is a map $s \colon B \to A$ such that $\pi \circ s = \mathrm{id}_B$. The set of sections of π is denoted by $\mathrm{Sec}(\pi)$. Then we see that a section of a trivial bundle $\mathrm{pr}_1 \colon B \times F \to B$ is indeed the graph of a function from B to F, so the sections of a trivial bundle correspond to the maps from B to F. There is also the concept of *local section*, a map $s \colon U \to A$, where U is an open subset of B, such that $\pi \circ s = \mathrm{id}_U$. The set of local sections with domain U is denoted by $\mathrm{Sec}_U(\pi)$. An extension of the concept of section is that of *section along a map*. If $f \colon C \to B$ is a map from a manifold C to the base space B, a section along f is a map $g \colon C \to A$ such that $\pi \circ g = f$. Note that if we set $f = \mathrm{id}_B$ we recover the previous definition of section. The set of sections along f is denoted by $\mathrm{Sec}(f)$. An example of section along a map is given by the derivative $\dot\gamma$ of a path $\gamma \colon I \to M$ on a manifold M, in this case $\dot\gamma$ is a section along γ of the tangent bundle $\tau_M \colon \mathrm{T}M \to M$. A section X of τ_M along a map $f \colon N \to M$ is also called *vector field along a map*. It has a corresponding differential operator $\mathrm{d}_X \colon \mathrm{C}^\infty(M) \to \mathrm{C}^\infty(N)$ defined as $(\mathrm{d}_X f)(p) = X(p) \cdot f$. In an analogous way, for each type of tensor bundle there are tensor fields along maps.

Let $\pi\colon A \to B$ and $\pi'\colon A' \to B'$ be two bundles. A *bundle morphism* from π to π' is a pair of maps (f, f_o) where $f\colon A \to A'$, $f_o\colon B \to B'$ and $\pi' \circ f = f_o \circ \pi$:

$$\begin{array}{ccc} A & \xrightarrow{f} & A' \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi'} \\ B & \xrightarrow[f_o]{} & B' \end{array}$$

The map f_o is called the *projection* of f and we say that f is a *morphism over* f_o. The important point is that f maps fibres to fibres (i.e., $f(A_b) \subset A'_{f_o(b)}$), so in fact f_o is determined by f. Reciprocally, every $f\colon A \to A'$ that maps fibres to fibres determines a unique projection $f_o\colon B \to B'$ and hence a bundle morphism. We denote the restrictions to the fibres $f|_{A_b}$ simply by f_b. Using adapted coordinates (x^i, u^α) on A and (y^j, v^β) on A', a map $f\colon A \to A'$ has a local description $f = (f^j(x^i, u^\alpha), f^\beta(x^i, u^\alpha))$. Then f induces a bundle morphism if and only if the components f^j only depend on the base coordinates x^i of B (that is, $f = (f^j(x^i), f^\beta(x^i, u^\alpha)))$, and in this case the projection $f_o\colon B \to B'$ has the local expression $f_o = (f^j(x^i))$. An important particular case is when $B = B'$ and $f_o = \mathrm{id}_B$, then f is simply called a *morphism* between bundles over B.

If (A, π, B, F) is a bundle and $f\colon C \to B$ a map from a manifold C to the base B, we can define a new bundle $(f^*(A), f^*(\pi), C, F)$ called the *pull-back of* π *by* f. The total space $f^*(A)$ (also denoted by $A \times_f C$) is the manifold

$$f^*(A) = A \times_f C = \{(a, c) \in A \times C \mid \pi(a) = f(c)\}$$

and the projection $f^*(\pi)$ is simply the projection to C: $f^*(\pi)(a, c) = c$. It is easy to see that π and $f^*(\pi)$ share the same typical fibre, because $f^*(A)_c = \{(a, c) \in A \times C \mid \pi(a) = f(c)\} \cong A_{f(c)}$, where $\cong$ means that the two fibres are diffeomorphic. It is clear that the sections of the pull-back $f^*(\pi)$ are in bijective correspondence with the sections along f. Similarly, given a bundle $\rho\colon D \to C$ with basis C, the bundle morphisms $F\colon D \to A$ over f are in bijective correspondence with the morphisms between ρ and the pull-back bundle $f^*(\pi)$. A particular case of pull-back occurs when $C \hookrightarrow B$ is a submanifold, in this case we denote the pull-back bundle by $\pi|_C\colon A|_C \to C$ and we call it the *restriction* of A to C.

Vector bundles

A bundle (E, π, B, F) whose fibres (and the typical fibre F) are vector spaces of dimension n, and such that it has local trivializations which, restricted to each fibre, are linear isomorphisms $\phi|_{E_b}\colon E_b \to \{b\} \times F$ between the fibres and F, is called a *vector bundle*. We will always use this kind of trivializations for vector bundles. The *rank* of a vector bundle is the dimension of its fibres.

If (u_α) is a basis of the vector space F, the components (u^α) with respect to this basis are coordinates on F which, together with a local chart (U, x^i) of B and local trivializations, induce an adapted coordinate system (x^i, u^α) on E as we described earlier. These adapted coordinates systems are called *vector bundle coordinates systems* and we will always use them when dealing with vector bundles. A *local frame* (on an open set $U \subset B$) of a vector bundle of rank k is a family of sections $(s_\alpha \colon U \to E)$ such that, for every $b \in U$, $(s_\alpha(b))$ is a basis of the vector space E_b. A vector bundle chart $(U, (x^i, u^\alpha))$ defines a local frame (s_β) on U by $u^\alpha(s_\beta(b)) = \delta^\alpha_\beta$ for every $b \in U$.

Since the fibres of a vector bundle π are vector spaces, the set $\mathrm{Sec}(\pi)$ (or $\mathrm{Sec}_U(\pi)$) of (local) sections is a $\mathrm{C}^\infty(B)$-module ($\mathrm{C}^\infty(U)$-module) under the pointwise operations $(s_1 + s_2)(b) = s_1(b) + s_2(b)$ and $(fs)(b) = f(b)s(b)$. Therefore, a local frame on U is just a basis of the $\mathrm{C}^\infty(U)$-module $\mathrm{Sec}_U(\pi)$.

The tangent and cotangent bundles of a manifold M are both vector bundles. The natural coordinates $(x^i, \dot{x}^i)$ and (x^i, p_i) are vector bundle coordinates, with corresponding local frames $(\frac{\partial}{\partial x^i})$ and $(\mathrm{d}x^i)$ respectively.

Let us point out that the tangent space $\mathrm{T}A$ of the total space of a bundle $\pi \colon A \to B$ is the total space of two different vector bundles: the tangent bundle $\tau_A \colon \mathrm{T}A \to A$ and the bundle defined by the tangent map $\mathrm{T}\pi \colon \mathrm{T}A \to \mathrm{T}B$. In case that π is a tangent bundle $\tau_M \colon \mathrm{T}M \to M$ itself, we see that $\mathrm{T}(\mathrm{T}M)$ has two structures as a vector bundle over $\mathrm{T}M$; the tangent bundle $\tau_{\mathrm{T}M}$ and $\mathrm{T}\tau_M$.

Let $\pi \colon E \to B$ be a vector bundle and $E' \subset E$ a submanifold such that $\pi|_{E'} \colon E' \to \pi(E')$ is itself a vector bundle under the restriction of the vector addition and scalar multiplication of the fibres of π to the fibres of $\pi|_{E'}$. We say that $\pi|_{E'}$ is a *vector subbundle* of π, although sometimes we also refer to the submanifold E' as the vector subbundle. It will be usual that $\pi(E') = B$, so the vector subbundle has the same basis as the vector bundle. The notation $\pi|_{E'}$ is also used for the restriction of a bundle to a submanifold of the basis, but the context will clarify the real meaning in this case.

A *vector bundle morphism* (f, f_o) between two vector bundles $\pi \colon E \to B$ and $\pi' \colon E' \to B'$ is a bundle morphism which is linear in each fibre (i.e., for each $b \in B$, $f_b \colon E_b \to E'_{f_o(b)}$ is a linear map). We say that f has constant rank if the rank of f_b is the same for all $b \in B$. In vector bundle coordinates, a vector bundle morphism reads as

$$f(x^i, u^\alpha) = (f^j_o(x^i), f^\beta_\alpha(x^i)u^\alpha).$$

The matrix of functions (f^β_α) is called the *local matrix representation* of the vector bundle morphism f.

If $f \colon M \to N$ is a map between manifolds, then the tangent map $\mathrm{T}f \colon \mathrm{T}M \to \mathrm{T}N$ ia a well known example of vector bundle morphism.

The *kernel* of a vector bundle morphism $f \colon E \to E'$ is, the subset

$$\mathrm{Ker}\, f = \{e \in E \mid f(e) = 0 \in E'_{f_o(\pi(e))}\} = \bigcup_{b \in B} \mathrm{Ker}\, f_b \subset E.$$

The following important property holds: if f has constant rank, $\operatorname{Ker} f$ is a vector subbundle of π and $\operatorname{Im} f$ is a vector subbundle of π'.

The usual operations that can be performed with vector spaces induce operations among vector bundles with the same basis. For example, if we have two vector bundles $\pi: E \to B$ and $\pi': E' \to B$, $\pi \oplus \pi': E \oplus E' \to B$ is the vector bundle whose fibre $(E \oplus E')_b$ is $E_b \oplus E'_b$; it is called the *Whitney sum*. Similarly, we have the tensor product $\pi \otimes \pi': E \otimes E' \to B$, which generalizes to the tensor product of a finite number of vector bundles, the alternating and symmetric products $\Lambda^r \pi: \bigwedge^r E \to B$ and $\mathsf{S}^r \pi: S^r E \to B$, the bundle of homomorphisms $\operatorname{Hom}(\pi, \pi'): \operatorname{Hom}(E, E') \to B$ and the *dual bundle* $\pi^*: E^* \to B$, where $E^* = \operatorname{Hom}(E, B \times \mathbf{R})$. For every system of vector bundle coordinates (x^i, u^α) on π we can define a system of dual coordinates (x^i, u_α) on π^* in such a way that the corresponding local frames s_α of π and s^α of π^* are dual: $s^\alpha(b)(s_\beta(b)) = \delta^\alpha_\beta$. For instance, the tangent and cotangent bundle of a manifold are dual bundles, and the standard coordinates $(x^i, \dot{x}^i)$ and (x^i, p_i) are dual coordinates, since the corresponding local frames $(\frac{\partial}{\partial x^i})$ and $(\mathrm{d}x^i)$ are dual frames. Also, if E' is a vector subbundle of E (with the same basis), we can define the *quotient bundle* E/E' as the vector bundle over B with fibres $(E/E')_b = E_b/E'_b$.

Finally, the operations with linear transformations between vector spaces can also be extended to morphisms of vector bundles over the same basis. In particular, a vector bundle morphism $f: E \to E'$ has a transposed morphism between the dual bundles, we denote it by ${}^t f: (E')^* \to E^*$.

Let $\pi: E \to B$ be a vector bundle and $f: E \to \mathbf{R}$ a function on the total space. For each $b \in B$, consider the restriction of f to the fibre E_b, $f_b: E_b \to \mathbf{R}$. It is a function on a vector space, so for every point e_b of the fibre, the derivative of f_b at e_b is a linear form $\mathrm{D}f_b(e_b) \in \operatorname{Hom}(E_b, \mathbf{R}) = E^*_b$. Therefore, we can define the map

$$\begin{array}{rccc} \mathcal{F}f: & E & \to & E^* \\ & e_b & \mapsto & \mathrm{D}f_b(e_b) \end{array} \tag{2.1}$$

This map is a morphism of bundles over B and we call it the *fibre derivative* of f. In adapted coordinate systems (dual, of course), its local expression is $\mathcal{F}f(x, u) = (x^i, \frac{\partial f}{\partial u^\alpha}(x, u))$. Similarly, we can define higher order derivatives. The *fibre hessian* of a function $f: E \to \mathbf{R}$ is the morphism $\mathcal{F}^2 f: E \to E^* \otimes E^*$ defined as $\mathcal{F}^2 f(e_b) = \mathrm{D}^2 f_b(e_b)$, where $\mathrm{D}^2 f_b(e_b): E_b \times E_b \to \mathbf{R}$ is the second derivative of f_b at e_b, hence a symmetric bilinear form. Its local expression is $\mathcal{F}^2 f(x, u) = (x^i, \frac{\partial f}{\partial u^\alpha \partial u^\beta}(x, u))$

The tangent bundle of the total space A of any bundle $\pi: A \to B$ has a distinguished vector subbundle, namely the *vertical bundle to* π, whose total space, denoted by $\mathrm{V}\pi$, is the kernel of $\mathrm{T}\pi$:

$$\mathrm{V}\pi = \{v \in \mathrm{T}A \mid \mathrm{T}\pi(v) = 0\}.$$

Now let us define the vertical lift, an important operation on vector bundles. A fibre E_b of a vector bundle $\pi: E \to B$ is a vector space, therefore for each $e_b \in E_b$ there exists an isomorphism $\lambda_{e_b}\colon E_b \to \mathrm{T}_{e_b}(E_b)$, defined by $\lambda_{e_b}(e'_b) = [t \mapsto e_b+te'_b] = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}(e_b+te'_b)$. The vector $\lambda_{e_b}(e'_b)$ is called the *vertical lift* of e'_b over e_b. Note that $\mathrm{T}_{e_b}(E_b)$, which is a linear subspace of $\mathrm{T}_{e_b}E$, is precisely the fibre $\mathrm{V}_{e_b}\pi$ of the vertical bundle. Gluing all these isomorphisms together we obtain an isomorphism

$$\begin{array}{rccc} \lambda_E\colon & E\times_B E & \to & \mathrm{V}\pi \\ & (e_b, e'_b) & \mapsto & \lambda_E(e_b,e'_b) = \lambda_{e_b}(e'_b) \end{array}.$$

Correspondingly to the vertical bundle to $\pi\colon A \to B$, there is a distinguished vector subbundle of the contangent bundle T^*A: the annihilator of $\mathrm{V}\pi$. Its sections are called *π-semibasic 1-forms*. We denote the bundle by $\mathrm{Sb}\,\pi$ (or $\mathrm{Sb}\,A$ when the projection is understood) and the set of π-semibasic 1-forms by $\Omega^1_0(\pi)$. We note that this bundle is isomorphic to the pullback of the cotangent bundle T^*B of the basis by π, that is $A \times_\pi \mathrm{T}^*B$.

A *connection* Γ on a bundle $\pi\colon A \to B$ is a vector-valued semibasic 1-form (that is, a section of the bundle $\mathrm{Sb}\,\pi \otimes \mathrm{T}A$) such that $i_\sigma\Gamma = \sigma$ for every $\sigma \in \Omega^1_0(\pi)$. In coordinates, a connection Γ has the expression

$$\Gamma = \mathrm{d}x^i \otimes \left(\frac{\partial}{\partial x^i} + \Gamma^\alpha_i \frac{\partial}{\partial u^\alpha}\right).$$

A connection Γ determines a subbundle H_Γ of the tangent bundle $\mathrm{T}A$, called the *horizontal bundle* by Γ, defined as the image of $\mathrm{T}A$ by Γ: $(H_\Gamma\pi)_a = \{\Gamma_a(v) \,|\, v \in \mathrm{T}_aA\}$. The tangent bundle of A can be written as the direct sum of the vertical bundle and the horizontal bundle by Γ: $\mathrm{T}A = \mathrm{V}\pi \oplus H_\Gamma$, so every connection induces an splitting of $\mathrm{T}A$. Conversely, if H is a vector subbundle of $\mathrm{T}A$ such that $\mathrm{T}A = \mathrm{V}\pi \oplus H$, then there exists a unique connection Γ such that $H = H_\Gamma$. Therefore, it is equivalent to give a connection on A or a splitting of $\mathrm{T}A$.

Affine bundles

In an analogous way as vector spaces give rise to vector bundles, we can define the notion of affine bundle. Let (E, π, B, V) be a vector bundle. A bundle (A, ρ, B, F) such that for each $b \in N$, the fibre A_b is an affine spaces modelled on E_b (equivalently, the typical fibre F is an affine space modelled on V), and such that has local trivializations which, restricted to each fibre, are affine isomorphisms $\phi|_{A_b}\colon A_b \to \{b\}\times F$ between the fibres and F is called an *affine bundle* modelled on π. We will always use this type of trivializations when dealing with affine bundles. Naturally, there is an action $A\times_B E \to A$ defined as $(a_b, e_b) \mapsto a_b + e_b$.

A vector bundle coordinate system (x^i, u^α) on E and a section $z \in \mathrm{Sec}(\rho)$ of the affine bundle A induce an adapted coordinate system (x^i, a^α) on A, where a^α are the affine functions $a^\alpha(a_b) = u^\alpha(a_b - z(b))$. We will always use these special coordinates systems on affine bundles, they are called *affine bundle coordinate systems*.

Every vector bundle can be considered an affine bundle modelled on itself, since each fibre it is a vector space, that can be considered an affine space modelled on itself. Furthermore, let $\rho: A \to B$ be an affine bundle modelled on a vector bundle $\pi: E \to B$. Then every global section $z \in \mathrm{Sec}(\rho)$ determines a vector bundle structure on ρ, since for every $b \in B$, the section z distinguishes a point $z(b) \in A_b$ and induces an isomorphism $A_b \cong E_b$ by the bijection $a_b \leftrightarrow a_b - z(b)$.

Let $\rho: A \to B$ and $\rho': A' \to B'$ be two affine bundles. A bundle morphism (f, f_o) from ρ to ρ' is an *affine bundle morphism* if it is affine in each fibre (that is, for each $b \in B$, $f_b: A_b \to A'_{f_o(b)}$ is an affine map). It has a linear part, which is a vector bundle morphism between the vector bundles over which are modelled ρ and ρ', defined by taking the linear part of the affine map in every fibre. In affine bundle coordinates, an affine bundle morphism (f, f_o) is written as

$$f(x^i, a^\alpha) = (f_o(x^i), f_\alpha^\beta(x^i)a^\alpha + f^\beta(x^i)),$$

where $(f_\alpha^\beta(x^i))$ is the local matrix representation of the linear part of the affine bundle morphism.

2.5 Geometry of the tangent bundle

Let M be a manifold of dimension m. Here we define some canonical geometric objects associated with the tangent bundle $\mathrm{T}M$ that will be useful in the geometric description of physical systems. We refer to [Cra 83, LR 89] for a more detailed discussion. The vertical lift of the tangent bundle, $\lambda_{\mathrm{T}M}: \mathrm{T}M \times_M \mathrm{T}M \to \mathrm{T}(\mathrm{T}M)$, will prove to be a very useful tool.

Let us start with the *Liouville vector field* $\Delta \in \mathfrak{X}(\mathrm{T}M)$, defined as

$$\Delta(v) = \lambda_{\mathrm{T}M}(v, v). \tag{2.2}$$

It can also be defined as the infinitesimal generator of the one-parameter group of dilations, that is, the flow of Δ is $F_\Delta^t(v) = e^t v$. It can be seen that in any natural system of coordinates $(x^i, \dot{x}^i)$, the local expression of the Liouville vector field is $\Delta = \dot{x}^i \frac{\partial}{\partial \dot{x}^i}$. It is worth noting here that a Liouville vector field $\Delta_E \in \mathfrak{X}(\mathrm{T}E)$ can be equally defined in any vector bundle $\pi: E \to B$.

One basic property of the Liouville vector field is the way it acts upon homogeneous functions. If $f \in \mathrm{C}^\infty(\mathrm{T}M)$ is homogeneous of degree k (that is, $f(\lambda v) = \lambda^k f(v)$ for $\lambda \in \mathbf{R}$) then $\Delta \cdot f = kf$.

Another canonical geometric object is the *vertical endomorphism* S of $\mathrm{T}M$. It is the $(1,1)$-tensor field on $\mathrm{T}M$ defined by

$$S(w_v) = \lambda_{\mathrm{T}M}(v, \mathrm{T}_v\tau_M(w)), \tag{2.3}$$

where $w_v \in \mathrm{T}_v\mathrm{T}M$. The local expression of the vertical endomorphism is $S = \frac{\partial}{\partial \dot{x}^i} \otimes \mathrm{d}x^i$. It follows from its very definition of S that both its kernel and image is the vertical bundle $\mathrm{V}\tau_M$. When we consider the tensor field acting on $\mathrm{T}^*(\mathrm{T}M)$, we will denote it by ${}^tS$, since it is the transposed of the endomorphism S.

Every vector field X on M defines two vector fields on $\mathrm{T}M$. The *vertical lift* of $X \in \mathfrak{X}(M)$ is the vector field $X^V \in \mathfrak{X}(\mathrm{T}M)$ defined as

$$X^V(u_p) = \lambda_{\mathrm{T}M}(u_p, X(p)).$$

Locally, if $X = X^i \frac{\partial}{\partial x^i}$ then $X^V = X^i \frac{\partial}{\partial \dot{x}^i}$, where we see that the vertical lift is indeed a vertical vector field. The other lift is the *complete lift* or *canonical lift*; it is the vector field $X^T \in \mathfrak{X}(\mathrm{T}M)$ whose flow is constituted by the tangent extensions of the flow of X, that is,

$$F^t_{X^T} = \mathrm{T}(F^t_X).$$

The local expression of the complete lift is $X^T = X^i \frac{\partial}{\partial x^i} + (\frac{\partial X^i}{\partial x^j})\dot{x}^j \frac{\partial}{\partial \dot{x}^i}$.

There is another distinguished class of vector fields on $\mathrm{T}M$. A vector field $X \in \mathfrak{X}(\mathrm{T}M)$ is called a *second order differential equation* (or SODE for short) if

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

Another characterization is that $\mathrm{T}(\tau_M) \circ X = \mathrm{Id}_{\mathrm{T}M}$, so a SODE is a section of the two structures of vector bundle that has $\mathrm{T}(\mathrm{T}M)$ over $\mathrm{T}M$.

The reason of the nomenclature becomes clear when we examine the local expression of a SODE, which is $X = \dot{x}^i \frac{\partial}{\partial x^i} + f^i(x, \dot{x}) \frac{\partial}{\partial \dot{x}^i}$, Hence, an integral curve of the SODE X is the solution of the system of differential equations

$$\left\{ \begin{array}{l} \frac{\mathrm{d}x^i}{\mathrm{d}t} = \dot{x}^i \\ \frac{\mathrm{d}\dot{x}^i}{\mathrm{d}t} = f^i(x, \dot{x}) \end{array} \right. ,$$

which, as it is well known, is equivalent to a system of m second-order differential equations in the variables x^i: $\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = f^i(x, \frac{\mathrm{d}x^i}{\mathrm{d}t})$. Therefore, an integral curve $\gamma\colon I \to \mathrm{T}M$ of X will be the derivative of the curve $\tau_M \circ \gamma\colon I \to M$.

2.6 Geometry of the cotangent bundle

Now we recall that the cotangent bundle T^*M of any manifold M has a canonical symplectic structure that can be constructed as follows.

Let θ_M be the 1-form on T^*M defined by its action on vectors of $\mathrm{T}(\mathrm{T}^*M)$ as

$$\langle \theta_M(\alpha_p), v_{\alpha_p} \rangle = \langle \alpha_p, \mathrm{T}\pi_M(v_{\alpha_p}) \rangle,$$

for every $p \in M$, $\alpha_p \in \mathrm{T}^*_pM$ and $v_{\alpha_p} \in \mathrm{T}_{\alpha_p}(\mathrm{T}^*M)$. We call θ_M the *Liouville 1-form.* In any natural system of coordinates (x^i, p_i) of T^*M, the Liouville 1-form has local expression

$$\theta_M = p_i \mathrm{d}x^i,$$

(so it is a π^*-semibasic 1-form). Therefore, the 2-form $\omega_M = -\mathrm{d}\theta_M$ has local expression

$$\omega_M = \mathrm{d}x^i \wedge \mathrm{d}p_i,$$

where we see that it is nondegenerate. Since, by definition, it is also closed, ω_M is a symplectic form of T^*M, called the *canonical symplectic form* of the contangent bundle. We also note that the natural coordinates of T^*M are symplectic coordinates.

2.7 Geometry of jet bundles

Jet bundles are the basic geometric structure to model time-dependent systems and classical field theories. Here we give an introduction to the subject and the most relevant aspects for this dissertation. For details we refer to [Sau 89].

Let $\pi: M \to B$ be a fibre bundle. We say that two sections ϕ and ψ of π are *k-equivalent at the point $b \in B$* if $\phi(b) = \psi(b)$ and, in some adapted coordinate system of the bundle, the local expressions of ϕ and ψ have the same derivatives up to order k.

It turns out that this definition is independent of the choice of the coordinate system and that this relation is an equivalence relation. The equivalence class containing a section ϕ is called *k-jet of ϕ at b* and is denoted by $\mathrm{j}^k_b\phi$.

The *k-th jet manifold of π* is the set of all k-jets, it is denoted by $\mathrm{J}^k\pi$. We have the following canonical projections:

$$\begin{array}{rccc} \pi_k\colon & \mathrm{J}^k\pi & \to & B \\ & \mathrm{j}^k_b\phi & \mapsto & b \end{array} ,$$

$$\begin{array}{rccc} \pi_{k,0}\colon & \mathrm{J}^k\pi & \to & M \\ & \mathrm{j}^k_b\phi & \mapsto & \phi(b) \end{array}$$

and, for $1 \leq l \leq k$,

$$\begin{array}{rccc} \pi_{k,l}\colon & \mathrm{J}^k\pi & \to & \mathrm{J}^l\pi \\ & \mathrm{j}^k_b\phi & \mapsto & \mathrm{j}^l_b\phi \end{array} .$$

Every adapted coordinate system (x^i, u^α) on M induces a coordinate system on $\mathrm{J}^k\pi$. To describe it, first we introduce the multi-index notation. Let n be the dimension of the base B. A *multi-index* is an n-tuple I of natural numbers. We denote by $I(i)$ the i-th component of I, $1 \leq i \leq n$. The length of a multi-index I is $|I| = \sum_{i=1}^n I(i)$. The multi-index 1_i has all its components equal to 0 but the i-th component, which is equal to 1. We can add and subtract two multi-indexes componentwise: $(I \pm J)(i) = I(i) \pm J(i)$, whenever the result is another multi-index. Now, given a coordinate system (x^i, u^α) on M, the induced coordinate system on $\mathrm{J}^k\rho$ is $(x^i, u^\alpha, u^\alpha_I)$, for all the multi-indices I

with $1 \leq |I| \leq k$. The coordinate u^α_I is defined by

$$u^\alpha_I(\mathrm{j}^k_p\phi) = \left(\prod_{i=1}^n \frac{\partial^{I(i)}}{\partial x^i}\right)\Bigg|_p (\phi^\alpha).$$

(For the case $k = 1$, this simplifies to $(x^i, u^\alpha, u^\alpha_i)$.)

Now, given an atlas of adapted charts on M, the corresponding collection of adapted charts on $\mathrm{J}^k\pi$ as described above is an atlas. Therefore, $\mathrm{J}^k\pi$ has a differential manifold structure induced by the fibre bundle structure of π. Moreover, $\pi_k\colon \mathrm{J}^k\pi \to B$, $\pi_{k,0}\colon \mathrm{J}^k\pi \to M$ and $\pi_{k,l}\colon \mathrm{J}^k\pi \to \mathrm{J}^l\pi$ are all fibre bundles.

For any local section ϕ of π with domain $W \subset B$, we can define a local section of π_k called the *k-th prolongation of ϕ*. This section $\mathrm{j}^k\phi\colon W \to \mathrm{J}^k M$ is given by

$$\mathrm{j}^k\phi(b) = \mathrm{j}^k_b\phi.$$

Also, any bundle morphism (f, f_o) between bundles $\pi\colon M \to B$ and $\pi'\colon M' \to B'$ such that f_o is a diffeomorphism (in particular, when $B = B'$ and $f_o = \mathrm{Id}_B$) can be prolonged to a map $\mathrm{j}^k f\colon \mathrm{J}^k\pi \to \mathrm{J}^k\pi'$ defined as

$$\mathrm{j}^k f(\mathrm{j}^k_b\phi) = \mathrm{j}^k_{f_o(b)}(\tilde{f}(\phi)),$$

where $\tilde{f}(\phi)$ is the section $f \circ \phi \circ f_o^{-1}$ of π'. It can be seen that both $(\mathrm{j}^k f, f)$ and $(\mathrm{j}^k f, f_o)$ are bundle morphisms.

A significant fact is the natural affine bundle structure of the bundles $\pi_{k,k-1}$ for $k \geq 1$. Its associated vector bundle is $\pi^*_{k-1}(S^k\mathrm{T}^*B) \otimes \pi^*_{k-1,0}(\mathrm{V}\pi)$. We recall that $S^k\mathrm{T}^*B \to B$ is the bundle of symmetric k-covectors and $\mathrm{V}\pi \subset \mathrm{T}M$ the vertical bundle to π.

A *jet field* is a section Γ of the bundle $\pi_{1,0}\colon \mathrm{J}^1\pi \to M$. The set of jet field are in bijective correspondence with the connections on π [Sau 89], so a jet field Γ induces a splitting $\mathrm{T}M = \mathrm{V}\pi \oplus H_\Gamma$ of the tangent bundle, with projections v_Γ and h_Γ. We will denote by $\tilde{\Gamma}$ the connection induced by a jet field Γ. The coordinate expressions of these objects are:

$$\Gamma(x^i, u^\alpha) = (x^i, u^\alpha, \Gamma^\alpha_i(x^i, u^\alpha)),$$

$$\tilde{\Gamma} = \mathrm{d}x^i \otimes \left(\frac{\partial}{\partial x^i} + \Gamma^\alpha_i \frac{\partial}{\partial u^\alpha}\right),$$

$$v_\Gamma(x^i, u^\alpha; \dot{x}^i, \dot{u}^\alpha) = (x^i, u^\alpha; 0, \dot{u}^\alpha - \dot{x}^i\,\Gamma^\alpha_i(x^i, u^\alpha)),$$

$$h_\Gamma(x^i, u^\alpha; \dot{x}^i, \dot{u}^\alpha) = (x^i, u^\alpha; \dot{x}^i, \dot{x}^i\,\Gamma^\alpha_i(x^i, u^\alpha)).$$

Now we will study with some more detail the case where the base of the fibre bundle is the real line: $B = \mathbf{R}$. These bundles are the appropriate models for time-dependent configuration spaces. In this case, we will denote the bundle by $\rho\colon M \to \mathbf{R}$. We will use the canonical identity coordinate t in $\mathbf{R}$.

Note that the sections of ρ are also paths in M. This fact allows to define the canonical embedding

$$\begin{array}{rccl} \iota\colon & \mathrm{J}^1\rho & \to & \mathrm{T}M \\ & \mathrm{j}^1_t\xi & \mapsto & \dot{\xi}(t) \ . \end{array} \qquad (2.4)$$

With the usual notation (t, q^i, v^i) for the coordinates on $\mathrm{J}^1\rho$ induced by adapted coordinates (t, q^i) on M, the embedding is written as $\iota(t, q^i, v^i) = (t, q^i; 1, v^i)$.

Likewise, the k-jet prolongation of a section of ρ is a path in $\mathrm{J}^k\rho$ and, for each k, there is an embedding

$$\begin{array}{rccl} \iota_k\colon & \mathrm{J}^k\rho & \longrightarrow & \mathrm{T}\mathrm{J}^{k-1}\rho \\ & \mathrm{j}^k_t\xi & \longmapsto & (\mathrm{j}^{k-1}\xi)^{\cdot}(t). \end{array} \qquad (2.5)$$

Note that ι_k is a vector field along $\rho_{k,k-1}\colon \mathrm{J}^k\rho \to \mathrm{J}^{k-1}\rho$, so it has an associated differential operator from $\mathrm{C}^\infty(\mathrm{J}^{k-1}\rho)$ to $\mathrm{C}^\infty(\mathrm{J}^k\rho)$, known as the *total time derivative operator* [CMF 92], and denoted by $\mathbf{T}^{(k)}$.

Using adapted local coordinates $(t, q^i_{(0)}, q^i_{(1)}, \dots, q^i_{(k-1)}, q^i_{(k)})$ on $\mathrm{J}^k\rho$, the embeddings ι_k are written as

$$\iota_k(t, q^i_{(0)}, q^i_{(1)}, \dots, q^i_{(k-1)}, q^i_{(k)}) = (t, q^i_{(0)}, q^i_{(1)}, \dots, q^i_{(k-1)}; 1, q^i_{(1)}, \dots, q^i_{(k)}).$$

The associated total time derivative operator $\mathbf{T}^{(k)}$ acts on a function $f \in \mathrm{C}^\infty(\mathrm{J}^{k-1}\rho)$ as

$$\mathbf{T}^{(k)} f = \frac{\partial f}{\partial t} + \sum_{l=0}^{k} q^i_{(l+1)} \frac{\partial f}{\partial q^i_{(l)}} \in \mathrm{C}^\infty(\mathrm{J}^k\rho).$$

As noted above, the bundle $\mathrm{J}^1\rho \to M$ is an affine bundle. In this particular case $(B = \mathbf{R}, k = 1)$, the associated vector bundle reduces to $\mathrm{V}\rho$. The affine structure becomes clear making use of the canonical embedding ι, since we can view $\mathrm{J}^1\rho$ and $\mathrm{V}\rho$ as an affine and a vector subbundle of $\mathrm{T}M$.

There is a canonical $(1,1)$-tensor field on $\mathrm{J}^1\rho$ that plays a similar role as the vertical endomorphism S of a tangent bundle. It is called the *vertical endomorphism of* $\mathrm{J}^1\rho$ and, by abuse of notation, we will also denote it by S. In local coordinates it is written as

$$S = (\mathrm{d}q^i - v^i \mathrm{d}t) \otimes \frac{\partial}{\partial v^i}. \qquad (2.6)$$

Using this vertical endomorphism we can characterize $\mathrm{J}^2\rho$ (or, more exactly, its image by the embedding ι_2) as the subbundle of $\mathrm{TJ}^1\rho$

$$\iota_2(\mathrm{J}^2\rho) = \{w \in \mathrm{TJ}^1\rho \mid i_w \mathrm{d}t = 1,\ S(w) = 0\}.$$

The local expression of S readily shows that $\mathrm{Im}(S) = \mathrm{V}\rho_{1,0}$. The kernel of S is called the *Cartan distribution on* $\mathrm{J}^1\rho$ and is denoted by $C\rho_{1,0}$. Locally, we can describe $C\rho_{1,0}$ as the distribution generated by the $n+1$ vector fields $\{\frac{\partial}{\partial t} + v^i \frac{\partial}{\partial q^i}, \frac{\partial}{\partial v^i}\}$. To sum up, we have the exact sequence

$$0 \longrightarrow C\rho_{1,0} \longrightarrow \mathrm{T}(\mathrm{J}^1\rho) \xrightarrow{S} \mathrm{V}\rho_{1,0} \longrightarrow 0 \qquad (2.7)$$

In general, the *Cartan distribution on* $\mathrm{J}^{k-1}\rho$ is defined as the distribution generated by the vectors tangent to $(k-1)$-jet prolongations of sections of ρ. We denote it by $C\rho_{k-1,k-2}$, and it is locally generated by the $n+1$ vector fields $\{\frac{\partial}{\partial t} + \sum_{l=0}^{k-2} q^i_{(l+1)} \frac{\partial}{\partial q^i_{(l)}}, \frac{\partial}{\partial q^1_{(k-1)}}, \dots, \frac{\partial}{\partial q^n_{(k-1)}}\}$.

2.8 Higher order tangent bundles

Let M be a manifold. We say that two curves $\gamma, \delta\colon \mathbf{R} \to M$ are *k-tangent at* 0 if $\gamma(0) = \delta(0)$ and, for every function $f \in \mathrm{C}^\infty(M)$, all the derivatives up to order k of the difference $f \circ \gamma - f \circ \delta \in \mathrm{C}^\infty(\mathbf{R})$ vanish at 0. This relation is an equivalence relation, the equivalence classes are called *tangent vectors of order k* and we denote the class of γ by $\gamma_0^{(k)}$. The set of all the equivalence classes is the *tangent bundle of order k*, denoted by $\mathrm{T}^k M$. It is clear from the definition that if $k = 1$ we recover the tangent bundle $\mathrm{T}M$, and $\mathrm{T}^0 M$ is identified with M.

It can be shown that $\mathrm{T}^k M$ has a differential structure inherited from M such that the projections, for $0 \le r < k$,

$$\begin{array}{cccc} \tau_{k,r}\colon & \mathrm{T}^k M & \to & \mathrm{T}^r M \\ & \gamma_0^{(k)} & \mapsto & \gamma_0^{(r)} \end{array}$$

are fibre bundles.

A local system of coordinates (x^i) on M induces coordinates $(x^i, x^i_{(1)}, \dots, x^i_{(k)})$ on $\mathrm{T}^k M$ defined as

$$x^i_{(r)}(\gamma_0^{(k)}) = \frac{\mathrm{d}^r(x^i \circ \gamma)}{\mathrm{d}t^r}(0),$$

for $r \le k$. They are called natural coordinates and in the case of the tangent bundle they are the already known ones.

The *k-th prolongation* of a curve $\gamma\colon I \to M$ is the curve $\gamma^{(k)}\colon I \to \mathrm{T}^k M$ defined as $\gamma^{(k)}(t) = (\gamma_t)_0^{(k)}$, where $\gamma_t(s) = \gamma(t+s)$. It is a section of $\tau_{(k,1)}$ along γ. In local coordinates, the k-th prolongation of a curve $(\gamma^i(t))$ has the expression $(\gamma^i(t), \frac{\mathrm{d}\gamma^i}{\mathrm{d}t}(t), \dots, \frac{\mathrm{d}^k\gamma^i}{\mathrm{d}t^k}(t))$. Obviously we have that the derivative $\dot\gamma$ of a curve is its first prolongation $\gamma^{(1)}$; we will also use the notation $\ddot\gamma = \gamma^{(2)}$.

For every three natural numbers k, r and s, with $k = r + s$, we can define the embedding

$$\begin{array}{cccc} \iota_{r,s}\colon & \mathrm{T}^k M & \to & \mathrm{T}^r(\mathrm{T}^s M) \\ & \gamma_0^{(k)} & \mapsto & (\gamma^{(s)})_0^{(r)} \end{array},$$

which is a morphism over $\mathrm{T}^s M$. Of particular interest is the embedding $\iota_{1,1}\colon \mathrm{T}^2 M \to \mathrm{T}(\mathrm{T}M)$, which identifies the bundle $\tau_{2,1}\colon \mathrm{T}^2 M \to \mathrm{T}M$ as an affine subbundle of $\tau_{\mathrm{T}M}\colon \mathrm{T}(\mathrm{T}M) \to \mathrm{T}M$. It turns out that a vector field $X \in \mathfrak{X}(\mathrm{T}M)$ is a SODE if and only if its image belongs to $\mathrm{T}^2 M$, so we can identify the second order differential equations with sections of $\tau_{2,1}\colon \mathrm{T}^2 M \to \mathrm{T}M$.

Chapter 3

Time independent systems

In this chapter we introduce the geometric structure which we call linearly singular system. These systems are suitable to model first order ordinary differential equations on manifolds that are affine in the velocities. In local coordinates, such a differential equation is written

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

where $\mathbf{A}$ is in general a singular matrix.

This kind of equations arise in several formalisms of mechanics (Lagrangian, Hamiltonian or unified formalisms), which is the main motivation for us to study them. They also arise in the study of higher order singular lagrangians and their "higher order differential equation" conditions, as well as many other systems that appear in technological applications, such as electric o chemical engineering, control theory or economics —see some references in [BCP 96, GMR 04].

In this chapter we will study time independent systems, leaving the time-dependent case and other generalizations to the next chapters.

Autonomous linearly singular systems were geometrically presented in [GP 91] and developed in [GP 92a, GP 02]. Since the matrix $\mathbf{A}$ may be singular, the system may not have solutions passing through each point of of the configuration manifold, and the solutions may not be unique. In these works was also given the consistency algorithm that should be performed in order to solve the corresponding equation of motion. If the algorithm ends, one obtains a submanifold where there exist solutions of the system. One also obtains the family of vector fields whose integral curves are the solutions of the differential equation. This algorithm is indeed a generalization of the presymplectic constraint algorithm [GNH 78].

In the first section of the chapter we review the main concepts and results introduced in the cited works of Gràcia and Pons. The second section is devoted to the study of Lagrangian systems and we show that they can be enclosed in the geometric setting of linearly singular systems.

3.1 Linearly singular systems

Let M be a manifold. We start by defining what we mean by "differential equation" both implicit and explicit, and its solutions.

Definition 3.1 *An* implicit differential equation *on M is a submanifold $D \subset \mathrm{T}M$. A path $\xi\colon I \to M$ is a* solution of D *when*

$$\dot{\xi}(I) \subset D. \tag{3.1}$$

In coordinates, if the submanifold D is implicitly described by some equations $F^a = 0$ and the path ξ is represented by some functions $x(t)$, then the local expression of the implicit differential equation is

$$F^a(x, \dot{x}) = 0. \tag{3.2}$$

We have a particular case when $D = X(M)$, with X a vector field on M. Then X defines an *explicit differential equation*, and ξ is a solution iff

$$\dot{\xi} = X \circ \xi. \tag{3.3}$$

Now, if the vector field X is locally described as $X = f^i \frac{\partial}{\partial x^i}$, then the local expression of the explicit differential equation is

$$\dot{x}^i = f^i(x). \tag{3.4}$$

Now we will define the systems which are the main purpose of our discussion. These are the systems whose solutions are determined by solving differential equations that are affine in the velocities.

Definition 3.2 *An* autonomous linearly singular system *on M is constituted by a vector bundle $\pi\colon F \to M$, a vector bundle morphism $A\colon \mathrm{T}M \to F$, and a section $f\colon M \to F$ of π:*

$$\begin{array}{ccc} \mathrm{T}M & \xrightarrow{A} & F \\ {\scriptstyle \tau_M}\downarrow & \overset{\pi}{\swarrow}\!\!\nearrow_{f} & \\ M & & \end{array} \tag{3.5}$$

We denote by $(A\colon \mathrm{T}M \to F, f)$ this autonomous linearly singular system.

Taking local coordinates $(x^i, \dot{x}^i)$ on $\mathrm{T}M$ and (x^i, u^α) on F, the local expressions of the maps are

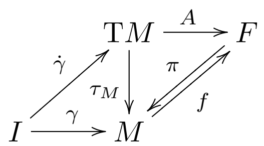
$$\pi(x^i, u^\alpha) = (x^i), \qquad b(x^i) = (x^i, b^\alpha(x^i)), \qquad A(x^i, \dot{x}^i) = (x^i, A^\alpha_j(x^i)\dot{x}^j).$$

We say that a path $\gamma\colon I \to M$ is a *solution path* if

$$A \circ \dot{\gamma} = b \circ \gamma. \tag{3.6}$$

Locally, $A^\alpha_j(\gamma(t))\dot{\gamma}^j(t) = b^\alpha(\gamma(t))$.

The following diagram shows all these data:



From equations (3.1) and (3.6) it follows that a linearly singular system $(A\colon \mathrm{T}M \to F, f)$ is associated with the implicit differential equation defined by

$$D = A^{-1}(f(M)) \subset \mathrm{T}M, \tag{3.7}$$

which has the same solutions as the linearly singular system.

Example 3.3 As we have said in the introduction, some mechanical systems can be modelled with linearly singular systems. To illustrate this, let us consider a simple (planar) pendulum of mass m and length R under constant gravity g. The equations of motion that determine its evolution can be written (though it is not the usual way) as:

$$\begin{aligned} \dot{x} &= v_x \\ \dot{y} &= v_y \\ \dot{v}_x &= -\frac{\tau}{R}x \\ \dot{v}_y &= -\frac{\tau}{R}y - g \\ 0 &= x^2 + y^2 - R^2 \end{aligned}$$

where (x, y) are the position coordinates, (v_x, v_y) the velocities and τ denotes the string tension. Thus, we take as configuration manifold of the system $M = \mathbf{R}^5$ with coordinates (x, y, v_x, v_y, τ). Then, the equations of motion are those given by the linearly singular system $(A\colon \mathrm{T}M \to F, f)$, where the vector bundle F is $M \times \mathbf{R}^5 \to M$, the section f is $(v_x, v_y, -\frac{\tau}{R}x, -\frac{\tau}{R}y - g, x^2 + y^2 - R^2)$ and the vector bundle morphism A has coordinate matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

In chapter 4 we will see more examples of autonomous linearly singular systems.

The solutions of a linearly singular system can be equivalently described as integral curves of vector fields. Let us remark that in general the solutions are restricted to a submanifold $S \subset M$ because the equation (3.6) may not have solutions passing through every point $x \in M$. Therefore, the equation of motion can be written as an equation for a vector field X and a submanifold S:

$$\begin{cases} X \text{ tangent to } S \\ A \circ X \underset{S}{\simeq} f, \end{cases} \tag{3.8}$$

where the notation $\underset{S}{\simeq}$ means equality at the points of S. If X is a solution of (3.8), then its integral curves are solutions of A. For this reason, the vector fields that are solutions of (3.8) are called *solution vector fields* of the linearly singular system A.

Regular and surjective systems

In some situations, the solutions of a linearly singular system are easy to find. The most evident case is when the differential equation is actually explicit.

Definition 3.4 *An autonomous linearly singular system* $(A\colon \mathrm{T}M \to F, f)$ *is* regular *when A is a vector bundle isomorphism.*

In this case, the differential equation associated with the linearly singular system is explicit and it is given by the vector field $X = A^{-1} \circ f$. Therefore, X is the (unique) solution of (3.8) and it is not restricted to any submanifold S.

A system may not be regular but may still have solutions on the whole manifold M, as is the case in the following systems.

Definition 3.5 *An autonomous linearly singular system* $(A\colon \mathrm{T}M \to F, f)$ *is* surjective *when A is a surjective vector bundle morphism.*

Now, the solutions of (3.8) can be expressed as $X_\circ + \Gamma$, where $X_\circ$ is a particular solution and Γ belongs to $\mathrm{Ker}\, A$. Again, since A is surjective, the solutions are defined on the whole manifold.

Constraint algorithm

When a linearly singular system is not surjective, a recursive algorithm can be applied to find its solutions. This algorithm is a generalization of the Gotay–Nester algorithm [GNH 78] for presymplectic systems, which, in turn, is a geometric version of the Dirac–Bergmann algorithm [BG 55, Dir 64] for systems with singular lagrangian. It is designed to find the maximal submanifold $S \subset M$ where (3.8) has solutions. In other words, we want to obtain the maximal surjective subsystem of the original system.

Now we will explain to some detail the first step of the algorithm. We start noting that, in order that a solution passes through a point $x \in M$, it is necessary that

$$f(x) \in \mathrm{Im}\, A_x, \tag{3.9}$$

so the solutions are necessarily contained in the primary constraint subset

$$M_1 = \{x \in M \mid f(x) \in \operatorname{Im} A_x\}, \tag{3.10}$$

which will be assumed to be a closed submanifold. This is the case when the vector bundle morphism A has constant rank. It can be seen that the primary constraint submanifold M_1 is locally described by the vanishing of the functions $\phi^\alpha := \langle s^\alpha, b\rangle$, where (s^α) is a local frame for $\operatorname{Ker} {}^tA \subset F^*$.

We have that a vector field X in M, in order to be a solution of the system, must satisfy the equation $A \circ X \underset{M_1}{\simeq} b$. Vector fields satisfying this condition always exist and are called primary vector fields. Given one primary vector field X_0, the others have the form $X \underset{M_1}{\simeq} X_0 + \sum_\mu f^\mu \Gamma_\mu$, where f^μ are functions uniquely determined on M_1 and (Γ_μ) is a local frame for $\operatorname{Ker} A$.

The tangency to M_1 forces the initial system to be restricted to $(A_1 \colon \mathrm{T}M_1 \to F_1, f_1)$, where $A_1 = A|_{\mathrm{T}M_1}$, $F_1 = F|_{M_1}$ and $f_1 = f|_{M_1}$. Thus, the first step is completed and we start with the second. A primary vector field X can be a solution of the system only if it is tangent to M_1, so we obtain the equation, for every constraint ϕ^α, $(X \cdot \phi^\alpha)(x) = 0$, for $x \in M_1$, or, equivalently, $(X_0 \cdot \phi^\alpha)(x) + \sum_\mu (\Gamma_\mu \cdot \phi^\alpha)(x) f^\mu(x) = 0$. These equations may provide new constraints that define the secondary constraint submanifold M_2, and may also determine some of the functions f^μ.

The algorithm follows recursively and, assuming that throughout the process the subsets M_i are always closed submanifolds, it ends with a final constraint submanifold S such that $f(S) \subset \operatorname{Im} A_S$; thus the system is surjective, so the equation

$$A_S \circ X = f_S \tag{3.11}$$

for a vector field X tangent to S has solutions. These solutions will be the primary vector fields determined through the process. Given a particular solution $X_\circ$, the set of all the solutions of (3.8) can be described as $X_\circ + \operatorname{Ker} A_S$, as we saw in the previous section.

Symmetries

Symmetries of differential equations is a vast topic that has been studied from different perspectives (see for instance [Olv 93]). For the case that interests us, symmetries of linearly singular system were discussed in [GP 02]. We will define the concept of symmetry for these systems following this paper.

Definition 3.6 *A* symmetry *of an implicit differential equation $D \subset \mathrm{T}M$, is a diffeomorphism φ of M that leaves D invariant, that is,* $(\mathrm{T}\varphi)(D) \subset D$.

Definition 3.7 *A* symmetry *of a linearly singular system $(A \colon \mathrm{T}M \to F, f)$ is a vector bundle automorphism (φ, Φ) of $\pi \colon F \to M$ such that*

$$f = \Phi_*[f] := \Phi \circ f \circ \varphi^{-1}, \qquad A = \Phi_*[A] := \Phi \circ A \circ (\mathrm{T}\varphi)^{-1}. \tag{3.12}$$

It is easily seen that the base map φ of a symmetry (φ, Φ) of the linearly singular system A is itself a symmetry of the associated implicit differential equation $D = A^{-1}(f(M))$. In [GP 02] is proved that, if A has constant rank, a kind of converse is also true: every symmetry φ of D is *locally* the base map of a symmetry of $(A\colon \mathrm{T}M \to F, f)$.

We will also give the infinitesimal counterparts of the previous facts. First we define the concepts of infinitesimal symmetry that we will use.

Definition 3.8 *An* infinitesimal symmetry *of an implicit differential equation $D \subset \mathrm{T}M$ is a vector field V on M such that its flow F_V^ε is constituted by (local) symmetries of D.*

Or, equivalently, the canonical lift of V to $\mathrm{T}M$, V^T, is tangent to D.

Definition 3.9 *An* infinitesimal symmetry *of a linearly singular system $(A\colon \mathrm{T}M \to F, f)$ is an infinitesimal automorphism (V, W) of the vector bundle $\pi\colon F \to M$ such that its flow $(\mathrm{F}_V^\varepsilon, \mathrm{F}_W^\varepsilon)$ is constituted by (local) symmetries of the linearly singular differential equation.*

The last property is equivalent to the following conditions:

$$\mathrm{T}f \circ V = W \circ f, \qquad \mathrm{T}A \circ V^T = W \circ A, \tag{3.13}$$

which are the infinitesimal version of (3.12).

Then, given a linearly singular system $(A\colon \mathrm{T}M \to F, f)$ and its associated implicit differential equation D, and under the assumption that A has constant rank, a vector field V on M is an infinitesimal symmetry of D if and only if there exists a vector field W on F such that (V, W) is an infinitesimal symmetry of the linearly singular system.

3.2 Lagrangian systems

Here we present a standard geometric formulation of autonomous Lagrangian systems that can be found in many reference books [AM 78, LR 89, Sou 97, JS 98].

The mathematical definition of Lagrangian system is very simple:

Definition 3.10 *An* autonomous Lagrangian system *consists of an n-dimensional manifold M, called the* configuration space *and a function on the tangent bundle of M, $L\colon \mathrm{T}M \to \mathbf{R}$, called the* lagrangian *of the system.*

Although we may study Lagrangian systems independently of any physical interpretation, it is generally assumed that a Lagrangian system is the mathematical model of some physical system, so that the points of the configuration space M represent possible configurations of the physical system, and the lagrangian L, which contains the dynamical information, should be chosen such that the solutions of the Lagrangian system represent the real motions of the physical system. The tangent space $\mathrm{T}M$ is also called *velocity space* and its elements (tangent vectors) correspond to the *states* of the system.

Example 3.11 One of the most studied class of Lagrangian systems is that of simple mechanical systems. They are characterized by the special form of its lagrangian:

$$L = T - U \circ \tau_M,$$

where $T\colon \mathrm{T}M \to \mathbf{R}$ is the *kinetic energy* associated with a Riemannian metric g, that is, $T(v_p) = g(v_p, v_p)$, and $U\colon M \to \mathbf{R}$ is a function called *potential energy*.

Roughly speaking, the lagrangian of a system is a sort of "cost" function, so that the solutions of the system are the curves on the configuration manifold that minimize in some sense the lagrangian. We next give the precise mathematical development of this idea, that leads to the Hamilton's principle.

Euler–Lagrange equations

The solutions of a Lagrangian system (M, L) will be curves γ on the configuration manifold. We require that the solutions are twice differentiable. Given two points p and q of the configuration space and an interval $[a, b]$ of $\mathbf{R}$, consider

$$\mathcal{C}^2([a,b],p,q) = \{\gamma\colon [a,b] \to M \mid \gamma \text{ is } C^2, \gamma(a) = p, \gamma(b) = q\},$$

the space of twice differentiable curves that are defined on $[a, b]$ and that start at p and end at q. This is the space of curves over which the minimization principle will be stated.

The function of $\mathcal{C}^2([a,b],p,q)$ to be minimized is the action functional associated with the lagrangian L, that is, the function $J_L\colon \mathcal{C}^2([a,b],p,q) \to \mathbf{R}$ defined as

$$J_L(\gamma) = \int_a^b L(\dot\gamma(t))\mathrm{d}t.$$

Now we have all the necessary ingredients to state the *Hamilton's principle*: a curve $\gamma \in \mathcal{C}^2([a,b],p,q)$ is a *solution of the Lagrangian system* defined by L if it minimizes J_L. A necessary condition for a curve γ to be a minimizer of J_L is that $\mathrm{d}J_L(\gamma) = 0$, that is to say, that γ is an extremal of J_L.

Remark 3.12 Sometimes [AM 78, MR 99] the Hamilton's principle is formulated in a different way, stating that the solutions of the system are all the extremals of J_L, not only the minimizers. It can be seen that if the matrix $\left(\partial^2 L/\partial\dot x^i\,\partial\dot x^j\right)$ is positive definite everywhere then all the extremals are minimizers. For instance, this is the case of the simple mechanical systems of example 3.11. In this dissertation we will not be concerned about this question.

It is well-known (see [AM 78]) that, locally, the equation $\mathrm{d}J_L(\gamma) = 0$ is equivalent to the *Euler–Lagrange equations* for L,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^i} \circ \dot{\gamma}\right) - \frac{\partial L}{\partial x^i} \circ \dot{\gamma} = 0,\ 1 \leq i \leq n. \tag{3.14}$$

Symplectic formulation

There are different approaches to write the Euler–Lagrange equations in an intrinsic way. We discuss here a classical one: the symplectic formulation of Lagrangian systems.

First we need to define some geometric objects associated with a Lagrangian system. The *energy* of the system is the function on $\mathrm{T}M$,

$$E_L = \Delta(L) - L,$$

where Δ is the Liouville vector field (2.2). Note that in the simple mechanical case (example 3.11), where $L = T - U$, since the kinetic energy T is homogeneous of degree 2 and the potential energy U is homogeneous of degree 0, we have that $E_L = \Delta(L) - L = 2T - (T - U) = T + U$, an equation that shows that E_L represents the classical total energy of the mechanical system.

The other relevant objects are the Poincaré–Cartan forms. The Poincaré–Cartan 1-form $\theta_L \in \Omega^1(\mathrm{T}M)$ is defined by

$$\theta_L = {}^tS \circ \mathrm{d}L,$$

where S is the vertical endomorphism (2.3), and the Poincaré–Cartan 2-form is

$$\omega_L = -\mathrm{d}\theta_L.$$

In local coordinates, the Poincaré–Cartan forms read as

$$\theta_L = \frac{\partial L}{\partial \dot{x}^i}\mathrm{d}x^i,$$

$$\omega_L = \frac{\partial^2 L}{\partial \dot{x}^i\,\partial x^j}\mathrm{d}x^i \wedge \mathrm{d}x^j + \frac{\partial^2 L}{\partial \dot{x}^i\,\partial \dot{x}^j}\mathrm{d}x^i \wedge \mathrm{d}\dot{x}^j.$$

Let us point out here that we can construct the Poincaré–Cartan forms following another route. Consider the fibre derivative $\mathcal{F}L\colon \mathrm{T}M \to \mathrm{T}^*M$ (recall equation (2.1)) of the lagrangian function. It is called the *Legendre transformation.* It locally reads as $\mathcal{F}L(x, \dot{x}) = (x^i, \partial L/\partial \dot{x}^i(x, \dot{x}))$. Now, recall the the canonical forms of the cotangent bundle θ_M and ω_M. It turns out that

$$\theta_L = \mathcal{F}L^*\theta_M \quad \text{and} \quad \omega_L = \mathcal{F}L^*\omega_M. \tag{3.15}$$

We already have the geometric objects needed for the symplectic formulation. Since ω_L is closed, the triplet $(M, \omega_L, \mathrm{d}E_L)$ is a presymplectic dynamical system. Consider its associated equation, for vector fields X on $\mathrm{T}M$,

$$i_X\omega_L = \mathrm{d}E_L. \tag{3.16}$$

The discussion of this equation splits into two cases, according to whether the Poincaré–Cartan 2-form ω_L is nodegenerate or not. This characteristic is so important that the Lagrangian systems are divided into two classes. We say that a lagrangian function L (and the Lagrangian system that it defines) is *regular* if any of these equivalent conditions hold:

- The Poincaré–Cartan 2-form ω_L is nondegenerate (hence symplectic, since it is closed by definition).
- The Legendre transformation $\mathcal{F}L$ is a local diffeomorphism.
- The fibre hessian of the lagrangian, $\mathcal{F}^2L \colon \mathrm{T}M \to \mathrm{T}^*M \otimes \mathrm{T}^*M$, is nondegenerate.

In coordinates, these conditions are equivalent to the regularity of the Hessian matrix

$$\left(\frac{\partial^2 L}{\partial \dot{x}^i \, \partial \dot{x}^j} \right).$$

The lagrangians (and Lagrangian systems) that are not regular are called *singular*.

Now, if the lagrangian is regular, since ω_L is symplectic, equation (3.16) has a unique solution, which we denote by Γ_L. It can be seen that Γ_L satisfies these two properties:

1. It is a second order differential equation. We recall that this implies that its integral curves $\xi(t)$ are the velocities of curves on M: $\xi(t) = (\tau_M \circ \xi)^{\cdot}(t)$.
2. The curves $\gamma(t) = (\tau_M \circ \xi)(t)$, where $\xi(t)$ is and integral curve of Γ_L, are the solutions of the Euler–Lagrange equations (3.14).

On the other hand, if the system is singular (we refer to [Car 90] for a detailed discussion on singular lagrangian systems), the Poincaré–Cartan 2-form ω_L is just presymplectic and equation (3.16) may be unsolvable or solvable only over a submanifold of M. Furthermore, the equation (where solvable) may have multiple solutions, because sections of $\operatorname{Ker}\omega_L$ can be added to any solution of (3.16) to obtain more solutions.

Finally, it is not assured that the solutions are SODE as in the regular case. Therefore, the equation

$$S(X) = \Delta \tag{3.17}$$

should be added to equation (3.16) to get rid of the vector fields that are not a SODE. In spite of all these problems, the SODE solutions of equation (3.16) are vector fields (possibly defined only on a submanifold of M) whose integral curves are the velocities of solutions of the Euler–Lagrange equations (3.14). In order to obtain these solutions, we can apply the geometric algorithm of presymplectic systems proposed by Gotay, Nester and Hinds [GNH 78] (see also [LR 89]). This algorithm is in fact the constraint algorithm for linearly singular systems given in section 3.1 when we consider the presymplectic system as a linearly singular system in the way that we will see later.

Before this, we will see two different formulations for degenerate Lagrangian systems where the second-order condition arises directly from the equations of motion.

Formulation with the operator K

There is an interesting alternative way to describe intrinsically the Euler–Lagrange equations with makes use of the so-called time-evolution operator K. In local coordinates, K was first introduced in [BGPR 86] (see also [Car 90]) as a differential operator $K\colon \mathrm{C}^\infty(\mathrm{T}^*M) \to \mathrm{C}^\infty(\mathrm{T}M)$. In [GP 89] is presented as a vector field along the Legendre transformation $\mathcal{F}L$, that is, a map $K\colon \mathrm{T}M \to \mathrm{T}(\mathrm{T}^*M)$ such that $K(u_p) \in \mathrm{T}_{\mathcal{F}L(u_p)}(\mathrm{T}^*M)$, satisfying the two following properties that determine it completely:

$$\begin{cases} \mathrm{T}\pi_M \circ K = \mathrm{id}_{\mathrm{T}M} \\ (\mathcal{F}L)^*(i_K\omega_Q) = \mathrm{d}E_L \end{cases} .$$

The second equation, which involves operations with tensor fields along a map, it is understood as follows: for every $v \in \mathrm{T}M$, $({}^t(\mathrm{T}_v\mathcal{F}L) \circ \widehat{\omega}_Q \circ K)(v) = \mathrm{d}E_L(v)$.

Its local expression in natural coordinates is

$$K(x,\dot{x}) = \left(x^i, \frac{\partial L}{\partial \dot{x}^i}(x,\dot{x}); \dot{x}^i, \frac{\partial L}{\partial x^i}(x,\dot{x}) \right) .$$

After some computations it can be seen that a path $\xi\colon I \to \mathrm{T}M$ is the velocity of a solution of the Euler–Lagrangre equations (3.14) if and only if $\mathrm{T}(\mathcal{F}L) \circ \dot{\xi} = K \circ \xi$. Moreover, if a vector field $X \in \mathfrak{X}(\mathrm{T}M)$ is a solution of the equation

$$\mathrm{T}(\mathcal{F}L) \circ X = K, \tag{3.18}$$

then it is automatically a SODE, so its integral curves are derivatives of curves on M that, as we have just said, are solutions of the Euler–Lagrange equations. Therefore, equation (3.18) is a good intrinsic way to write the Euler–Lagrange equations for singular lagrangians. It is worth mentioning that the operator K it is useful in the theory of singular lagrangians not only to express the equations of motion but also to relate the Lagrangian and Hamiltonian formulation [BGPR 86, CL 87, Pon 88], to study symmetries [GP 88, FP 90, GP 00] or to study Lagrangian systems with generic singularities [PV 00].

First-order formulation

Yet there is another formulation, developed geometrically by Skinner and Rusk [Ski 83, SR 83] of the dynamical equations of singular systems. We can say that this approach is a mixed velocity-momentum description of Lagrangian systems because the Euler–Lagrange equations are seen to be equivalent to a first-order differential equation on the Whitney sum $\mathrm{T}^*M \oplus \mathrm{T}M$.

Then let us consider the Whitney sum of T^*M and $\mathrm{T}M$, which we denote by $W = \mathrm{T}^*M \oplus \mathrm{T}M$. The natural projections of W on T^*M and $\mathrm{T}M$ are denoted by pr_1 and pr_2 respectively. We can define a presymplectic form on W as the pull-back of the canonical symplectic form of T^*M:

$$\omega_W = \mathrm{pr}_1^*(\omega_M)$$

and the function $E_W \in \mathrm{C}^\infty(W)$,

$$E_W = \langle \mathrm{pr}_1, \mathrm{pr}_2 \rangle - \mathrm{pr}_2^*(L).$$

In [Ski 83] it was shown that the equation

$$i_Z \omega_W = \mathrm{d}E_W \tag{3.19}$$

for a vector field $Z \in \mathfrak{X}(W)$ is equivalent to equation (3.16) together with the SODE condition (3.17). More precisely, a solution Z of (3.19) is pr_2-related to a SODE X (that is, $\mathrm{Tpr}_2 \circ Z = X \circ \mathrm{pr}_2$) solution of (3.16). Therefore, the projection to M of integral curves of Z are solutions of the Euler–Lagrange equations (3.14).

Since (3.19) is a presymplectic equation, the Gotay–Nester–Hinds algorithm [GNH 78] can be applied. See [SR 83] for details and other procedures to generate the final constraint submanifold.

Hamiltonian formalism

Now we will briefly describe the Hamiltonian formalism of a Lagrangian system (M, L). Its main general characteristics are that the dynamics take place at the cotangent bundle T^*M, which is naturally endowed with a symplectic structure, and the equations of motion locally take the form of the well-known Hamilton equations. An important reason to have a Hamiltonian description of a system is that it is the first step towards quantization. We will follow the exposition given in [Car 90].

The link between the Lagrangian and Hamiltonian formalism is the Legendre transformation $\mathcal{F}L \colon \mathrm{T}M \to \mathrm{T}^*M$ associated with the lagrangian L. Recall that the lagrangian is regular if $\mathcal{F}L$ is a local diffeomorphism.

If $\mathcal{F}L$ is a global diffeomorphism we say that the lagrangian function L (and also the Lagrangian system) is *hyperregular*. In this case, we can relate a Hamiltonian system on T^*M (see section (2.2)) to the Lagrangian system as follows. We define the hamiltonian function as $H = E_L \circ \mathcal{F}L^{-1} \in \mathrm{C}^\infty(\mathrm{T}^*M)$. Taking into account that $\omega_L = \mathcal{F}L^*\omega_M$, the Hamiltonian vector field X_H, which is the solution of the Hamiltonian equation

$$i_Y \omega_M = \mathrm{d}H, \tag{3.20}$$

is $\mathcal{F}L$-related to the solution Γ_L of equation (3.16). Therefore, the solutions of the Euler–Lagrange equations (3.14) are the curves $\gamma(t) = (\pi_M \circ \eta)(t)$, where $\eta(t)$ is an integral curve of X_H. In natural coordinates (x^i, p_i) of T^*M, the equations for the integral curves of X_H are the Hamilton equations

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial x^i}.$$

If the lagrangian L is singular we can not construct a hamiltonian function in T^*M. Nevertheless, in some cases we can obtain a system on the submanifold $P_1 = \mathcal{F}L(\mathrm{T}M) \subset \mathrm{T}^*M$ which is $\mathcal{F}L$-related to the presymplectic system $(\mathrm{T}M, \omega_L, \mathrm{d}E_L)$. We say that the lagrangian is *almost-regular* if $\mathcal{F}L$ is a submersion onto P_1 and the fibres $\mathcal{F}L^{-1}(\mathcal{F}L(v))$ are connected for each $v \in \mathrm{T}M$. This concept was introduced by Gotay and Nester in [GN 79], where they prove that in this case the energy E_L is $\mathcal{F}L$-projectable, that is, there exists a function $H_1 \in \mathrm{C}^\infty(P_1)$ such that $H_1 \circ \mathcal{F}L = E_L$. Let ω_1 be the pull-back of ω_M to P_1. It is clear that $(P_1, \omega_1, \mathrm{d}H_1)$ is a presymplectic system $\mathcal{F}L$-related to the lagrangian one $(\mathrm{T}M, \omega_L, \mathrm{d}E_L)$. Thus, for almost-regular lagrangians there is a hamiltonian formalism of the dynamics with equation

$$i_Y \omega_1 = \mathrm{d}H_1, \tag{3.21}$$

for $Y \in \mathfrak{X}(P_1)$. As in the Lagrangian formalism, this equation may lead to constraints, so the solutions (if exist) are found by means of the constraint algorithm.

Lagrangian systems and linearly singular systems

Now, realize that equation (3.16) is the equation for a vector field to be the solution of the autonomous linearly singular system $(\widehat{\omega}_L\colon \mathrm{T}(\mathrm{T}M) \to \mathrm{T}^*(\mathrm{T}M), \mathrm{d}E_L)$:

$$\begin{array}{ccc} \mathrm{T}(\mathrm{T}M) & \xrightarrow{\widehat{\omega}_L} & \mathrm{T}^*(\mathrm{T}M) \\ \downarrow & \nearrow_{\mathrm{d}E_L} & \\ \mathrm{T}M & & \end{array} . \tag{3.22}$$

In fact, every presymplectic system (Q, ω, α) is equivalent to the linearly singular system $(\widehat{\omega}\colon \mathrm{T}Q \to \mathrm{T}^*Q, \alpha)$.

Or, if the system is singular and we want to deal jointly with equation (3.16) and the SODE problem, we can add equation (3.17) to the autonomous linearly singular system (3.22) by means of a Whitney sum, obtaining the linearly singular system $(\widehat{\omega}_L \oplus S\colon \mathrm{T}(\mathrm{T}M) \to \mathrm{T}^*(\mathrm{T}M) \oplus \mathrm{T}(\mathrm{T}M), \mathrm{d}E_L \oplus \Delta)$:

$$\begin{array}{ccc} \mathrm{T}(\mathrm{T}M) & \xrightarrow{\widehat{\omega}_L \oplus S} & \mathrm{T}^*(\mathrm{T}M) \oplus \mathrm{T}(\mathrm{T}M) \\ \downarrow & \nearrow_{\mathrm{d}E_L \oplus \Delta} & \\ \mathrm{T}M & & \end{array} .$$

The formulation for singular systems using the time-evolution operator K led to equation (3.18), which is equivalent to the linearly singular system $(\mathring{\mathrm{T}}(\mathcal{F}L)\colon \mathrm{T}(\mathrm{T}M) \to \mathrm{T}(\mathrm{T}^*M) \times_{\mathcal{F}L} \mathrm{T}M, \mathring{K})$:

$$\begin{array}{ccc} \mathrm{T}(\mathrm{T}M) & \xrightarrow{\mathring{\mathrm{T}}(\mathcal{F}L)} & \mathrm{T}(\mathrm{T}^*M) \times_{\mathcal{F}L} \mathrm{T}M \\ \downarrow & \nearrow_{\mathring{K}} & \\ \mathrm{T}M & & \end{array} ,$$

where $\mathring{\mathrm{T}}(\mathcal{F}L) = (\mathrm{T}(\mathcal{F}L), \tau_{\mathrm{T}M})$ and $\mathring{K} = (K, \mathrm{id}_{\mathrm{T}M})$. This maps are used because the operator K is not the section of a bundle but a section along $\mathcal{F}L$. Therefore, we use the pull-back $\mathrm{T}(\mathrm{T}^*M) \times_{\mathcal{F}L} \mathrm{T}M$ as the vector bundle of the linearly singular system and the section $\mathring{K}$ and vector bundle morphism $\mathring{\mathrm{T}}(\mathcal{F}L)$ corresponding to K and $\mathrm{T}\mathcal{F}L$.

Finally, the mixed formulation of Skinner and Rusk (equation (3.19)) is equivalent to linearly singular system on $W = \mathrm{T}^*M \oplus \mathrm{T}M$, $(\widehat{\omega}_W\colon \mathrm{T}(\mathrm{T}W) \to \mathrm{T}^*(\mathrm{T}W), \mathrm{d}E_W)$:

$$\begin{array}{ccc} \mathrm{T}W & \xrightarrow{\widehat{\omega}_W} & \mathrm{T}^*W \\ \downarrow & \nearrow_{\mathrm{d}E_W} & \\ W & & \end{array} .$$

Therefore, the three alternative formulations of singular Lagrangian systems that we have seen can be interpreted as linearly singular systems that are not consistent. We should use the constraint algorithm of section (3.1) to obtain the maximal submanifold where a solutions exist.

Also the Hamiltonian formalism that we have seen is equivalent to a linearly singular system, since it leads to a presymplectic system $(P_1, \omega_1, \mathrm{d}H_1)$ (equation (3.21)). The corresponding linearly singular system is

$$\begin{array}{ccc} \mathrm{T}P_1 & \xrightarrow{\widehat{\omega}_1} & \mathrm{T}^*P_1 \\ \downarrow & \nearrow_{\mathrm{d}H_1} & \\ P_1 & & \end{array} .$$

Forced systems

During this discussion, we have assumed that there are no "external forces" acting on the system. In the context of Lagrangian systems, forces are represented by bundle morphisms $F\colon \mathrm{T}M \to \mathrm{T}^*M$, in other words, forces are 1-forms along τ_M. In coordinates, a force is written as $F(x^i, \dot{x}^i) = F_i(x^i, \dot{x}^i)\mathrm{d}x^i$. It is usual to identify the force F with the semibasic 1-form on $\mathrm{T}M$ given by $(\mathrm{T}\tau_M)^*F$; it has the same local expression.

Without going into details, in presence of forces it is applied the so-called Lagrange–d'Alembert principle, which states how a force affects the motion of a system. The resulting equations are the *forced Euler–Lagrange equations*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^i}\circ\dot{\gamma}\right)-\frac{\partial L}{\partial x^i}\circ\dot{\gamma}=F_i\circ\dot{\gamma},\ 1\leq i\leq n.$$

The symplectic formulation of the forced equations has the simple form

$$i_X\omega_L=\mathrm{d}E_L+F. \tag{3.23}$$

Viewed as linearly singular systems, the only difference between the unforced and forced system is in the section of $\pi_{\mathrm{T}M}$. The forced system is equivalent to the linearly singular system $(\widehat{\omega}_L\colon \mathrm{T}(\mathrm{T}M)\to\mathrm{T}^*(\mathrm{T}M), \mathrm{d}E_L+F)$.

Chapter 4

Generalized nonholonomic systems

The main goal of this chapter is to study the nonholonomic mechanical systems within the framework of linearly singular differential equations.

A lagrangian system with nonholonomic constraints may be considered, thinking in a more general way, as a singular differential equation defined by some constraints and some multipliers:

$$\dot{x} = g(x) + \sum_{\mu} u^{\mu}\, h_{\mu}(x), \qquad \phi_{\alpha}(x) = 0.$$

Such an equation can be described geometrically as a linearly singular system, the type of implicit differential equations that we discussed in the previous chapter, that is, a differential equation where the velocities are not isolated because of a linear factor multiplying them:

$$A(x)\dot{x} = b(x).$$

The idea of modelling mechanical systems as implicit differential equations is found in earlier papers by Tulczyjew [MT 78, MMT 95], and it has also been used to deal with nonholonomic constraints [Tul 86, ILMM 96].

In this chapter we see that a system with constraints and multipliers, and in particular any nonholonomic mechanical system, can be described as a linearly singular system. Therefore, all the methods and results about linearly singular systems can be applied directly to nonholonomic systems.

More precisely, the combination of two operations that can be performed on linearly singular systems —restriction to a subsystem and projection to a quotient— can be applied to obtain what we call a generalized nonholonomic system. We deal with the regularity, consistency and equations of motion of these derived systems in section 4.3. To prove some results of this section, previously we give some lemmas about linear algebra in section 4.2

In the previous chapter, we studied the symmetries of a linearly singular system. Here in section 4.4 we discuss the relation between the symmetries of a system with nonholonomic constraints and the symmetries of its original unconstrained system, both modelled on linearly singular differential equations. We will also study their constants of motion.

In section 4.5 we show how a lagrangian system with nonholonomic constraints can be described in terms of a generalized nonholonomic systems. A system constituted by a relativistic particle moving in spacetime under the action of an electromagnetic field and a potential is studied in section 4.7, where we see that a nonholonomic constraint can convert a singular lagrangian into a regular system. Two additional examples are studied in section 4.8.

4.1 Nonholonomic Lagrangian systems

Consider a Lagrangian system (M, L) as defined in section 3.2. For different physical reasons, that in principle are not related with the lagrangian L, it could happen that some states of the system are impossible to attain. We say then that the system has constraints. Here we will only deal with smooth constraints on the positions and velocities of the system:

Definition 4.1 *A* constrained Lagrangian system *is a Lagrangian system* (M, L) *on a manifold* M *together with a submanifold* C *of* $\mathrm{T}M$*, called the* constraint submanifold*.*

Obviously, C represents the states that the physical system can reach. We will assume that all the configurations are reachable, that is, $\tau_M(C) = M$. If the constraint C is derived from a constraint in the configuration space (that is, $C = \mathrm{T}N$ with N a submanifold of M) we say that the constraint is *holonomic*. More generally, a *semi-holonomic* constraint C is the total space of an integrable distribution, so through each state $v \in \mathrm{T}M$ there passes a submanifold S of M such that $\mathrm{T}_v S = C_v$. Therefore, in this case the constraint leads to some conservation laws and induces a foliation on M by integral submanifolds.

In this dissertation we are interested in the remaining case, when the constraint C is not integrable. In this case the constraint C is said to be *nonholonomic*.

Obviously, we can not expect that the constrained systems behaves as the free system. One possible approach is to assume that there is a force, which we call the *constraint* or *reaction force*, acting upon the free system in such a way that the solutions of the forced system correspond to the motions of the constrained system. The question is which force or set of forces have to be chosen in order to give a real description of the physical system. In most cases, in particular when the constraints are generated by interactions such as sliding and rolling between different components of the system, it is used the rule proposed by Chetaev [Che 34], which can be expressed in geometric terms as follows: in a nonholonomic Lagrangian system (M, L, C), the constraint force belongs to the vector subbundle ${}^tS((\mathrm{T}C)^\vdash) \subset \mathrm{T}^*(\mathrm{T}M)|_C$. This subbundle is called the *Chetaev bundle*. The Chetaev's rule is a generalization of the so-called d'Alembert's principle, valid for linear or affine constraints (that is, when C is a vector subbundle or an affine subbundle of $\mathrm{T}M$). We remark that when the constraint is semi-holonomic (or holonomic), the result of applying the d'Alembert's principle is that on each integral submanifold S the system behaves as the unconstrained system $(S, L|_S)$, as should be expected.

We are not going to discuss these principles, we refer the reader to [Mar 98]. In any case, we note that in some systems the Chetaev's rule may not lead to the right equations of motion; this is due to the fact that the reaction forces depend on the nature of the constraints. If this is the case, we can consider more generally that the constraint force belongs to a given vector subbundle $F \subset \mathrm{T}^*(\mathrm{T}M)|_C$, called the *bundle of forces*. In [ILMM 96] is given a more general formulation, where the constraint forces are also given independently of the constraint manifold and, furthermore, allows restrictions on the accelerations of the system. See [CILM 04] for an even more general discussion that allows higher–order constraints.

There is another approach to constrained system, in which the dynamics are derived from a variational principle (a certain natural generalization of Hamilton's principle). This model, originally proposed by Kozlov [Koz 83], is known with the name of vakonomic mechanics (mechanics of **v**ariational **a**xiomatic **k**ind). This approach is natural when dealing with some problems in engineering or economics, but we will not follow it here. Some references on this topic are [Arn 83, LM 95, CLMM 02]. Unified geometrical approaches to both nonholonomic and vakonomic mechanics are given in [LMM 00, GMM 03].

Through this dissertation we will follow the nonholonomic approach to constrained systems and, particularly, we will accept the Chetaev's rule. Therefore, the systems that we will discuss are the following:

Definition 4.2 *A* nonholonomic Lagrangian system *is a constrained Lagrangian system (M, L, C) such that $\tau_M(C) = M$ and the Chetaev's rule provides the dynamics.*

Since a nonholonomic Lagrangian system (M, L, C) is a forced system, the forced equation (3.23) can be used to find the solutions. Recall that a solution of equation (3.23) is a vector field X of $\mathrm{T}M$ such that its integral curves are velocities of solutions of the forced Euler–Lagrange equations. In this case, the only thing we know about the force is that, on the constraint submanifold C, it belongs to the Chetaev bundle, so equation (3.23) takes the form

$$(i_X\omega_L - \mathrm{d}E_L)|_C \in {}^tS((\mathrm{T}C)^\vdash). \tag{4.1}$$

Of course, the velocity of a solution of the nonholonomic Lagrangian system must be a curve in C. This is ensured by equation

$$X|_C \in \mathrm{T}C, \tag{4.2}$$

so that an integral curve of X with initial condition in C it is a curve in C. The value of X outside C is irrelevant in both equations, which agrees with the concept of constrained system: since the points outside the constraint are unattainable, the value of a solution vector field (which, roughly speaking, represent the accelerations) outside C is not important. Therefore, it is more logical to consider the vector field as defined only on C, that is, $X \in \mathfrak{X}(C)$. Then, equations (4.1) and (4.2) sum up to

$$\left\{ \begin{array}{l} (\widehat{\omega}_L|_C) \circ X - \mathrm{d}E_L|_C \in {}^tS((\mathrm{T}C)^\vdash) \\ X \in \mathfrak{X}(C) \end{array} \right. , \tag{4.3}$$

where $\widehat{\omega}_L|_C$ denotes the restriction of $\widehat{\omega}_L$ to $\mathrm{T}(\mathrm{T}M)|_C$. If the constraint C is locally defined by the vanishing of m constraint functions ϕ^α, then equations (4.3) have the local expression

$$\left\{ \begin{array}{rcll} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^i} \circ \dot{\gamma}\right) - \frac{\partial L}{\partial x^i} \circ \dot{\gamma} & = & \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot{x}^i} \circ \dot{\gamma}, & 1 \le i \le n \\ \phi^\alpha \circ \dot{\gamma} & = & 0, & 1 \le \alpha \le m \end{array} \right. , \tag{4.4}$$

where λ^α (called *Lagrange multipliers*) are m functions to be determined. Thus, there are $m+n$ equations with $m+n$ unknowns.

In case that the Chetaev's rule is not valid but we have some information that provides with a bundle of forces $F \subset \mathrm{T}^*(\mathrm{T}M)|_C$, the equations of motion are formulated as (4.3), replacing the Chetaev bundle ${}^tS((\mathrm{T}C)^\vdash)$ by F.

4.2 Some lemmas about linear algebra

Here we present three lemmas about vector spaces and linear maps that are not widely known and will be useful to prove some results of the next section. These lemmas are stated and proved for vector spaces, but of course nothing changes essentially if vector bundles are considered instead.

Lemma 4.3 *Let $f\colon E \to F$ be a linear map between vector spaces, and $E_\circ \subset E$ and $F_\circ \subset F$ vector subspaces. Denote $j\colon E_\circ \to E$ the inclusion, $p\colon F \to F/F_\circ$ the projection to the quotient, and consider the composition $\bar{f} = p \circ f \circ j$. Then:*

1. *$\bar{f}$ is injective iff $E_\circ \cap f^{-1}(F_\circ) = \{0\}$.*
   Assuming f injective, this also amounts to $f(E_\circ) \cap F_\circ = \{0\}$.
2. *$\bar{f}$ is surjective iff $f(E_\circ) + F_\circ = F$.*
   Assuming f surjective, this also amounts to $E_\circ + f^{-1}(F_\circ) = E$.

3. *When f is surjective, $\bar{f}$ is bijective iff $E_\circ \oplus f^{-1}(F_\circ) = E$.*
When f is injective, $\bar{f}$ is bijective iff $f(E_\circ) \oplus F_\circ = F$.

Proof. First note that

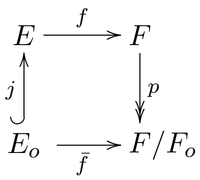
$$\operatorname{Ker} \bar{f} = E_\circ \cap f^{-1}(F_\circ), \quad \operatorname{Im} \bar{f} = (f(E_\circ) + F_\circ)/F_\circ. \tag{4.5}$$

These equalities are clear: the kernel is constituted by the vectors in $E_\circ$ mapped to $F_\circ$ by f, and the image of a subspace $F' \subset F$ by p is $(F' + F_\circ)/F_\circ$. This readily yields the first assertions about injectivity and surjectivity.

Their equivalent formulations when f is injective [or surjective] can be proved using the formulas for $f(E_1 \cap E_2)$ and $f^{-1}(F_1 \cap F_2)$ [or for the sum], as well as $f^{-1}(f(E_\circ)) = E_\circ + \operatorname{Ker} f$, $f(f^{-1}(F_\circ)) = F_\circ \cap \operatorname{Im} f$.

Finally, the assertions about the bijectivity of $\bar{f}$ are a trivial consequence of the other ones. ∎

The elements of this lemma are showed in this diagram:



Now let us study a linear equation on $E_\circ$ defined as in the preceding lemma by $\bar{f}$ and the class of an element $b \in F$. Recall that a linear equation $f(x) = b$ is *consistent* iff $b \in \operatorname{Im} f$.

Lemma 4.4 *The linear equation $\bar{f}(x) = \bar{b}$ is equivalent to the couple of equations $f(x) - b \in F_\circ$, $x \in E_\circ$. It is consistent iff $b \in f(E_\circ) + F_\circ$; in this case the solution is unique iff $E_\circ \cap f^{-1}(F_\circ) = \{0\}$.* ∎

The proof of this lemma is straightforward.

Finally, let $E \subset G$ be a subspace of a vector space. Recall that the *annihilator* (or orthogonal) of E is the subspace

$$E^{\vdash} = \{\gamma \in G^* \mid (\forall x \in E)\ \langle \gamma, x \rangle = 0\} \subset G^*.$$

This space has a close relationship with G/E. Indeed, the transpose map of $G \to G/E$ defines a canonical isomorphism

$$\delta \colon (G/E)^* \to E^{\vdash},$$

such that, for $\alpha \in E^{\vdash}$ and $z \in G$, $\langle \delta^{-1}(\alpha), z + E \rangle = \langle \alpha, z \rangle$.

Lemma 4.5 *Let $E, F \subset G$ be vector subspaces. Let $(\alpha^1, \dots, \alpha^p)$ be a basis for the annihilator $E^{\vdash} \subset G^*$, and $(v_1, \dots, v_q)$ a basis for F. Consider the matrix $D = (D^i_j)_{\substack{1\le i\le p\\ 1\le j\le q}}$ with elements $D^i_j = \langle \alpha^i, v_j \rangle$. Then:*

1. *$E + F = G$ iff $\operatorname{rank} D = p$.*
2. *$E \cap F = \{0\}$ iff $\operatorname{rank} D = q$.*
3. *$E \oplus F = G$ iff D is square invertible.*

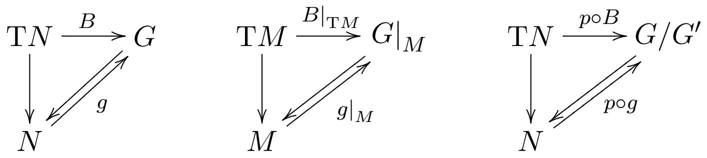
Proof. Consider the linear map $\varepsilon\colon F \to G/E$ defined as the composition of the inclusion $F \hookrightarrow G$ and the projection to the quotient $G \twoheadrightarrow G/E$. It is clear that $E + F = G$ iff ε is surjective, and $E \cap F = \{0\}$ iff ε is injective, so the only thing to prove is that the given matrix is the matrix D of ε in appropriate bases: the basis (v_j) for F, and the basis $(\bar\alpha_i)$, the dual basis of $\bar\alpha^i = \delta^{-1}(\alpha^i)$, for G/E.

Then, if $\varepsilon(v_j) = \bar\alpha_i \, D^i_j$, we have $D^i_j = \langle \bar\alpha^i, \varepsilon(v_j) \rangle = \langle \alpha^i, v_j \rangle$, which is what we wanted to prove. ∎

4.3 Generalized nonholonomic systems

The geometric setting

Among the various operations that can be performed with a linearly singular system $(B\colon \mathrm{T}N \to G, g)$, we are especially interested in the subsystem defined on a submanifold $j\colon M \hookrightarrow N$, and the projection $p\colon G \to G/G'$ to a quotient with respect to a vector subbundle $G' \subset G$:



Suppose that the original system admits solutions Y on a submanifold $N_{\mathrm{f}} \subset N$. Then the subsystem on M has solutions on the submanifolds of $M \cap N_{\mathrm{f}}$ over which a solution Y of the initial system is tangent.

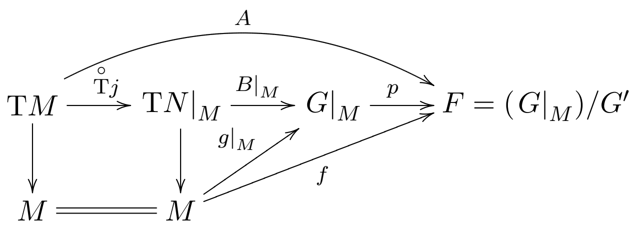
On the other hand, the quotient system has, in general, more solutions than the initial system: if Z is any vector field on N tangent to N_{f} with values in $B^{-1}(G')$ then $Y + Z$ is a solution of the quotient system on N_{f}; there may also exist solutions defined on a submanifold greater than N_{f}.

The dynamics of systems with nonholonomic constraints is a mixture of both constructions: the presence of some constraints, combined with a certain degree of arbitrariness expressed through some multipliers. This combination may result advantageous: though in general Y is not tangent to the submanifold M, it may happen that for some vector fields Z in $B^{-1}(G')$ one has solutions $Y + Z$ tangent to M, or at least to a "big" submanifold of M.

We will call a *generalized nonholonomic system* the linearly singular system $(A\colon \mathrm{T}M \to F, f)$ defined from $(B\colon \mathrm{T}N \to G, g)$ by a *constraint submanifold* $M \subset N$ and a *subbundle of constraint forces* $G' \subset G|_M$ as follows:

- $F = (G|_M)/G'$,
- $A = p \circ B|_M \circ \mathring{\mathrm{T}}j$, and
- $f = p \circ g|_M$,

where $p\colon G|_M \to (G|_M)/G'$ is the projection to the quotient, and $\mathring{\mathrm{T}}j$ denotes the tangent map of j with the image restricted to M. All this is shown in the following diagram:



Regularity and consistency

Before discussing the equations of motion, we want to study some general properties of the generalized nonholonomic system $(A\colon \mathrm{T}M \to F, f)$, namely, whether A is surjective (we will also say that the system is surjective) or bijective (the system is regular), or the equation $A \circ X = f$ is everywhere consistent.

Let us denote

$$H = B^{-1}(G') \subset \mathrm{T}N|_M,$$

which is a vector subbundle whenever the morphism B has constant rank.

Proposition 4.6 *With the preceding notations, the generalized nonholonomic system is surjective iff*

$$B(\mathrm{T}M) + G' = G|_M.$$

Assuming that the original system is surjective, the nonholonomic system is surjective iff

$$\mathrm{T}M + H = \mathrm{T}N|_M,$$

and it is regular iff in addition

$$\mathrm{T}M \cap H = \{0\}.$$

Proof. We want to decide whether $A = p \circ B|_M \circ \mathrm{T}\overset{\circ}{j}$ (the composition of an inclusion, a morphism and a projection) is surjective or injective, and this is given by lemma (4.3).■

The preceding result could be refined also in the case where B is injective, but this does not seem so interesting. As an immediate consequence, we have:

Corollary 4.7 *Suppose that the original system is surjective (or, more particularly, regular). Then the generalized nonholonomic system is regular iff*

$$\mathrm{T}N|_M = \mathrm{T}M \oplus H. \qquad \blacksquare$$

These relations can be given a more concrete form in terms of constraints and frames. Consider a local basis $(\Gamma_\mu)_{1 \le \mu \le m_\circ}$ of sections for the subbundle $H \subset \mathrm{T}N|_M$ (they are vector fields in N, but defined only on M). Consider also a set of $a_\circ$ constraints ϕ^α, linearly independent at each point, that locally define the submanifold $M \subset N$. Finally, consider the matrix

$$D^\alpha_\mu = \langle \mathrm{d}\phi^\alpha|_M, \Gamma_\mu \rangle = \Gamma_\mu \cdot \phi^\alpha, \tag{4.6}$$

whose elements are functions on M.

Proposition 4.8 *With the preceding notations,*

1. $\mathrm{T}M \cap H = 0$ *iff* $\mathrm{rank}(D^\alpha_\mu) = m_\circ$.
2. $\mathrm{T}M + H = \mathrm{T}N|_M$ *iff* $\mathrm{rank}(D^\alpha_\mu) = a_\circ$.
3. $\mathrm{T}M \oplus H = \mathrm{T}N|_M$ *iff* (D^α_μ) *is a square invertible matrix.*

Proof. It is a consequence of lemma (4.5), since the $\mathrm{d}\phi^\alpha|_M$ constitute a basis for the annihilator of $\mathrm{T}M$ in $(\mathrm{T}N|_M)^*$. ■

The connection of such a matrix with the notion of regularity and consistency of a constrained system was already noted in [CR 93, LM 96].

Equations of motion

From the definition of the generalized nonholonomic system $(A\colon \mathrm{T}M \to F, f)$, it is clear that a path $\xi\colon I \to N$ is a solution of the equation of motion iff it is contained in M and

$$B \circ \dot{\xi} - g \circ \xi \in G'. \tag{4.7}$$

If some sections Δ_ν constitute a frame for G', then this equation can be written as

$$B \circ \dot{\xi} = g \circ \xi + \sum_\nu v^\nu \, \Delta_\nu \circ \xi, \tag{4.8}$$

for some multipliers $v^\nu(t)$.

In the same way, for a submanifold $S \subset M$ and a vector field X on M tangent to S, the equation of motion $A \circ X \underset{S}{\simeq} f$ can be written as

$$B \circ X - g \underset{S}{\in} G', \tag{4.9}$$

where the equation must only hold on the points of S. This equation may be also written as

$$B \circ X \underset{S}{\simeq} g + \sum_{\nu} v^{\nu} \Delta_{\nu}, \tag{4.10}$$

for some multipliers $v^{\nu}(x)$.

Of course, we can apply the constraint algorithm to find the solutions of this linearly singular system. However, there is an alternative way to solve the problem when the original problem is regular, or at least consistent. Under this hypothesis, let Y be a vector field on N, solution of the equation of motion of the linearly singular system $(B\colon \mathrm{T}N \to G, g)$:

$$B \circ Y = g.$$

(For most applications the original system is regular, and then the unique solution of this equation is the vector field $Y = B^{-1} \circ g$.)

Using Y, the equations of motion become

$$\dot{\xi} - Y \circ \xi \in H \tag{4.11}$$

for a path ξ in M, and

$$X - Y \underset{S}{\subset} H, \tag{4.12}$$

for a vector field X on M that should be tangent to S.

These equations can be expressed in a more concrete form in terms of the local basis (Γ_{μ}) of sections for the subbundle $H \subset \mathrm{T}N|_M$:

$$\dot{\xi} = Y \circ \xi + \sum_{\mu} u^{\mu}\, \Gamma_{\mu} \circ \xi, \tag{4.13}$$

for some functions $u^{\mu}(t)$, and

$$X \underset{S}{\simeq} Y + \sum_{\mu} u^{\mu}\, \Gamma_{\mu}, \tag{4.14}$$

for some functions u^{μ} on M.

Let us examine whether this last equation has solutions. The requirement for X of being tangent to M is $X \cdot \phi^{\alpha} \underset{M}{\simeq} 0$, which reads

$$\sum_{\mu} D^{\alpha}_{\mu}\, u^{\mu} + Y \cdot \phi^{\alpha} \underset{M}{\simeq} 0, \tag{4.15}$$

where (D^{α}_{μ}) is the matrix defined by (4.6). From this it is clear that the generalized nonholonomic system is regular iff the matrix (D^{α}_{μ}) is invertible on M, and in this case the equation (4.15) directly determines the functions u^μ that give the solution X expressed in (4.14). More generally, the nonholonomic system has solutions if the matrix (D^α_μ) has rank $a_\circ$.

Geometrically, the decomposition $\mathrm{T}N|_M = \mathrm{T}M \oplus H$ stated in Corollary 4.7 has two associated projectors $\mathcal{P}$, $\mathcal{Q}$. Writing $Y = \mathcal{P} \circ Y + \mathcal{Q} \circ Y$ on M, the following result is clear:

Proposition 4.9 *With the preceding notations, if the original system is consistent, with a solution Y, and the generalized nonholonomic system is regular, with solution X, the latter can be obtained as*

$$X = \mathcal{P} \circ Y|_M. \tag{4.16}$$

■

Such projectors were studied, in the context of nonholonomic lagrangian systems, in [LM 96].

4.4 Symmetries and constants of motion

Let us consider a generalized nonholonomic system $(A\colon \mathrm{T}M \to F, f)$, obtained from a linearly singular system $(B\colon \mathrm{T}N \to G, g)$ by means of a restriction to a submanifold $M \subset N$ and a projection to the quotient $p\colon G|_M \to (G|_M)/G'$, where $G' \subset G|_M$ is a vector subbundle.

Recall the definitions of symmetry and infinitesimal symmetry given in section 3.1. Our aim is to study the relation between the symmetries of the original linearly singular system on N and the symmetries of the generalized nonholonomic system on M. In the next proposition, we give sufficient conditions on a symmetry of the original system in order to define a symmetry of the constrained system:

Proposition 4.10 *Let (ψ, Ψ) be a symmetry of $(B\colon \mathrm{T}N \to G, g)$. Suppose that ψ leaves the submanifold $M \subset N$ invariant, and Ψ leaves the subbundle $G' \subset G|_M$ invariant. Then (φ, Φ), where $\varphi = \psi|_M$, and $\Phi\colon (G|_M)/G' \to (G|_M)/G'$ is the map induced on the quotient from Ψ, is a symmetry of $(A\colon \mathrm{T}M \to F, f)$.*

Proof. We have

$$A \circ \mathrm{T}\varphi = p \circ B \circ \mathrm{T}j \circ \mathrm{T}(\psi|_M) = p \circ B \circ \mathrm{T}\psi \circ \mathrm{T}j = p \circ \Psi \circ B \circ \mathrm{T}j = \Phi \circ p \circ B \circ \mathrm{T}j = \Phi \circ A,$$

and

$$f \circ \varphi = p \circ g \circ \psi|_M = p \circ \Psi \circ g|_M = \Phi \circ p \circ g|_M = \Phi \circ f,$$

so the two conditions for being a symmetry are satisfied. ■

We can obtain a similar result for infinitesimal symmetries, by making use of their infinitesimal characterization (3.13):

Proposition 4.11 *Let $(V, \bar{V})$ be an infinitesimal symmetry of $(B\colon \mathrm{T}N \to G, g)$. Suppose that V is tangent to the submanifold $M \subset N$, and $\bar{V}$ is tangent to the subbundle $G' \subset G|_M$. Then $(U, \bar{U})$, where $U = V|_M$ and $\bar{U}\colon (G|_M)/G' \to \mathrm{T}((G|_M)/G')$ is the vector field induced on the quotient from $\bar{V}$, is an infinitesimal symmetry of $(A\colon \mathrm{T}M \to F, f)$.*

Proof. The proof runs as in proposition 4.10:

$$\mathrm{T}f \circ U = \mathrm{T}p \circ \mathrm{T}g \circ V|_M = \mathrm{T}p \circ \bar{V} \circ g|_M = \bar{U} \circ p \circ g|_M = \bar{U} \circ f,$$

$$\begin{aligned} \mathrm{T}A \circ U^T &= \mathrm{T}p \circ \mathrm{T}B \circ \mathrm{T}(\mathrm{T}j) \circ (V^T)\big|_{\mathrm{T}M} = \mathrm{T}p \circ \mathrm{T}B \circ V^T \circ \mathrm{T}j = \\ &= \mathrm{T}p \circ \bar{V} \circ B \circ \mathrm{T}j = \bar{U} \circ p \circ B \circ \mathrm{T}j = \bar{U} \circ A. \end{aligned}$$

■

We now consider constants of motion. Suppose that the original system has a solution $Y \in \mathfrak{X}(N)$, and let us consider a function $h \in \mathrm{C}^\infty(N)$ such that $Y \cdot h = 0$. Under which conditions is $h|_M$ a constant of motion of the generalized nonholonomic system?

Suppose that both the original system and the nonholonomic system are regular, so that $\mathrm{T}N|_M = \mathrm{T}M \oplus H$; let $\mathcal{P}$ be the projector to the first factor, which, according to proposition 4.9, relates the dynamics of both systems as $X = \mathcal{P} \circ Y$. Then we have a simple characterization:

Proposition 4.12 *With the preceding hypothesis, write $X = Y - \Gamma$, where Γ is a section of $H \subset \mathrm{T}N|_M$. Let h be a constant of motion of the unconstrained system. Then $h|_M$ is a constant of motion of the generalized nonholonomic system iff $\Gamma \cdot h = 0$.*

Proof. It is straightforward:

$$X \cdot h = (Y - \Gamma) \cdot h = Y \cdot h - \Gamma \cdot h.$$

(Note that Y and Γ, considered as sections of $\mathrm{T}N|_M$, map functions on N to functions on M.) ■

4.5 Nonholonomic Lagrangian systems revisited

In this section we will show that the dynamics of a nonholonomic lagrangian system as described in section 4.1 falls into the class of generalized nonholonomic systems of section 4.3. We will consider the case of the lagrangian being regular, which amounts to ω_L being a symplectic form.

Consider a nonholonomic lagrangian system (M, L, C). We will consider only the case where the constraint submanifold C restricts the velocities, not the configuration coordinates. In a more formal way, this is described by the conditions given in the next proposition:

Proposition 4.13 *Let $C \subset \mathrm{T}M$ be a submanifold. The following conditions are equivalent:*

1. *The projection $C \to M$ (restriction of the tangent bundle projection $\tau_M\colon \mathrm{T}M \to M$) is a submersion.*
2. $(\mathrm{T}C)^{\vdash} \cap \mathrm{Sb}(\mathrm{T}M)\big|_C = 0$.
3. *The submanifold $C \subset \mathrm{T}M$ can be locally described by the vanishing of some constraints ϕ^i whose fibre derivatives $\mathcal{F}\phi^i$ are linearly independent at each point of C.*
4. *The submanifold $C \subset \mathrm{T}M$ can be locally described by the vanishing of some constraints ϕ^i such that the 1-forms $\Delta^i = {}^t S \circ \mathrm{d}\phi^i$ are linearly independent at each point of C.*

Proof.

$1 \Leftrightarrow 2$:

Let $i\colon C \hookrightarrow \mathrm{T}M$ denote the inclusion and $p\colon C \to M$ denote the projection.

We have that, for $v \in C \subset \mathrm{T}M$, $(\mathrm{T}_v C)^{\vdash} = \mathrm{Ker}\,\mathrm{T}_v^* i$ and $\mathrm{Sb}_v(\mathrm{T}M) = \mathrm{Im}\,\mathrm{T}_v^* \tau_M$. Since $p = \tau_M \circ i$, $\mathrm{T}_v p = \mathrm{T}_v \tau_M \circ \mathrm{T}_v i$ and, passing to the dual, $\mathrm{T}_v^* p = \mathrm{T}_v^* i \circ \mathrm{T}_v^* \tau_M$. Using this fact, it can be seen that $\mathrm{Ker}\,\mathrm{T}_v^* i \cap \mathrm{Im}\,\mathrm{T}_v^* \tau_M = \mathrm{T}_v^* \tau_M(\mathrm{Ker}\,\mathrm{T}_v^* p)$.

So, $(\mathrm{T}_v C)^{\vdash} \cap \mathrm{Sb}_v(\mathrm{T}M) = \mathrm{T}_v^* \tau_M(\mathrm{Ker}\,\mathrm{T}_v^* p)$, and, since $\mathrm{T}_v^* \tau_M$ is injective, $(\mathrm{T}_v C)^{\vdash} \cap \mathrm{Sb}_v(\mathrm{T}M) = 0$ if and only if $\mathrm{Ker}\,\mathrm{T}_v^* p = 0$, or, passing to the dual, if and only if $\mathrm{T}_v p$ is surjective.

This argument is valid for every $v \in C$, so $(\mathrm{T}C)^{\vdash} \cap \mathrm{Sb}(\mathrm{T}M)\big|_C = 0$ if and only if p is a projection.

$2 \Leftrightarrow 4$:

Take $v \in C$. We know that $\mathrm{Sb}_v(\mathrm{T}M)$ is the kernel of ${}^t S_v\colon \mathrm{T}_v^*(\mathrm{T}M) \to \mathrm{T}_v^*(\mathrm{T}M)$. Hence, $(\mathrm{T}_v C)^{\vdash} \cap \mathrm{Sb}_v(\mathrm{T}M) = (\mathrm{T}_v C)^{\vdash} \cap \mathrm{Ker}\,{}^t S_v = \mathrm{Ker}({}^t S\big|_{(\mathrm{T}_v C)^{\vdash}})$, so $(\mathrm{T}_v C)^{\vdash} \cap \mathrm{Sb}_v(\mathrm{T}M) = 0$ if and only if ${}^t S\big|_{(\mathrm{T}_v C)^{\vdash}}$ is injective.

Let ϕ^i be a set of functions such that $\mathrm{d}\phi^i_v$ are linearly independent and C is locally described around v by the vanishing of the ϕ^i. By theorem 2.4 such constraints functions always exist and $(\mathrm{T}_v C)^{\vdash}$ is generated by $\langle \mathrm{d}\phi^i_v \rangle$. Therefore, ${}^t S\big|_{(\mathrm{T}_v C)^{\vdash}}$ is injective if and only if ${}^t S(\mathrm{d}\phi^i_v)$ are linearly independent.

Since we can prove this for every $v \in C$, $(\mathrm{T}_v C)^{\vdash} \cap \mathrm{Sb}_v(\mathrm{T}M) = 0$ if and only if $\Delta^i = {}^t S \circ \mathrm{d}\phi^i$ are linearly independent at each point of C.

$3 \Leftrightarrow 4$:

For every $v_q \in \mathrm{T}M$, $\mathrm{T}_v^* \tau_M\colon \mathrm{T}_q^* M \to \mathrm{Sb}_v(\mathrm{T}M)$ is a linear isomorphism. Therefore, since $\Delta^i(v) = \mathrm{T}_v^* \tau_M(\mathcal{F}\phi^i(v))$, the cotangent vectors $\Delta^i(v)$ are linearly independent if and only if the corresponding $\mathcal{F}\phi^i(v)$ are linearly independent.

This holds, in particular, for $v \in C$.

■

In coordinates, these conditions mean that $\left(\dfrac{\partial \phi^i}{\partial v^k}\right)$ has maximal rank.

Therefore, from now on, *we assume that the projection $C \to M$ is a surjective* submersion.

For the sake of simplicity, we use the notation $G' = {}^t S((\mathrm{T}C)^\vdash) \subset \mathrm{Sb}(\mathrm{T}M)|_C$ for the Chetaev bundle and $H = \widehat{\omega}_L^{-1}(G') \subset \mathrm{V}(\mathrm{T}M)|_C$ its image by $\widehat{\omega}_L^{-1}$. Suppose that $M \subset \mathrm{T}Q$ is defined by the vanishing of some independent constraints ϕ^i as in the preceding proposition. Then $(\mathrm{T}C)^\vdash$ is spanned by the $\mathrm{d}\phi^i|_C$. We denote by Δ^i and Γ_i their corresponding images in G' (through ${}^t S$) and H (through $\widehat{\omega}_L^{-1}$).

The following diagram shows all these objects:

$$
\begin{array}{ccccccc}
\mathrm{T}C & \hookrightarrow & \mathrm{T}(\mathrm{T}M)|_C & \xrightarrow{\widehat{\omega}_L} & \mathrm{T}^*(\mathrm{T}M)|_C & \hookleftarrow & (\mathrm{T}C)^\vdash = \langle \mathrm{d}\phi^i|_C \rangle \\
 & & \Big\updownarrow S & & \Big\updownarrow {}^t S & & \\
\langle \Gamma_i \rangle = H & \hookrightarrow & \mathrm{V}(\mathrm{T}M)|_C & \xrightarrow{\widehat{\omega}_L} & \mathrm{Sb}(\mathrm{T}M)|_C & \hookleftarrow & G' = \langle \Delta^i \rangle
\end{array}
$$

So we have two subbundles $\mathrm{T}C, H \subset \mathrm{T}(\mathrm{T}M)|_C$. We have $\operatorname{rank} \mathrm{T}C = m$ and $\operatorname{rank}(\mathrm{T}C)^\vdash = n - m$; the conditions in proposition 4.13 imply also that $\operatorname{rank} H = \operatorname{rank} G' = n - m$.

Theorem 4.14 *The nonholonomic lagrangian system (M, L, C) is equivalent to the generalized nonholonomic system defined from the lagrangian system $(\widehat{\omega}_L \colon \mathrm{T}(\mathrm{T}M) \to \mathrm{T}^*(\mathrm{T}M), \mathrm{d}E_L)$ by the constraint submanifold $C \subset \mathrm{T}M$ and the Chetaev bundle $G' = {}^t S((\mathrm{T}C)^\vdash) \subset \mathrm{T}^*(\mathrm{T}M)|_C$.*

$$
\begin{array}{ccccccc}
\mathrm{T}C & \xrightarrow{\mathring{\mathrm{T}}j} & \mathrm{T}(\mathrm{T}M)|_C & \xrightarrow{\widehat{\omega}_L|_C} & \mathrm{T}^*(\mathrm{T}M)|_C & \longrightarrow & \mathrm{T}^*(\mathrm{T}M)|_C / G' \\
\big\downarrow & & \big\downarrow & \nearrow_{\mathrm{d}E_L|_C} & & & \\
C & = & C & & & &
\end{array}
$$

Proof. The equation of motion for a path $\xi = \dot{\gamma}$ such that $\xi(t) \in M$ is

$$
\dot{\xi} = X_L \circ \xi + \sum_i u^i \, \Gamma_i \circ \xi. \tag{4.17}
$$

Instead, let us write the equations of motion for vector fields: according to (4.9), for a second-order vector field X on $\mathrm{T}M$, tangent to C, the equation is

$$
i_X \omega_L - \mathrm{d}E_L \underset{C}{\in} G', \tag{4.18}
$$

which is equation (4.3).

■

If, in addition to $(\mathrm{T}C)^\vdash \cap \mathrm{Sb}(\mathrm{T}M)|_C = 0$, we have $\mathrm{T}C \cap H = 0$, then $\mathrm{T}(\mathrm{T}M)|_C = \mathrm{T}C \oplus H$, and so there is a unique solution X of the equation of motion, which can be obtained from Y through the projector to $\mathrm{T}C$ as described by proposition 4.9.

4.6 Implicit hamiltonian systems

In this section we will show that implicit hamiltonian systems can be considered as generalized nonholonomic systems. Thus we can apply the techniques and results studied so far to this kind of systems.

Firstly, we will introduce the basic concepts regarding implicit hamiltonian systems and its underlying geometric structures: the Dirac structures. For more details we refer to [Bla 00, Cou 90, DS 99, Dor 93].

Dirac structures

Let M be a manifold with dimension n. Consider a vector field $X \in \mathfrak{X}(M)$ and a one-form $\alpha \in \Omega^1(M)$. We say that the pair (X, α) belongs to a vector subbundle $\mathcal{D} \subset \mathrm{T}M \oplus \mathrm{T}^*M$ (denoted $(X, \alpha) \in \mathcal{D}$) if $(X(x), \alpha(x)) \in \mathcal{D}(x)$ for every point x in M.

A *Dirac structure* on M is a vector subbundle $\mathcal{D} \subset \mathrm{T}M \oplus \mathrm{T}^*M$ such that $\mathcal{D} = \mathcal{D}^\perp$, where

$$\mathcal{D}^\perp = \{(Y, \beta) \in \mathrm{T}M \oplus \mathrm{T}^*M \mid \langle \alpha, Y\rangle + \langle \beta, X\rangle = 0, \text{ for all } (X, \alpha) \in \mathcal{D}\}.$$

The notation $\mathcal{D}^\perp$ is justified by the fact that $\mathcal{D}^\perp$ is the orthogonal to $\mathcal{D}$ with respect to the symmetric pairing $\langle\langle \cdot, \cdot \rangle\rangle$ on $\mathrm{T}M \oplus \mathrm{T}^*M$ defined by

$$\langle\langle (X, \alpha), (Y, \beta)\rangle\rangle = \langle \alpha, Y\rangle + \langle \beta, X\rangle, \quad (X, \alpha), (Y, \beta) \in \mathrm{T}M \oplus \mathrm{T}^*M.$$

Let us remark that in [Bla 00, DS 99], this is called a *generalized* Dirac structure, and keep the nomenclature "Dirac structure" only for those that satisfy some closedness (or integrability) condition.

It can be seen that a Dirac structure has constant dimension n. The next two examples of Dirac structure show the connection with symplectic and Poisson systems:

Example 4.15 *Let $\omega \in \Omega^2(M)$ be a nondegenerate two-form on M (in case that ω is closed, it is a symplectic form). Then*

$$\mathcal{D} = \{(X, \alpha) \in \mathrm{T}M \oplus \mathrm{T}^*M \mid \alpha = i_X\omega\}$$

is a Dirac structure on M.

Example 4.16 *Let Λ be a bivector field on M (i.e., an antisymmetric $(2,0)$-tensor field). Then*

$$\mathcal{D} = \{(X, \alpha) \in \mathrm{T}M \oplus \mathrm{T}^*M \mid X = i_\alpha\Lambda\}$$

is a Dirac structure on M.

Recall that Λ defines an almost-Poisson bracket by

$$\{f, g\} = \Lambda(\mathrm{d}f, \mathrm{d}g).$$

This is a Poisson bracket (satisfies the Jacobi identity) if $[\Lambda, \Lambda] = 0$, where $[\cdot, \cdot]$ denotes the Schouten-Nijenhuis bracket [Vai 94].

A Dirac structure $\mathcal{D}$ defines the following distributions and codistributions:

- $G_0 = \{X \in \mathfrak{X}(M) \mid (X, 0) \in \mathcal{D}\}$
- $G_1 = \{X \in \mathfrak{X}(M) \mid \exists\, \alpha \in \Omega^1(M) \text{ such that } (X, \alpha) \in \mathcal{D}\}$
- $P_0 = \{\alpha \in \Omega^1(M) \mid (0, \alpha) \in \mathcal{D}\}$
- $P_1 = \{\alpha \in \Omega^1(M) \mid \exists\, X \in \mathfrak{X}(M) \text{ such that } (X, \alpha) \in \mathcal{D}\}$

We define the annihilator of a distribution G as

$$\operatorname{ann} G = \{\alpha \in \Omega^1(M) \mid \langle \alpha, X \rangle = 0 \text{ for all } X \in G\},$$

and the kernel of a codistribution P as

$$\ker P = \{X \in \mathfrak{X}(M) \mid \langle \alpha, X \rangle = 0 \text{ for all } \alpha \in P\}.$$

It follows that $G_0 = \ker P_1$ and $P_0 = \operatorname{ann} G_1$. Moreover, $G_1 = \ker P_0$ and $P_1 = \operatorname{ann} G_0$, with equality if and only if G_1, respectively P_1, is constant dimensional.

Implicit hamiltonian systems

As a second ingredient (together with a Dirac structure) to define an implicit hamiltonian system we need a function $H \in \mathrm{C}^\infty(M)$, called the Hamiltonian function, which represents the energy of the system.

The *implicit hamiltonian system* on M corresponding to $\mathcal{D}$ and H is specified by the equation

$$(\dot\gamma(t), \mathrm{d}H(\gamma(t))) \in \mathcal{D}(\gamma(t)), \quad \text{for each } t \in I, \tag{4.19}$$

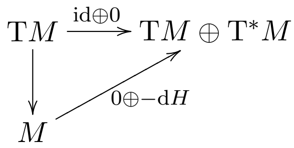
where the solutions are paths $\gamma\colon I \to M$ with domain an interval $I \subset \mathbf{R}$.

It is clear that if it exists a solution of the implicit hamiltonian system passing through a point $x \in M$, then $\mathrm{d}H(x)$ belongs to the co-distribution P_1. This is an algebraic constraint which defines a submanifold of M:

$$M_1 = \{x \in M \mid \mathrm{d}H(x) \in P_1(x)\}.$$

We have then a system with some constraints and certain degree of arbitrariness expressed by equation (4.19). Therefore, an implicit hamiltonian system is a generalized nonholonomic system, as the following proposition states.

Proposition 4.17 *The implicit hamiltonian system on a manifold M corresponding to a Dirac structure $\mathcal{D}$ and a Hamiltonian H is equivalent (has the same solutions) to* *the generalized nonholonomic system defined from the linearly singular system* $(\mathrm{id} \oplus 0\colon \mathrm{T}M \to \mathrm{T}M \oplus \mathrm{T}^*M, 0 \oplus -\mathrm{d}H)$



by the submanifold $M_1 = \{x \in M \mid \mathrm{d}H(x) \in P_1(x)\}$ *and the subbundle* $\mathcal{D} \subset \mathrm{T}M \oplus \mathrm{T}^*M$.

Proof. We use the notation $p\colon \mathrm{T}M \oplus \mathrm{T}^*M \to (\mathrm{T}M \oplus \mathrm{T}^*M)/\mathcal{D}$ for the projection. A path $\xi\colon I \to M_1$ is solution of the generalized nonholonomic system if

$$p \circ (\mathrm{id} \oplus 0) \circ \dot{\xi} = p \circ (0 \oplus -\mathrm{d}H) \circ \xi,$$

which is equivalent to

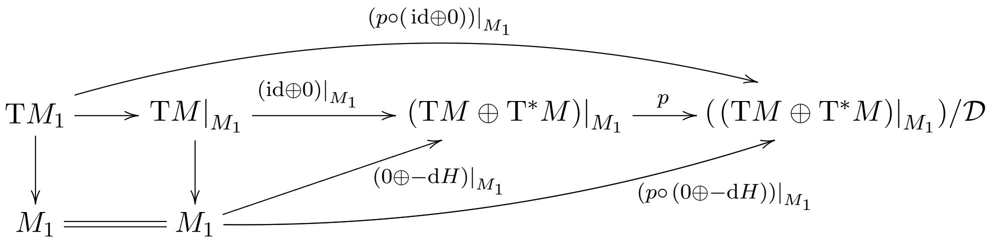
$$p \circ (\dot{\xi} \oplus (\mathrm{d}H \circ \xi)) = 0,$$

or

$$(\dot{\xi}, \mathrm{d}H \circ \xi) \in \mathcal{D}.$$

This is equation (4.19) plus the requirement that the path is in the submanifold M_1. Since we have already seen that all the solutions of the implicit hamiltonian system are in M_1, the result is proved. ∎

The generalized nonholonomic system equivalent to the implicit hamiltonian system is shown in the diagram:



(4.20)

Let us now prove a result about regularity of implicit hamiltonian systems, given in [vdS 98], using the theory of generalized nonholonomic systems.

Proposition 4.18 *Consider the implicit hamiltonian system* $(M, \mathcal{D}, H)$. *Assume that* P_1 *is constant dimensional and* $\{\Gamma_1, \ldots, \Gamma_m\}$ *is a basis of vector fields for* G_0.

Then the implicit hamiltonian system reduces to an explicit one on M_1 *if and only if the matrix* $(\Gamma_i \cdot \Gamma_j \cdot H(x))_{i,j=1,\ldots m}$ *is invertible for all* $x \in M_1$.

Proof. The implicit hamiltonian system reduces to an explicit one on M_1 if and only if the system has one and only one solution passing through every point in M_1. Taking into account the equivalence stated in proposition (4.17), that is to say that the generalized nonholonomic system (4.20) is regular.

Therefore, we can use the results of section 4.3. With the notations of that section, the vector subbundle H is

$$H = (\mathrm{id} \oplus 0)^{-1}(\mathcal{D}) = G_0,$$

and a basis of sections of H is $(\Gamma_i)_{1\leq i\leq m}$.

Since the manifold M_1 is defined by the equation $\mathrm{d}H \in P_1$, $\ker P_1 = G_0$ and $\{\Gamma_1, \ldots, \Gamma_m\}$ is a basis of vector fields for G_0, a set of linearly independent constraints defining M_1 is

$$\phi^j = \Gamma_j \cdot H, \ 1 \leq j \leq m.$$

Thus, the matrix (4.6) is in this case

$$D_i^j = \Gamma_i \cdot \Gamma_j \cdot H$$

and the result follows directly from proposition 4.8 and corollary 4.7.

■

In [BS 01] is shown that, under the assumption that the distribution G_1 is constant dimensional, a Dirac structure $\mathcal{D}$ on a manifold M induces a Dirac structure $\bar{\mathcal{D}}$ on every submanifold $\bar{M}$ of M such that $G_1(\bar{x}) \cap \mathrm{T}_{\bar{x}}\bar{M}$, $\bar{x} \in \bar{M}$, is constant dimensional. This restricted Dirac structure is given by

$$\begin{aligned}\bar{\mathcal{D}} = \{(\bar{X}, \bar{\alpha}) \in \mathrm{T}\bar{M} \oplus \mathrm{T}^*\bar{M} \mid \exists X \,\text{such that}\; \bar{X} \sim_\iota X \,\text{and}\\ \exists\, \alpha \,\text{such that}\; \bar{\alpha} = \iota^*\alpha \,\text{with}\; (X, \alpha) \in \mathcal{D}\},\end{aligned}$$

where ι denotes the inclusion $\bar{M} \subset M$.

We remark that when the constraint algorithm is applied to the generalized nonholonomic system (4.20), the equations obtained on M_1 and the successive constraint submanifolds that may appear, in general do not correspond to an implicit hamiltonian system associated with the restricted Dirac structures or any other Dirac structures on the constraint submanifolds.

4.7 Relativistic particle with a nonholonomic constraint

In this section we study the motion of a relativistic particle as a nonholonomic constrained system. We will consider two possible lagrangian functions, a regular one (deeply studied in [KM 01]) and a singular one.

Let us consider a particle with mass m and charge e moving in spacetime. We model spacetime as a 4-dimensional manifold Q, endowed with a metric tensor g of signature $(1, 3)$. Suppose furthermore that the particle is subject to the action of an electromagnetic field $F = \mathrm{d}A$, where $A \in \Omega^1(Q)$, and a potential $U \in \mathrm{C}^\infty(Q)$.

Recall that there are some relevant objects associated with the metric g, namely, the isomorphism $\widehat{g}\colon \mathrm{T}Q \to \mathrm{T}^*Q$ (we will denote $X^\flat = \widehat{g} \circ X$)), the Levi-Civita connection ∇, the differential forms $\theta_g = \widehat{g}^*(\theta_Q) \in \Omega^1(\mathrm{T}Q)$ and $\omega_g = \widehat{g}^*(\omega_Q) = -\mathrm{d}\theta_g \in \Omega^2(\mathrm{T}Q)$, the energy $E_g(u_q) = \frac{1}{2}g(u_q, u_q) \in \mathrm{C}^\infty(\mathrm{T}Q)$, and the geodesic vector field S_g, which satisfies $i_{S_g}\omega_g = \mathrm{d}E_g$. We denote $v = \sqrt{2E_g}$.

We will study two different lagrangian functions, namely

$$L_1(u_q) = -mc\, g(u_q, u_q)^{1/2} - \frac{e}{c}\langle A(q), u_q\rangle - U(q),$$

and

$$L_2(u_q) = -\frac{1}{2}m\, g(u_q, u_q) - \frac{e}{c}\langle A(q), u_q\rangle - U(q).$$

Forgetting the potential, L_1 is the singular lagrangian commonly used in relativistic mechanics to describe a particle in an electromagnetic field; it is defined only on the open set of time-like vectors of $\mathrm{T}Q$ (that is, the tangent vectors u_q such that $g(u_q, u_q) > 0$). The lagrangian L_2 appears in [KM 01]. Our aim is to compare both systems, and to introduce the nonholonomic constraint $v^2 = c^2$ to them.

The lagrangians L_1 and L_2 have, respectively, associated Lagrange's 1-forms:

$$\theta_1 = {}^tJ \circ \mathrm{d}L_1 = -\frac{mc}{v}\theta_g - \frac{e}{c}\tau_Q^* A$$

and

$$\theta_2 = {}^tJ \circ \mathrm{d}L_2 = -m\theta_g - \frac{e}{c}\tau_Q^* A;$$

the Lagrange's 2-forms are

$$\omega_1 = -\mathrm{d}\theta_1 = -\frac{mc}{v}\omega_g - \frac{c}{v^2}\mathrm{d}v \wedge \theta_g + \frac{e}{c}\tau_Q^* F$$

and

$$\omega_2 = -\mathrm{d}\theta_2 = -m\omega_g + \frac{e}{c}\tau_Q^* F;$$

and the lagrangian energies are

$$E_1 = \Delta \cdot L_1 - L_1 = U$$

and

$$E_2 = \Delta \cdot L_2 - L_2 = -\frac{1}{2}mv^2 + U.$$

The symplectic formulation of the equations of motion for the lagrangians L_1 and L_2 are, respectively,

$$i_X\omega_1 = \mathrm{d}E_1, \tag{4.21}$$

and

$$i_X\omega_2 = \mathrm{d}E_2, \tag{4.22}$$

for second-order vector fields X. For any 2-form ω, we will also denote $i_X\omega$ by $\widehat{\omega}(X)$.

It is worth writing down the Euler–Lagrange equations of motion for a path γ, which are, for lagrangians L_1 and L_2:

$$\frac{mc}{g(\dot\gamma,\dot\gamma)^{1/2}}\left((\nabla_t\dot\gamma)^\flat - \frac{g(\dot\gamma,\nabla_t\dot\gamma)}{g(\dot\gamma,\dot\gamma)}\dot\gamma^\flat\right) + \frac{e}{c}i_{\dot\gamma}F - \mathrm{d}U = 0, \tag{4.23}$$

and

$$m(\nabla_t\dot\gamma)^\flat + \frac{e}{c}i_{\dot\gamma}F - \mathrm{d}U = 0. \tag{4.24}$$

Let us now consider equations (4.21) and (4.22).

As $\widehat{\omega}_1$ is not surjective, equation (4.21) could have no solutions. We denote by $T = \dot q^i \frac{\partial}{\partial q^i}$ the natural vector field along τ_Q and $\xi^\vee$ the vertical lift of a vector field $\xi\colon \mathrm{T}Q \to \mathrm{T}Q$ along τ_Q. We have that $\operatorname{Ker}\omega_1 = \langle\Delta,\Sigma\rangle$, where

$$\Sigma = S_g - \frac{ev}{mc^2}((i_T F)^\sharp)^\vee. \tag{4.25}$$

We can see that $\widehat{\omega}_1(\frac{v}{mc}(\operatorname{grad} U)^\vee) = \mathrm{d}U - (\frac{1}{v^2}i_T\mathrm{d}U)\theta_g$ and that $\theta_g \notin \operatorname{Im}\widehat{\omega}_1$. Therefore equation (4.21) has solutions if and only if $i_T\mathrm{d}U = 0$, that is, the potential U is constant, which, in practice, is the same as taking U equal to 0.

Since Σ is a second-order vector field, in absence of potential the solutions of equation (4.21) are $X_1 = \Sigma + \mu\Delta$, where μ is an arbitrary function. If, in addition, there is no electromagnetic field, then the solutions are $S_g + \mu\Delta$, and their integral curves are reparametrized geodesics.

On the other hand, equation (4.22) is regular, and its solution is

$$X_2 = S_g + \frac{1}{m}(\operatorname{grad} U)^\vee - \frac{e}{mc}((i_T F)^\sharp)^\vee. \tag{4.26}$$

This can be proved making use of the relations $i_{Z^\vee}\omega_g = -\tau_Q^*(Z^\flat)$ for vector fields Z along τ_Q, and $i_S(\tau_Q^* F) = \tau_Q^*(i_T F)$. In this case, in absence of electromagnetic field and potential, the solutions are the geodesics of g.

Now we introduce the nonholonomic constraint

$$\phi(u_q) := g(u_q,u_q) - c^2 = 0, \tag{4.27}$$

which defines a submanifold $M \subset \mathrm{T}Q$.

The subbundle of constraint forces is $\langle {}^tJ(\mathrm{d}\phi)\rangle\big|_M = \langle\theta_g\rangle|_M$, therefore, according to equation (4.18), the equations of motion for both lagrangians become

$$i_X\omega_1 \underset{M}{\simeq} \mathrm{d}E_1 + \lambda\theta_g, \tag{4.28}$$

and

$$i_X\omega_2 \underset{M}{\simeq} \mathrm{d}E_2 + \lambda\theta_g, \tag{4.29}$$

for second-order vector fields X tangent to M.

Note that if a path γ satisfies the constraint then it also satisfies the equation $0 = \dfrac{\mathrm{d}}{\mathrm{d}t} g(\dot\gamma,\dot\gamma) = 2g(\dot\gamma, \nabla_t\dot\gamma)$, so looking at equations (4.23) and (4.24) we realize that the two *constrained* systems have the *same* equations of motion:

$$\begin{cases} m(\nabla_t\dot\gamma)^\flat + \dfrac{e}{c} i_{\dot\gamma} F - \mathrm{d}U = \lambda\dot\gamma^\flat, \\ g(\dot\gamma,\dot\gamma) = c^2. \end{cases} \tag{4.30}$$

The multiplier λ can be found by contracting the equation with $\dot\gamma$, which gives $\lambda = -\dfrac{1}{c^2} i_{\dot\gamma}\mathrm{d}U$.

We are going to see this equivalence of the solutions of both Euler–Lagrange equations by computing the solutions of equations (4.28) and (4.29).

First let us analyse equation (4.29). From $\Delta\cdot\phi = 2v^2 \underset{M}{\simeq} 2c^2 \neq 0$ and $i_\Delta\omega_2 = m\theta_g$, it follows that $\mathrm{T}M \oplus \widehat\omega_2^{-1}(\langle\theta_g\rangle|_M) = (\mathrm{T}Q)|_M$, so, by proposition 4.6, the system is regular. Its solution is $X = X_2 + \frac{\lambda}{m}\Delta$, where the multiplier λ is found by imposing that X is tangent to M:

$$0 = X\cdot\phi = X_2\cdot\phi + \frac{\lambda}{m}\Delta\cdot\phi \underset{M}{\simeq} \frac{2}{m} i_T\mathrm{d}U + 2\frac{\lambda}{m}c^2. \tag{4.31}$$

Therefore, the solution of the second system is

$$X = S_g + \frac{1}{m}(\operatorname{grad} U)^\vee - \frac{e}{mc}((i_T F)^\sharp)^\vee - \frac{1}{mc^2}(i_T\mathrm{d}U)\Delta. \tag{4.32}$$

Now let us analyse equation (4.28). Since $Y = \frac{1}{m}(\operatorname{grad} U)^\vee - \frac{1}{mc^2}(i_T\mathrm{d}U)\Delta$ is a vector field tangent to M and $\widehat\omega_1(Y) \underset{M}{\simeq} \mathrm{d}U - (\frac{1}{c^2} i_T\mathrm{d}U)\theta_g$, the system is consistent.

We can see that

$$\mathrm{T}M \cap \widehat\omega_1^{-1}(\langle\theta_g\rangle|_M) = \mathrm{T}M \cap \operatorname{Ker}\widehat\omega_1 = \langle\Sigma\rangle|_M, \tag{4.33}$$

so the system is not regular. Then, the solutions of the equation are $Y + \mu\Sigma$. Since Y is vertical, in order to be a second-order vector field the function μ must be equal to one, so the solution is $Y + \Sigma \underset{M}{\simeq} X$, exactly the same as for the lagrangian L_2.

4.8 Examples

Example 1

Consider the differential equation on $N = \mathbf{R}^2$ defined by the vector field $Y = \frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$. We restrict this system to a generalized nonholonomic one by means of the construction of section 4.3, taking the submanifold $M = \mathbf{R}\times\{a\} \subset N$ and the subbundle $C = \langle x\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\rangle \subset \mathrm{T}N|_M$.

In this case $\mathrm{T}N|_M = \mathrm{T}M \oplus C$ and the projectors associated with this decomposition are

$$\begin{array}{rccc} \mathcal{P}\colon & \dfrac{\partial}{\partial x} & \longmapsto & \dfrac{\partial}{\partial x} \\ & \dfrac{\partial}{\partial y} & \longmapsto & -x\dfrac{\partial}{\partial x}, \end{array}$$

$$\mathcal{Q}\colon \begin{array}{ccl} \frac{\partial}{\partial x} & \longmapsto & 0 \\ \frac{\partial}{\partial y} & \longmapsto & x\frac{\partial}{\partial x} + \frac{\partial}{\partial y}. \end{array}$$

Thus $X = \mathcal{P} \circ Y|_M = (1 - ax)\, \frac{\partial}{\partial x}\big|_M$ is the solution of the generalized nonholonomic system.

Let us study the infinitesimal symmetries of both systems. We can see that a vector field $V \in \mathcal{X}(N)$ is an infinitesimal symmetry of the unconstrained system if it has the form $V = V^1(ye^{-x})\frac{\partial}{\partial x} + e^x V^2(ye^{-x})\frac{\partial}{\partial y}$, where V^1 and V^2 are arbitrary smooth functions.

On the other hand, since the constrained system is one-dimensional, its infinitesimal symmetries are the vector fields $U = kX$, with $k \in \mathbf{R}$. Observe that, in principle, an infinitesimal symmetry of Y does not lead to an infinitesimal symmetry of X by restriction to M, even when $Y|_M \in \mathfrak{X}(M)$. Nevertheless, if we also require that $V^T(C) \subset \mathrm{T}C$, then we obtain $V^1(t) = k(1 + a\ln(t/a))$ and $V^2(t) = 0$, so that actually $V|_M = k(1 - ax)\,\frac{\partial}{\partial x}\big|_M$ is an infinitesimal symmetry of X.

Example 2

Here we discuss an example of a particle with a nonholonomic constraint, due to Rosenberg [Ros 77]. This example has been discussed in some papers about reduction, as for instance [BGM 96, BKMM 96, BŚ 93, CL 99]. Consider a particle moving in $\mathbf{R}^3$ with lagrangian function

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

subject to the nonholonomic constraint

$$\phi = \dot{z} - y\dot{x}.$$

Using the notation of section 4.5, we have $N = \mathrm{T}\mathbf{R}^3$, $\omega_L = \mathrm{d}x \wedge \mathrm{d}\dot{x} + \mathrm{d}y \wedge \mathrm{d}\dot{y} + \mathrm{d}z \wedge \mathrm{d}\dot{z}$ and $\mathrm{d}E_L = \dot{x}\mathrm{d}\dot{x} + \dot{y}\mathrm{d}\dot{y} + \dot{z}\mathrm{d}\dot{z}$, so the unconstrained dynamics is the well-known free dynamics described by the vector field

$$X_L = \widehat{\omega}_L^{-1}(\mathrm{d}E_L) = \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + \dot{z}\frac{\partial}{\partial z}.$$

The constraint submanifold is $M = \{\dot{z}{=}y\dot{x}\}$, with tangent bundle

$$\mathrm{T}M = \mathrm{Ker}(\mathrm{d}\phi) = \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y} + \dot{x}\frac{\partial}{\partial \dot{z}}, \frac{\partial}{\partial z}, \frac{\partial}{\partial \dot{x}} + y\frac{\partial}{\partial \dot{z}}, \frac{\partial}{\partial \dot{y}} \right\rangle\Bigg|_M,$$

and the vector subbundle $C \subset \mathrm{T}N|_M$ is

$$C = \langle \widehat{\omega}_L^{-1}({}^tJ(\mathrm{d}\phi)) \rangle = \left\langle y\frac{\partial}{\partial \dot{x}} - \frac{\partial}{\partial \dot{z}} \right\rangle\Bigg|_M.$$

Note that $\mathrm{T}N|_M$ splits as $\mathrm{T}N|_M = \mathrm{T}M \oplus C$, so the only solution X of the constrained lagrangian system is the projection of $X_L|_M$ to $\mathrm{T}M$ according to this decomposition:

$$X = \left(\dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + \dot{z}\frac{\partial}{\partial z} - \frac{y\dot{y}\dot{x}}{y^2+1}\frac{\partial}{\partial \dot{x}} + \frac{\dot{y}\dot{x}}{y^2+1}\frac{\partial}{\partial \dot{z}} \right)\Bigg|_M .$$

We choose $(x, y, z, \dot{x}, \dot{y})$ as coordinates on M. With this system of coordinates, the vector field X reads as

$$X = \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + y\dot{x}\frac{\partial}{\partial z} - \frac{y\dot{y}\dot{x}}{y^2+1}\frac{\partial}{\partial \dot{x}}.$$

Now let us look for the symmetries and constants of motion of both systems. For the free particle, the constants of motion are the functions g on N that are invariant by X_L, $X_L \cdot g = 0$. The solutions of this partial differential equation have the form $G(\dot{x}, \dot{y}, \dot{z}, \dot{x}y - \dot{y}x, \dot{y}z - \dot{z}y)$, where G is an arbitrary function with five variables. On the other hand, the infinitesimal symmetries are the vector fields V that commute with X_L, $[V, X_L] = 0$. They are linear combinations of the six vector fields $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial}{\partial z}$, $x\frac{\partial}{\partial x} + \dot{x}\frac{\partial}{\partial \dot{x}}$, $y\frac{\partial}{\partial y} + \dot{y}\frac{\partial}{\partial \dot{y}}$ and $z\frac{\partial}{\partial z} + \dot{z}\frac{\partial}{\partial \dot{z}}$, with constants of motion as coefficients.

For the constrained particle, the constants of motion (written in coordinates of M), have the form

$$F\left(\dot{y}, \dot{x}\sqrt{y^2+1}, \dot{y}x - \operatorname{arcsinh}(y)\dot{x}\sqrt{y^2+1}, \dot{y}z - \dot{x}(y^2+1)\right), \tag{4.34}$$

and the infinitesimal symmetries are linear combinations of the five vector fields

$$\frac{\partial}{\partial x}, \qquad \frac{\partial}{\partial z}, \qquad \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + y\dot{x}\frac{\partial}{\partial z} - \frac{\dot{x}\dot{y}y}{y^2+1}\frac{\partial}{\partial \dot{x}},$$

$$\frac{\operatorname{arcsinh}(y)}{\dot{y}}\frac{\partial}{\partial x} + \frac{\sqrt{y^2+1}}{\dot{y}}\frac{\partial}{\partial z} + \frac{1}{\sqrt{y^2+1}}\frac{\partial}{\partial \dot{x}},$$

$$\frac{\dot{x}(y - \operatorname{arcsinh}(y)\sqrt{y^2+1})}{\dot{y}^2}\frac{\partial}{\partial x} + \frac{y}{\dot{y}}\frac{\partial}{\partial y} - \frac{\dot{x}}{\dot{y}^2}\frac{\partial}{\partial z} - \frac{\dot{x}y^2}{\dot{y}(y^2+1)}\frac{\partial}{\partial \dot{x}} + \frac{\partial}{\partial \dot{y}},$$

with constants of motion as coefficients.

Finally, we can illustrate proposition 4.12, and show the relation between the constants of motion of both systems. Consider a constant of motion $g = G(\dot{x}, \dot{y}, \dot{z}, \dot{x}y - \dot{y}x, \dot{y}z - \dot{z}y)$ of X_L; its restriction to M will be a constant of motion of X iff $Z \cdot g = 0$, where Z is the section of C

$$Z = X_L|_M - X = \frac{\dot{x}\dot{y}}{y^2+1}\left(y\frac{\partial}{\partial \dot{x}} - \frac{\partial}{\partial \dot{z}} \right)\Bigg|_M .$$

The condition implies that

$$g|_M = H\left(\dot{y}, \sqrt{\dot{z}^2+\dot{x}^2}, \dot{z} + \dot{y}x - \dot{x}y - \operatorname{arcsinh}(\dot{z}/\dot{x})\sqrt{\dot{z}^2+\dot{x}^2}, \dot{y}z - \dot{z}y - \dot{x}\right),$$

and we see that $g|_M$ coincides with (4.34).

Chapter 5

Vector hulls of affine spaces and affine bundles

Every affine space A has a canonical immersion $A \hookrightarrow \widehat{A}$ as a hyperplane in a vector space, the vector hull of A. This immersion satisfies a universal property with respect to the vector-valued affine functions defined on A. In addition, any affine map $f\colon A \to B$ between affine spaces has a canonical prolongation to a linear map $\widehat{f}\colon \widehat{A} \to \widehat{B}$ between their vector hulls, which we call the vector prolongation of f. Both constructions indeed define a functor from the category of affine spaces to the category of vector spaces.

These facts are not widely known, but they are greatly clarifying, both for affine geometry and its applications.

We devote the first sections to perform a systematic study of the properties of the vector hull, regardless of its explicit definition. As we will show, the fact that $A \subset \widehat{A}$ is a hyperplane and the universal property it satisfies are enough to obtain many interesting results. After this, we study a functorial construction of the vector hull of A in terms of a certain set of affine vector fields on A. This set is used in several references as the definition of the vector hull, but we give a very simple characterisation of it that enlightens its vector space structure. This definition of the vector hull is compared with other different ones that can be found in the literature. As we will show, there are several different but equivalent vector hull functors.

To show the interest of this study, we give some applications of the vector hull in linear algebra. Some results about barycentric calculus, the linear representation of an affine group and the projective completion of an affine space can be easily proved and understood using vector hulls.

In addition to affine spaces, we will also study affine bundles. In the same way as functors on vector spaces give raise to functors on vector bundles, the vector hull functor can be applied to an affine bundle A, yielding a vector bundle $\widehat{A}$ over the same base space, together with an affine bundle inclusion $A \to \widehat{A}$; of course, this is called the vector hull of A. The properties of the vector hull in the framework of affine bundles follow closely those of affine spaces.

Important examples of affine bundles are provided by jet manifolds. As an example, the jet space $\mathrm{J}^1 M$ of a bundle $M \to \mathbf{R}$ is an affine bundle over M, modelled on the vertical tangent bundle $\mathrm{V}M$. There is a canonical affine immersion of $\mathrm{J}^1 M$ in the tangent bundle $\mathrm{T}M$, and it turns out that this bundle is a model of the vector hull of $\mathrm{J}^1 M$ [MMS 02]. The example of the first jet bundle can be extended to other jet bundles, namely higher-order jet bundles and jet bundles over an arbitrary base.

These results can be applied to the geometric study of differential equations, as we will see in the next chapter.

The chapter is organised as follows. In section 5.1 we set some basic notations and give an account of some useful properties of hyperplanes in a vector space. In section 5.2 we describe the vector hull as a solution to a universal problem for affine maps; we also define the vector prolongation of an affine map. Section 5.3 gives a brief description of some geometric applications of the vector hull that, in our opinion, justify giving it its due importance. In section 5.4 we provide with an explicit construction of the vector hull of A in terms of certain vector fields on it. In section 5.5 we review other constructions of the vector hull that can be found in the bibliography and we discuss the essential uniqueness of the vector hull construction. Section 5.6 considers the vector hull of affine bundles. The cases of jet bundles over $\mathbf{R}$ and over an arbitrary base are considered in sections 5.7 and 5.8. Finally, in section 5.9 we consider a particular interesting case, that of the second-order tangent bundle of a manifold.

5.1 Some facts about affine spaces

We assume that the reader is acquainted with the basics about affine spaces, which can be found in many books on linear algebra and linear geometry —see especially [Fre 73] [Ber 77]. In this section we will fix the notation, and also point out some specific results that will be needed later on. We don't suppose the spaces to be finite-dimensional unless stated otherwise. Throughout the paper K denotes a field, which is the ground field of all vector spaces considered, unless stated otherwise.

Given a vector space E (over K), an *affine space* A modelled on E is defined by an action of E on A, denoted by $(\mathbf{u}, p) \mapsto p + \mathbf{u}$. This action is simply transitive, which means that the equation $q = p + \mathbf{u}$ determines a unique vector $\mathbf{u}$, which is denoted by $\overrightarrow{pq} = q - p$.

For any vector $\mathbf{u} \in E$, the bijections $T_{\mathbf{u}} \colon A \to A$ given by $T_{\mathbf{u}}(p) = p + \mathbf{u}$ are called *translations*, and they form a vector space $\vec{A}$ that may be canonically identified with E. On the other hand, the choice of a point $p \in A$ defines a bijection $\varphi_p \colon \vec{A} \to A$, $\mathbf{u} \mapsto p + \mathbf{u}$. With this bijection, A acquires a vector space structure that depends on the point chosen. In the opposite way, any vector space is trivially an affine space.

A map $f\colon A \to B$ between affine spaces is called an *affine map* if there exists an associated linear map $\vec{f}\colon \vec{A} \to \vec{B}$, called the *linear part* of f, such that $f(p + \mathbf{u}) = f(p) + \vec{f}(\mathbf{u})$, for any $p \in A$ and $\mathbf{u} \in \vec{A}$. Note that f is injective [or surjective] iff $\vec{f}$ also is.

We denote by $\mathscr{A}(A, B)$ the set of affine maps between the affine spaces A and B. We also denote by $\mathscr{L}(E, F)$ the set of linear maps between two vector spaces E and F; in particular, the dual space of E is $E^* = \mathscr{L}(E, K)$.

The assignment of the linear part, $f \mapsto \vec{f}$, defines a surjective map $\mathscr{A}(A, B) \twoheadrightarrow \mathscr{L}(\vec{A}, \vec{B})$.

Let A be an affine space modelled on E. A subset $A' \subset A$ is an *affine subspace* if there exists a vector subspace $E' \subset E$ such that $A' = p + E'$ for a certain point $p \in A'$. Then A' is clearly an affine space modelled on E', and the inclusion $A' \hookrightarrow A$ is an affine map, whose associated linear map is nothing but the inclusion $E' \hookrightarrow E$.

Let S be a set. If F is a vector space, the set $\mathscr{F}(S, F)$ of maps from S to F is a vector space. If B is an affine space, then $\mathscr{F}(S, B)$ is an affine space modelled on the vector space $\mathscr{F}(S, \vec{B})$.

Instead of S, consider an affine space A. Then the subset of affine maps is a vector subspace $\mathscr{A}(A, F) \subset \mathscr{F}(A, F)$.

Now consider the set of affine maps between two affine spaces, $\mathscr{A}(A, B) \subset \mathscr{F}(A, B)$. It is an affine subspace modelled on the vector space $\mathscr{A}(A, \vec{B})$. Indeed, if $g\colon A \to B$ and $V\colon A \to \vec{B}$ are affine, then $g + V$ also is. And if $g\colon A \to B$ is a given affine map and $g'\colon A \to B$ is another one, then $g' - g$ is an affine map $g' - g\colon A \to \vec{B}$.

Hyperplanes

We will be especially interested in *hyperplanes* of a vector space E, that is, affine subspaces of codimension 1. A hyperplane $H \subset E$ will be called *proper* if it does not contain 0, otherwise it will be called a vector hyperplane. As usual, we denote by $\langle S \rangle$ the vector subspace spanned by a subset $S \subset E$.

Our first statement is a trivial but useful decomposition:

Lemma 5.1 *Let E be a vector space, $H \subset E$ a proper hyperplane. If $p \in H$, then*

$$E = \langle p \rangle \oplus \vec{H}.$$

■

Now we note that giving a proper hyperplane in E amounts to giving a non-vanishing linear form on E:

Proposition 5.2 *Let E be a vector space. If $w\colon E \to K$ is a non-vanishing linear form, then $H = w^{-1}(1)$ is a proper hyperplane of E; moreover, $\vec{H}$ is identified with* $\operatorname{Ker} w = w^{-1}(0)$. *Conversely, given a proper hyperplane $H \subset E$, there exists a unique non-vanishing linear form $w\colon E \to K$ such that $H = w^{-1}(1)$.*

Proof. The first statement is trivial. Let us prove the converse. Given $p \in H$, we have the decomposition $E = \langle p \rangle \oplus \vec{H}$; then we can define $w(cp + \mathbf{v}) = c$ for $(cp, \mathbf{v}) \in \langle p \rangle \times \vec{H}$. Any other decomposition $E = \langle p' \rangle \oplus \vec{H}$ yields the same linear form, since $cp + \mathbf{v} = cp' + (c(p - p') + \mathbf{v})$. ∎

This linear form w, associated with H (and E) in this way, will be called the *weight function.* A point in E belongs to H iff it has unit weight.

As a first application, we have another useful decomposition:

Corollary 5.3 *Let E be a vector space, $H \subset E$ a proper hyperplane. If $x \in E$ then either $x \in \vec{H}$ or x can be uniquely written as $x = \lambda p$, where $\lambda \neq 0$ and $p \in H$.*

In other words, one has the disjoint union

$$E = \vec{H} \sqcup K^{\times} H,$$

where $K^{\times} = K - \{0\}$.

Proof. Let $w \colon E \to K$ be the weight function associated with H. If $w(x) = 0$ then $x \in \vec{H}$, otherwise $x = w(x) \frac{x}{w(x)}$. ∎

Proposition 5.4 *Consider a linear map $T \colon E \to F$ between vector spaces, and an affine subspace $H \subset E$. The map T is determined by its restriction $T|_H$ iff $H = E$ or H is a proper hyperplane.*

Proof. Clearly T is determined by $T|_H$ iff $\langle H \rangle = E$ (otherwise one could extend the linear map in different ways). ∎

Proposition 5.5 *Let $H \subset E$ be a proper hyperplane in a vector space, $w \colon E \to K$ its weight function, and $h \colon H \to F$ an affine map with values in another vector space. There exists a unique linear map $\bar{h} \colon E \to F$ prolonging h, and is defined by*

$$\bar{h}(x) = \begin{cases} \vec{h}(x) & \text{if } w(x) = 0 \\ w(x) h\left(\frac{x}{w(x)}\right) & \text{if } w(x) \neq 0 \end{cases}$$

Proof. Uniqueness of the prolongation is a consequence of the preceding proposition. By lemma 5.1, every $x \in E$ has a unique decomposition as $x = \lambda p + \mathbf{u}$. Taking into account that $w(\lambda p + \mathbf{u}) = \lambda$, one computes

$$\bar{h}(\lambda p + \mathbf{u}) = \lambda h(p) + \vec{h}(\mathbf{u}).$$

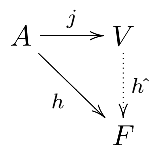
Using this formula it is clear that $\bar{h}$ is a prolongation of h and that it is a linear map. ∎

5.2 A universal problem for affine maps

The vector hull

Let A be an affine space. Consider a vector space V, and an affine map $j\colon A \to V$ satisfying the following **universal property** for affine functions:

> for every vector space F and affine function $h\colon A \to F$, there exists a unique linear function $\hat{h}\colon V \to F$ such that $h = \hat{h} \circ j$.



Proposition 5.6 *A couple (V, j) satisfies the universal property iff j is injective and $j(A) \subset V$ is a proper hyperplane.*

Proof. Affine functions $h\colon A \to F$ separate points, therefore the map j is necessarily injective. The subset $j(A) \subset V$ is an affine subspace. According to proposition 5.4, $\hat{h}$ is determined iff $j(A)$ is either a proper hyperplane or the whole space V; the later possibility cannot occur, since a constant function $h \neq 0$ is affine but can never be linear. The existence (and explicit construction) of $\hat{h}$ is a consequence of proposition 5.5. ∎

Definition 5.7 *The couple (V, j) is called a vector hull of A.*

As usual in universal problems, the solution is unique, up to isomorphism, but let us write the proof anyway. Consider a second vector hull, $j'\colon A \to V'$, of A. The universal property for the first one and the linear map j' yields a linear map $f\colon V \to V'$ such that $j' = f \circ j$. The only thing to prove is that f is an isomorphism. But reversing the role of both vector hulls, the same reasoning leads to $g\colon V' \to V$ such that $j = g \circ j'$. Therefore $j = g \circ f \circ j$, which leads to $g \circ f = \mathrm{Id}_V$, and similarly in the reverse order, which shows that g is the inverse of f.

On the existence of the vector hull, we shall see later on that it can be given several explicit constructions. However, in practice we will find the vector hull directly from an analysis of a certain vector space, as expressed in the following corollary:

Corollary 5.8 *Let V be a vector space, $w\colon V \to K$ a nonzero linear form. If $A_1 = w^{-1}(1)$ and $A_0 = w^{-1}(0)$, then A_1 is an affine space modelled on the vector space A_0, and the inclusion $j\colon A_1 \hookrightarrow V$ is a vector hull of A_1.* ∎

From now on we shall denote by $j\colon A \to \widehat{A}$ (any construction of) the vector hull of A. We will often identify A with its image $j(A) \subset \widehat{A}$ in the vector hull; with this identification, the function $\hat{h}$ is a prolongation of h; we call it the *homogenisation* of h.

We know (proposition 5.2) that there is a unique linear form $w\colon \widehat{A} \to K$ (the weight function) such that $j(A) = w^{-1}(1)$. So, in the same way as A, its translation space $\vec{A}$ can also be identified with a subspace of $\widehat{A}$, namely $w^{-1}(0)$.

With these identifications, the right-hand side of the equality $q = p + \mathbf{u}$ can be interpreted as an addition in the vector hull; in the same way, $\overrightarrow{pq} = q - p$ is a subtraction in the vector hull.

In addition, the weight function w is nothing but the homogenisation of the constant function equaling 1 on A: $w = (1_A)\hat{}$.

Let us summarise all this in a diagram:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \vec{A} & \xrightarrow{i_A} & \widehat{A} & \xrightarrow{w_A} & K & \longrightarrow & 0 \\ & & & & {\scriptstyle j_A}\big\uparrow & & & & \\ & & & & A & & & & \end{array} \qquad \begin{array}{ccc} \widehat{A} & \xrightarrow{h\hat{}} & F \\ {\scriptstyle j_A}\big\uparrow & \nearrow_{h} & \\ A & & \end{array}$$

Here we have put a subscript A to indicate, if needed, which is the initial affine space.

As a comment on notation, we are denoting by $h\hat{}$ a certain prolongation of an affine function; later on we will denote by $\widehat{h}$ a certain prolongation of an affine map; both will be closely related but different objects. (This has some reminiscence of the distinction between the differential $\mathrm{d}f$ of a function on a manifold and the tangent map $\mathrm{T}f$.)

Proposition 5.9 *Let A be an affine space. For any vector space F, the homogenisation map*

$$\mathscr{A}(A, F) \to \mathscr{L}(\widehat{A}, F), \quad h \mapsto h\hat{},$$

is a linear isomorphism.

In particular, $\mathscr{A}(A, K) \cong \widehat{A}^*$, and $\mathscr{A}(A, K)^* \cong \widehat{A}^{**}$.

Proof. The linearity of the map $h \mapsto h\hat{}$ is a consequence of the uniqueness of the homogenisations: $(h + k)\hat{}$ and $h\hat{} + k\hat{}$ coincide on $j(A)$, therefore coincide everywhere; a similar argument holds for λh.

The map $h \mapsto h\hat{}$ is injective since different functions h have necessarily different prolongations. And it is surjective since every linear map $\widehat{A} \to F$ restricts to an affine map $A \to F$ of which it is a prolongation. ■

Note that in finite dimension we obtain that $\widehat{A}$ is isomorphic to $\mathscr{A}(A, K)^*$.

We finish the study of the homogenisation with the following result, whose proof is straightforward:

Proposition 5.10 *Let $h\colon A \to F$ an affine function, $h\hat{}\colon \widehat{A} \to F$ its homogenisation.*

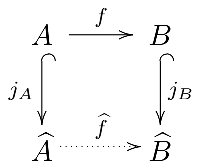
- *$h\hat{}$ is surjective iff the affine subspace $h(A) \subset F$ spans F.*
- *$h\hat{}$ is injective iff h is injective and nowhere vanishing.*
- *$h\hat{}$ is bijective iff h is injective and $h(A) \subset F$ is a proper hyperplane.* ■

The vector prolongation

Up to now, we have immersed every affine space in its vector hull. Now we examine whether we can assign a linear map to a morphism of affine spaces.

Proposition 5.11 *Given an affine map $f\colon A \to B$, there is a unique linear map $\widehat{f}\colon \widehat{A} \to \widehat{B}$ such that $\widehat{f} \circ j_A = j_B \circ f$.*

Proof. Application of the universal property of the vector hull $(\widehat{A}, j_A)$ of A to the affine function $j_B \circ f\colon A \to \widehat{B}$ yields the desired map $\widehat{f} = (j_B \circ f)\hat{}$. ∎

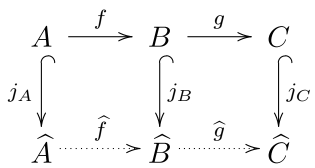


We call $\widehat{f}$ the *vector prolongation* of f since, considering the affine spaces as subsets of their vector hulls, $\widehat{f}$ is actually a prolongation of f. The fact that $\widehat{f}(A) \subset B$ implies, more generally, that $\widehat{f}$ preserves the weight:

$$w_B \circ \widehat{f} = w_A.$$

Proposition 5.12 *The vector prolongation satisfies $\widehat{\mathrm{Id}_A} = \mathrm{Id}_{\widehat{A}}$ and $\widehat{g \circ f} = \widehat{g} \circ \widehat{f}$.*

Proof. The uniqueness of the vector prolongation is used to prove both statements. The first one is obvious. For the second one, consider the diagram



∎

Proposition 5.13 *Let $f\colon A \to B$ be an affine map. The map f is injective [or surjective] iff $\widehat{f}$ also is. In particular, f is an affine isomorphism iff $\widehat{f}$ is a linear isomorphism.*

Proof. We first prove the direct implications.

As in the category of vector spaces and linear maps, if an affine map f is injective then it has a left-inverse, that is to say, an affine map $g\colon B \to A$ such hat $g \circ f = \mathrm{Id}_A$. Then by the preceding proposition $\widehat{g} \circ \widehat{f} = \mathrm{Id}_{\widehat{A}}$, which shows that $\widehat{f}$ also has a left-inverse, and thus it is injective.

In a similar way, if f is surjective then it has a right-inverse $g\colon B \to A$, which satisfies $f \circ g = \mathrm{Id}_B$. The proof goes on in the same way.

The converse statements are almost immediate: if $\widehat{f}$ is injective, then its restriction f also is; and if $\widehat{f}$ is surjective, taking into account that it preserves weights, we conclude that its restriction f also is. ■

Remark 5.14 This result may be useful to identify vector hulls of some affine spaces. If A is an affine subspace of B, then proposition 5.13 shows that the vector hull $\widehat{A}$ can be identified as a vector subspace of $\widehat{B}$. Moreover, if A is an affine subspace of a vector space F, and A does not contain the zero vector, then proposition 5.10 shows that the vector hull $\widehat{A}$ can be identified as a vector subspace of F.

This is illustrated by the following proposition, where we identify the vector hull of the set of affine maps from A to B.

Proposition 5.15 *Let A, B be affine spaces. Consider the vector prolongation map*

$$\mathrm{Vh}\colon \mathscr{A}(A,B) \to \mathscr{L}(\widehat{A},\widehat{B}), \quad f \mapsto \widehat{f}.$$

1. Vh *is an affine injection.*
2. *The image of* Vh *is the set* $\{T \in \mathscr{L}(\widehat{A},\widehat{B}) \mid T(A) \subset B\}$.
3. *The vector hull of* $\mathscr{A}(A,B)$ *is identified with* $\{T \in \mathscr{L}(\widehat{A},\widehat{B}) \mid T(\vec{A}) \subset \vec{B}\}$.

Proof. The map Vh is injective since obviously two different maps f, f' cannot have the same vector prolongation.

Remember that $\mathscr{A}(A,B)$ has an affine space structure, modelled on the vector space $\mathscr{A}(A,\vec{B})$. The affinity of the vector prolongation is a consequence of

$$\widehat{f+h} = \widehat{f} + i_B \circ \hat{h},$$

for affine maps $f\colon A \to B$ and $h\colon A \to \vec{B}$. By proposition 5.4, to prove this equality it is enough to show that these linear maps agree on the hyperplane A:

$$(\widehat{f} + i_B \circ \hat{h}) \circ j_A = \widehat{f} \circ j_A + i_B \circ \hat{h} \circ j_A = j_B \circ f + i_B \circ h = j_B \circ (f+h) = (\widehat{f+h}) \circ j_A.$$

The second statement is obvious: a linear map $T\colon \widehat{A} \to \widehat{B}$ restricts to an affine map $A \to B$ iff the inclusion $T(A) \subset B$ holds.

For the third statement, consider the linear form $w\colon \{T \in \mathscr{L}(\widehat{A},\widehat{B}) \mid T(\vec{A}) \subset \vec{B}\} \to K$ defined by $w(T) = w_B(T(a))$, that is, the weight of T is the weight of any of its images. This map is well defined, is linear, and its value is 1 iff $T(A) \subset B$. We conclude by applying corollary 5.8. ■

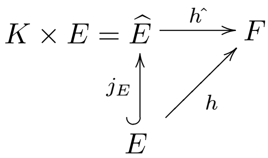
The vector hull of a vector space

Here we describe the particular, but important, case where where the affine space is a vector space E. Due to the characterisation (proposition 5.6) of the vector hull, it can be canonically identified with:

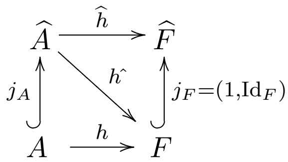
- the vector space $\widehat{E} = K \times E$, with
- the affine inclusion $j_E(u) = (1, u)$;

then the weight function is $w_E(\lambda, u) = \lambda$.

Consider an affine map $h\colon E \to F$ between vector spaces: $h(u) = h_0 + h_1 \cdot u$, where $h_0 \in F$ and $h_1\colon E \to F$ is linear. Then we have that $h\hat{}\colon K \times E \to F$ is the linear map $h\hat{}(\lambda, u) = h_0\lambda + h_1 \cdot u$, which justifies calling it *homogenisation.*



Now let us see how is the vector prolongation. Given an affine map $h\colon A \to F$, its vector prolongation $\widehat{h}\colon \widehat{A} \to \widehat{F} = K \times F$ is given by $\widehat{h} = (w_A, h\hat{})$. So we must carefully distinguish between $h\hat{}$ and $\widehat{h}$. Let us show this in a (noncommutative) diagram:



This can be applied in particular to an affine map $h\colon E \to F$ between vector spaces, $h(u) = h_0 + h_1 \cdot u$. Its vector prolongation is the map $\widehat{h}\colon K \times E \to K \times F$ given by $\widehat{h}(\lambda, u) = (\lambda, h_0\lambda + h_1 \cdot u)$.

Remark 5.16 Consider a vector space F and an affine subspace $A \subset F$, with inclusion $h\colon A \hookrightarrow F$. If $0 \notin A$ then, by proposition 5.10, one can consider $\widehat{A}$ as a subspace of F via $h\hat{}\colon \widehat{A} \hookrightarrow F$. On the other hand, if $0 \in A$ then $\widehat{A}$ can not be identified with subspace of F, but applying proposition 5.13, which gives $\widehat{h}\colon \widehat{A} \hookrightarrow \widehat{F} = K \times F$, it can always be considered as a subspace of $\widehat{F}$.

Coordinate description

Consider a point $a_0 \in A$. Identifying A and $\vec{A}$ as subsets of $\widehat{A}$, $\vec{A} \subset \widehat{A}$ is a hyperplane and a_0 is not in $\vec{A}$, and we have the decomposition $\widehat{A} = \langle a_0 \rangle \oplus \vec{A}$.

Cartesian coordinates

Write $e_0 = a_0$, and consider a basis $(e_i)_{i\in I}$ of $\vec{A}$. Then, putting $\hat{I} = \{0\} \cup I$, we see that $(e_\mu)_{\mu\in\hat{I}}$ is a basis for $\widehat{A}$; thus every point in $\widehat{A}$ can be uniquely written as $x = x^0 e_0 + x^i e_i$. It is clear that $w(x) = x^0$; therefore a point $x \in \widehat{A}$ belongs to A iff $x^0 = 1$, and belongs to $\vec{A}$ iff $x^0 = 0$.

Now we shall use these coordinates to express the homogenisation and the vector prolongation.

Besides the coordinates on A, consider a vector space F, with a basis, and an affine map $h\colon A \to F$ given in the corresponding coordinates of A and F by $y^j = b^j + T^j_i x^i$. Then its homogenisation is the linear map $\hat{h}\colon \widehat{A} \to F$ given by $(x^0; x^i) \mapsto (b^j x^0 + T^j_i x^i)$.

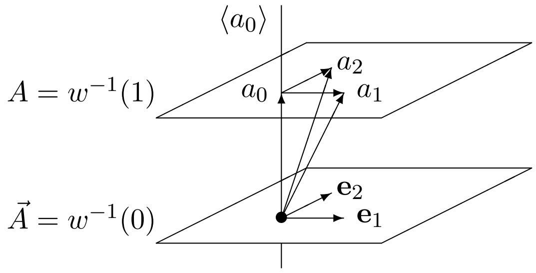
Now consider another affine space B, and an affine map $f\colon A \to B$ given in the same type of coordinates as $y^j = b^j + T^j_i x^i$. Then its vector prolongation is the linear map $\widehat{f}\colon \widehat{A} \to \widehat{B}$ given by $y^\nu = T^\nu_\mu x^\mu$, with $T^j_0 = b^j$, $T^0_0 = 1$, $T^0_i = 0$. So the matrix of $\widehat{f}$ in this basis is

$$\left(\begin{array}{c|c} 1 & 0 \\ \hline b^j & T^j_i \end{array}\right).$$

Barycentric coordinates

In addition to the preceding basis $(e_0; e_i)$ for $\widehat{A}$, one can also work with another one that only consists of points of A: if $a_i = a_0 + e_i$, then the family $(a_\mu)_{\mu\in\hat{I}}$ is another basis of $\widehat{A}$. If $x = \lambda^\mu a_\mu$, now $w(x) = \sum \lambda^\mu$; thus $x \in A$ iff $\sum \lambda^\mu = 1$, and $x \in \vec{A}$ iff $\sum \lambda^\mu = 0$.

In geometric terms, one can say that the first basis is a cartesian frame for A, the second basis is an affine frame for A.



Two basis for $\widehat{A}$

5.3 Some geometric applications of the vector hull

Barycentric calculus

In affine geometry one defines the notion of barycentre of some points $p_1, \ldots, p_N \in A$ with masses (or weights) $\lambda_1, \ldots, \lambda_N \in K$, with $\sum \lambda_k = 1$. This is the point p such that $\overrightarrow{op} = \lambda_1 \overrightarrow{op_1} + \ldots \lambda_N \overrightarrow{op_N}$, where o is an arbitrary fixed origin. Of course, one needs to show that this notion does not depend on the origin choosed.

Within the vector hull $\widehat{A}$ the *barycentre* is simply $p = \lambda_1 p_1 + \ldots \lambda_N p_N$, which belongs to A because its weight is $w(p) = \sum \lambda_k = 1$. For instance, if $K = \mathbf{R}$ the middle point of the segment $[p,q]$ is $(p+q)/2$.

Therefore, working in the vector hull of A provides with a neat presentation of barycentric calculus. As an example, we can easily prove an interesting characterisation of affine functions:

Proposition 5.17 *A map $f\colon A \to B$ between affine spaces is affine iff it preserves barycentres, in the sense that $f(\sum \lambda_k p_k) = \sum \lambda_k f(p_k)$ for any finite family of points $p_k \in A$ and scalars $\lambda_k \in K$ adding up to 1.*

Proof. The direct implication is immediate: if f is affine, then its vector prolongation $\widehat{f}$ is linear, therefore working in $\widehat{A}$ and $\widehat{B}$ we have $\widehat{f}(\sum \lambda_k p_k) = \sum \lambda_k \widehat{f}(p_k)$; all the $\widehat{f}$ can be replaced by f since all their arguments are in A.

To prove the converse, consider an affine frame (a_μ) of A. This is also a basis of $\widehat{A}$, therefore there exists a unique linear map $F\colon \widehat{A} \to \widehat{B}$ that coincides with f on the set $\{a_\mu\}$: $F(a_\mu) = f(a_\mu)$. Any point in A is $\lambda^\mu a_\mu$ with $\sum \lambda^\mu = 1$. Then

$$F(\lambda^\mu a_\mu) = \lambda^\mu F(a_\mu) = \lambda^\mu f(a_\mu) = f(\lambda^\mu a_\mu),$$

where we have used that f preserves barycentres. This shows that F coincides with f everywhere on A. Since F is linear, f is affine. ∎

The converse statement can be given an alternative proof that does not need the usage of an affine frame. To prove it, we must show that, for every p, the map $f(p+\mathbf{u}) - f(p)$ is a linear function of $\mathbf{u}$ —this will be $\vec{f}(\mathbf{u})$. Consider the barycentre of three points p, q, r with masses $1-\lambda-\mu$, λ, μ. By hypothesis $f((1-\lambda-\mu)p+\lambda q+\mu r) = (1-\lambda-\mu)f(p) + \lambda f(q) + \mu f(r)$. Thinking inside the vector hulls this can be rewritten as $f(p+\lambda(q-p)+\mu(r-p)) = f(p) + \lambda(f(q)-f(p)) + \mu(f(r)-f(p))$, and also

$$f(p+\lambda\mathbf{u}+\mu\mathbf{v}) = f(p) + \lambda(f(p+\mathbf{u})-f(p)) + \mu(f(p+\mathbf{v})-f(p)).$$

We apply this equation in two particular instances. Taking $\mu = 0$ we have $f(p+\lambda\mathbf{u}) - f(p) = \lambda(f(p+\mathbf{u})-f(p))$, that is, homogeneity. And taking $\lambda = \mu = 1$ we have $f(p+\mathbf{u}+\mathbf{v}) - f(p) = (f(p+\mathbf{u})-f(p)) + (f(p+\mathbf{v})-f(p))$, that is, additivity. Both properties show the linearity.

The linear representation of the affine group

Consider the *affine group* $\mathbf{GA}(A)$, whose elements are the affine automorphisms f of A. By proposition 5.13, their vector prolongations $\widehat{f}$ are linear automorphisms of $\widehat{A}$, and from proposition 5.12 we conclude:

Proposition 5.18 *Let A be an affine space. The map $f \mapsto \widehat{f}$ defines an injective group homomorphism $\mathbf{GA}(A) \hookrightarrow \mathbf{GL}(\widehat{A})$.* ∎

For a vector space E we have $\mathbf{GA}(E) \hookrightarrow \mathbf{GL}(K \times E)$, and more particularly, taking $E = K^n$, we obtain the well-known linear representation of the affine group in one more dimension, $\mathbf{GA}_n(K) \hookrightarrow \mathbf{GL}_{n+1}(K)$.

The projective completion of an affine space

In courses on projective geometry it is explained that an affine space admits a projective completion, that is to say, it can be attached a projective hyperplane at infinity such that the whole stuff is a projective space. This can be done by adding an additional coordinate, but many authors seem to be unaware that there is a canonical way to perform this operation.

First recall that the projective space $\mathbf{P}(E)$ of a vector space E can be defined as the quotient $E–\{0\}/\sim$, where $x \sim y$ iff they are linearly dependent; $\mathbf{P}(E)$ is also the set of 1-dimensional subspaces of E, so that its points are the lines $\langle x \rangle \subset E$, where $x \in E - \{0\}$. If $f\colon E \to F$ is an injective linear map then for $x \neq 0$, $f(\langle x \rangle) = \langle f(x) \rangle \neq \{0\}$, therefore one obtains an injective map between the projective spaces, $\mathbf{P}(f)\colon \mathbf{P}(E) \to \mathbf{P}(F)$, called a projective linear map. (If f is not injective, the domain of $\mathbf{P}(f)$ has to be changed to $\mathbf{P}(E) - \mathbf{P}(\operatorname{Ker} f)$ instead of all the space.)

Let A be an affine space, with vector hull $j\colon A \hookrightarrow \widehat{A}$. Let $a \in A$; $j(a)$ is not zero, the subspace $\langle j(a) \rangle \subset \widehat{A}$ is a point in the projective space, $\langle j(a) \rangle \in \mathbf{P}(\widehat{A})$. Moreover, given two different points $a, a' \in A$, we have $\langle j(a) \rangle \neq \langle j(a') \rangle$. Therefore we have defined an inclusion

$$A \overset{k}{\hookrightarrow} \mathbf{P}(\widehat{A}), \qquad a \mapsto \langle j(a) \rangle.$$

Consider also the inclusion $i\colon \vec{A} \hookrightarrow \widehat{A}$. Since it is an injective linear map, it passes to the projective spaces and yields another inclusion

$$\mathbf{P}(\vec{A}) \overset{\mathbf{P}(i)}{\hookrightarrow} \mathbf{P}(\widehat{A}), \qquad \langle \mathbf{u} \rangle \mapsto \langle i(\mathbf{u}) \rangle.$$

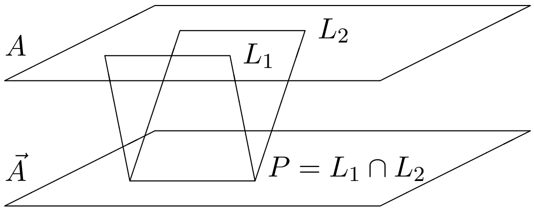
We will identify A and $\mathbf{P}(\vec{A})$ with their images in $\mathbf{P}(\widehat{A})$.

Proposition 5.19 $\mathbf{P}(\widehat{A}) = A \sqcup \mathbf{P}(\vec{A})$ *(disjoint union).*

Proof. Let w be the weight function associated with the proper hyperplane $A \subset \widehat{A}$. A point in the projective space is the linear span of a certain $x \in \widehat{A} - \{0\}$. If $w(x) = 0$ then $x \in \vec{A}$ and $\langle x \rangle \in \mathbf{P}(\vec{A})$. If $w(x) \neq 0$ then $x/w(x) \in A$ and $\langle x \rangle \in k(A) \cong A$. ■

Remark 5.20 The identification of points in A with one-dimensional subspaces of $\widehat{A}$ not contained in $\vec{A}$ is indeed more general: non-empty affine subspaces of A correspond to vector subspaces of $\widehat{A}$ not contained in $\vec{A}$.

The space $A^{\mathbf{P}} = \mathbf{P}(\widehat{A})$ is called the *projective completion* (or projective closure) of A. It has the same dimension as A, and can be understood as the result of attaching to A a projective hyperplane $A_\infty = \mathbf{P}(\vec{A})$, called the *hyperplane at infinity.*



Two parallel lines in A meet at a point in $A_\infty = \mathbf{P}(\vec{A})$

Proposition 5.21 *Any injective affine map $f\colon A \to B$ between affine spaces has a prolongation $f^{\mathbf{P}}\colon A^{\mathbf{P}} \to B^{\mathbf{P}}$ to their projective completions, which is a projective linear map.*

Proof. First we consider the vector prolongation, which is an injective linear map $\widehat{f}\colon \widehat{A} \to \widehat{B}$. Then we take as $f^{\mathbf{P}}$ the corresponding projective linear map $\mathbf{P}(\widehat{f})\colon \mathbf{P}(\widehat{A}) \to \mathbf{P}(\widehat{B})$. ■

5.4 A construction of the vector hull

Homothetic vector fields

Let A be an affine space of dimension > 0, that is, not reduced to a point.

The set of maps $X\colon A \to \vec{A}$ is a vector space $\mathscr{F}(A, \vec{A})$. We call any of these maps a *vector field*. (Of course, the terminology is justified since, when A is real and finite-dimensional, the tangent bundle of A (as a smooth manifold) is trivial, $\mathrm{T}A = A \times \vec{A}$, and its sections are maps $p \mapsto (p, X(p))$.)

In this space, the subset of affine maps $X\colon A \to \vec{A}$ is a vector subspace $\mathscr{A}(A, \vec{A})$. We call any of these maps an *affine vector field*.

As an affine map, X has an associated linear map $\vec{X}\colon \vec{A} \to \vec{A}$. If this endomorphism is a homothety (*i.e.*, a multiple of the identity), let us call X a *homothetic vector field*. In this section, let us denote by $\widehat{A}$ the set of homothetic vector fields on A.

Proposition 5.22 *The set $\widehat{A}$ of homothetic vector fields of A is a vector subspace of $\mathscr{A}(A, \vec{A})$, and the map $w\colon \widehat{A} \to K$ defined by*

$$\vec{X} = -w(X)\mathrm{Id}$$

is a nonzero linear form.

Proof. The proof is immediate, as illustrated by the following diagram. ■

$$\begin{array}{ccc}
\widehat{A} & \xrightarrow{-w} & K \\
\big\downarrow & & \big\downarrow \\
\mathscr{A}(A, \vec{A}) & \longrightarrow & \mathscr{L}(\vec{A}, \vec{A}) \\
X & \longmapsto & \vec{X}
\end{array}$$

(The minus sign in the definition of w will be explained later.)

Now consider the following special vector fields on A:

- Constant vector field $Y_{\mathbf{u}}(p) = \mathbf{u}$, where $\mathbf{u} \in \vec{A}$.
- Central vector field $Z_a^\lambda(p) = \lambda\,\overrightarrow{pa} = -\lambda\,\overrightarrow{ap}$, where $\lambda \in K^\times = K - \{0\}$ and $a \in A$.

They are homothetic, and indeed they exhaust all the homothetic vector fields on A:

Proposition 5.23 *Let $X \in \widehat{A}$ be a homothetic vector field. If $w(X) = 0$, then X is a constant vector field $Y_{\mathbf{u}}$. If $w(X) = \lambda \neq 0$, then X is a central vector field Z_a^λ.*

Proof. Suppose that $\vec{X} = 0$. Then $X(p+\mathbf{v}) = X(p) + \vec{X}(\mathbf{v}) = X(p)$, so that X takes a constant value $\mathbf{u}$.

Now suppose that $\vec{X} = -\lambda\,\mathrm{Id}_{\vec{A}}$. First, note that X necessarily vanishes somewhere: its value in a certain point, $X(p+\mathbf{v}) = X(p) - \lambda\mathbf{v}$, may be set to zero by taking $\mathbf{v} = \lambda^{-1}X(p)$. Let $a \in A$ such that $X(a) = 0$; then $X(a+\mathbf{v}) = X(a) - \lambda\mathbf{v} = -\lambda\mathbf{v}$, so that $X(p) = -\lambda\overrightarrow{ap}$, that is, $X = Z_a^\lambda$. ■

Therefore we have shown that the vector space $\widehat{A}$ of homothetic vector fields is the disjoint union of constant vector fields (which can be identified with $\vec{A}$) and central vector fields (which can be identified with $K^\times \times A$). This can be written as

$$\widehat{A} = \vec{A} \sqcup (K^\times \times A).$$

The following rules of calculus can be easily checked:

Proposition 5.24 *The homothetic vector fields satisfy the following relations:*

1. $\lambda Y_{\mathbf{u}} = Y_{\lambda\mathbf{u}}$
2. $Y_{\mathbf{u}} + Y_{\mathbf{v}} = Y_{\mathbf{u}+\mathbf{v}}$
3. $\lambda Z_a^\mu = Z_a^{\lambda\mu}$
4. $Z_a^\lambda + Z_b^\mu = Z_c^{\lambda+\mu}$ *if* $\lambda+\mu \neq 0$, *where* $c = (\lambda a + \mu b)/(\lambda+\mu) = a + \frac{\mu}{\lambda+\mu}\overrightarrow{ab}$
5. $Z_a^\lambda + Z_b^{-\lambda} = Y_{\lambda\overrightarrow{ba}}$
6. $Z_a^\lambda + Y_{\mathbf{u}} = Z_{a+\lambda^{-1}\mathbf{u}}^\lambda$ ■

Note in particular that $Y_{\mathbf{u}}$ is a linear function of $\mathbf{u}$, and also the special cases

3'. $\lambda Z_a^1 = Z_a^\lambda$

5'. $Z_a^1 + Z_b^{-1} = Y_{\overrightarrow{ba}}$

6'. $Z_a^1 + Y_{\mathbf{u}} = Z_{a+\mathbf{u}}^1$

The last one shows that Z_a^1 is an affine function of a, with associated linear function Y. So, we have also proved the following:

Proposition 5.25 *The map given by*

$$i\colon \vec{A} \hookrightarrow \widehat{A}, \quad i(\mathbf{u}) = Y_{\mathbf{u}}$$

is linear, and its image coincides with $w^{-1}(0)$.
The map given by

$$j\colon A \hookrightarrow \widehat{A}, \quad j(a) = Z_a^1$$

is affine, and its image coincides with $w^{-1}(1)$. ■

Let us remark that the choice of a sign in w as well as in the definition of the central vector fields is necessary to describe A as $w^{-1}(1)$ and to have $\vec{j} = i$.

Gathering the preceding results, and according to corollary 5.8, we have shown that:

Corollary 5.26 *The couple* $(\widehat{A}, j)$ *is a vector hull of* A. ■

Associated with concrete constructions $\widehat{A}$ of the vector hull of affine spaces, we have defined the vector prolongations $\widehat{f}$ of affine maps, which can be essentially described from propositions 5.5 and 5.11. Now that we identify the vector hull with the space of homothetic vector fields, and each point $a \in A$ is mapped to the central vector field $j(a) = Z_a^1$, we want to see how is the vector prolongation of an affine map in this construction.

Indeed, it is straightforward. Let $f\colon A \to B$ be an affine map. Since f maps a to $f(a)$, its vector prolongation $\widehat{f}$ maps Z_a^1 to $Z_{f(a)}^1$; therefore, using $\lambda Z_a^1 = Z_a^\lambda$ and the linearity of $\widehat{f}$, we have $\widehat{f}(Z_a^\lambda) = Z_{f(a)}^\lambda$. In the same way, $\widehat{f}(Y_{\mathbf{u}}) = Y_{\vec{f}(\mathbf{u})}$. This completes the description of $\widehat{f}$.

The vector hull functor

Proposition 5.27 *The assignment* $A \rightsquigarrow \widehat{A}$, $f \rightsquigarrow \widehat{f}$, *is a covariant functor from the category of K-affine spaces* **Aff** *to the category of K-vector spaces* **Vect**.

Proof. It is a consequence of proposition 5.12. ■

The case of a vector space

The special case of a vector space E can be studied. Vector fields on E are maps $X\colon E \to E$. The affine ones have the form $X(u) = a + T(u)$ where T is linear. And the homothetic ones have the form $X(u) = a - \lambda u$. So, denoting this one by $X_{\lambda,a}$, we obtain a bijection

$$K \times E \to \widehat{E}, \quad (\lambda, a) \mapsto X_{\lambda,a},$$

which reproduces the results from section 5.2. For $\lambda = 0$ we obtain the constant vector fields. For $\lambda \neq 0$ we obtain the central vector field $Z_{\lambda^{-1}a}^\lambda$.

The case of a point

In this section we have implicitly used that $\dim A > 0$. If the affine space is reduced to a point, $A = \{a\}$, then the only vector field on A is the zero constant, and the weight function of proposition 5.22 is not well-defined. In this case one takes as the vector hull $\widehat{A} = Ka$.

5.5 Other constructions of the vector hull

Constructions of the vector hull that are different from that given in the preceding section are seen in the literature. We will give a short review of them and show that all the constructions are equivalent.

Bibliographic review

Here we have a look at some definitions of the vector hull that can be found in the bibliography. We use our own notations.

A construction of the vector hull of an affine space A appears as an exercise of Bourbaki's *Algèbre* [Bou 70]. Indeed, this is already found in the third edition of its second chapter [Bou 62], and to our knowledge this is the first appearance of the vector hull. There $\widehat{A}$ is constructed as the quotient $K^{(A)}/N$, where $K^{(A)}$ is the vector space having all the points of A as a basis, and N is the subspace generated by all the relations

$$[p + \mathbf{u}] - [p] - [q + \mathbf{u}] + [q], \quad [p + \lambda\mathbf{u}] - [p] - \lambda[q + \mathbf{u}] + \lambda[q],$$

with $p, q \in A$, $\mathbf{u} \in \vec{A}$, $\lambda \in K$. Then the map $j \colon A \to \widehat{A}$, $p \mapsto [p]$, is a solution to the universal problem for affine functions.

In Frenkel's book on geometry [Fre 73] the vector hull of A is constructed directly as the union $\mathscr{D}(A) = \mathscr{C}(A) \cup \mathscr{C}'(A)$ of constant vector fields and central vector fields —the homothetic vector fields in proposition 5.23— and is identified with the set $\widehat{A} = \vec{A} \cup (K^\times \times A)$. Of course, one needs to show that this is a vector space —this is the contents of our proposition 5.24. The vector hull, which is not given a name, is widely used to study barycentres, the vector prolongation, the projective completion... This exposition is acknowledged to Glaeser.

In the article [BB 75] (included in the book [Sch 75]) Bamberger and Bourguignon define "equiprojective k-vector fields" on an affine space A: they are maps $X \colon A \to \bigwedge^k \vec{A}$ such that, for each $p, q \in A$, $X(p) \wedge \vec{pq} = X(q) \wedge \vec{pq}$. After a short study of some general properties, they consider the particular case of equiprojective vector fields, $\mathrm{Eq}(A)$, which again coincide with homothetic vector fields, thus giving another definition of $\widehat{A}$. This space there is called the *vectorialised* of A.

Published in the same year by the IREM of Strasbourg there is a book devoted to barycentric calculus and its applications [IREM 75], where the vector hull is defined as

$\widehat{A} = \vec{A} \cup (K^\times \times A)$, and identified with the homothetic vector fields as in Frenkel's book. Shortly after, the well-known book on geometry by Berger [Ber 77] appeared, following the same presentation; in this book the vector hull, called *universal space*, is widely used in affine and projective geometries.

The definition $\widehat{A} = \vec{A} \cup (K^\times \times A)$ can be found subsequently in a few textbooks, as for instance Tisseron's [Tis 83], where one can also find $\widehat{A} = \mathscr{A}(A, K)^*$, the dual space of affine functions. As we have shown, this construction is equivalent to the other ones in finite dimension, otherwise the space obtained in this way is "too big".

The vector hull also appears in an article with applied scope by Ramshaw [Ram 89]: if A is an affine space, a vector space $\hat{A}$ with a linear form $w \colon \hat{A} \to K$ (the weight) such that $A = w^{-1}(1)$ is called the *homogenisation* of A. However, no explicit construction of $\hat{A}$ is given. The vector hull and the vector prolongation are essentially used to homogenise affine maps and polynomial maps in view of applications to computer-aided geometric design.

A vector field $X \colon A \to \vec{A}$ defines a map $f_X \colon A \to A$, given by $f_X(p) = p + X(p)$, and conversely. This induces a bijection between homothetic vector fields and their associated maps, and these maps constitute the definition of the vector hull in Gallier's book [Gal 01] on geometry and applications: the *homogenisation* $\widehat{A}$ is defined as the (disjoint) union of the translations $T_{\mathbf{u}}$ together with the dilatations $H_{a,\lambda}$, where $H_{a,\lambda}(x) = a + \lambda\overrightarrow{ax}$. However, the vector space structure is not quite clear without the consideration of the vector fields themselves. But indeed the affine space $\mathscr{A}(A, A)$ has a privileged point, the identity map Id_A, therefore this space also has a canonical structure of vector space, which can be used to define the vector hull as pointed out recently by Bertram [Ber 04].

Vector hulls are considered in an article by Martínez, Mestdag and Sarlet [MMS 02], devoted to algebroid constructions in the affine framework. In these reference the vector hull is defined as the dual space of affine functions, $\mathscr{A}(A, K)^*$, and is called the *bidual* of A. This also appears in Mestdag's thesis [Mes 03], which develops explicitly the case of an affine bundle, and also uses the term *vector hull*.

Finally, another construction very similar to Bourbaki's has appeared in a recent paper by Grabowska, Grabowski and Urbański [GGU 03] devoted to affine-valued differential geometry. In this paper the vector hull is defined as the quotient $K^{(A)}/N$, where N is the subspace generated by the relations

$$[p + \lambda(p' - p'')] - [p] - \lambda[p'] + \lambda[p''].$$

In this paper affine bundles are also considered.

Besides these constructions, from a more elementary viewpoint one can take $\widehat{A} = K \oplus \vec{A}$. This can be found in several textbooks on geometry aiming to construct the projective completion of an affine space. The main drawback of this construction is that, opposite to all the preceding ones, the inclusion $j\colon A \hookrightarrow \widehat{A}$ depends on the choice of a privileged point in A.

Equivalence of the vector hull functors

Given several constructions of a "same" object, one may wonder if all them are equivalent. Of course, two vector hulls of A are isomorphic —A is immersed as a proper hyperplane— but the question goes further on. In the preceding paragraphs we have pointed out that there are several vector hull *functors*. Are they equivalent?

In general, if two functors solve a universal problem they are equivalent. More precisely, there exists a *natural equivalence* between them.

For the vector hull this means the following. Suppose we have two vector hull functors $A \rightsquigarrow \widehat{A}$, $f \rightsquigarrow \widehat{f}$ and $A \rightsquigarrow \widetilde{A}$, $f \rightsquigarrow \widetilde{f}$, which give solutions to the universal problem for affine functions. Then they are equivalent, in the following sense: for each affine space A there exists a vector space isomorphism $\varphi_A\colon \widehat{A} \to \widetilde{A}$ such that, for each affine map $f\colon A \to B$, this diagram is commutative:

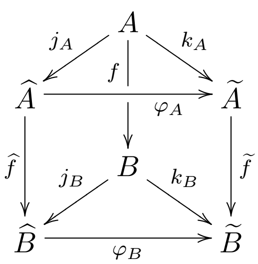
$$\begin{array}{ccc}\widehat{A} & \xrightarrow{\widehat{f}} & \widehat{B}\\ {\scriptstyle\varphi_A}\big\downarrow & & \big\downarrow{\scriptstyle\varphi_B}\\ \widetilde{A} & \xrightarrow{\widetilde{f}} & \widetilde{B}\end{array}$$

that is to say, the isomorphisms φ_A also relate the vector prolongations defined by both functors.

The construction of these isomorphisms and checking this property is just a matter of applying the universal property. Let us sketch the proof. The inclusion $k_A\colon A \to \widetilde{A}$ of A in the second vector hull, $\widetilde{A}$, is an affine map that, under the first vector hull functor, has a homogenisation from $\widehat{A}$ to $\widetilde{A}$: this is φ_A. Reversing both functors yields an inverse ψ_A, since one has for instance $\psi_A \circ \varphi_A = \mathrm{Id}_{\widehat{A}}$:

$$\begin{array}{ccc}A & \xrightarrow{j_A} & \widehat{A}\\ & {\scriptstyle k_A}\searrow & \big\downarrow{\scriptstyle \varphi_A=(k_A)^{\widehat{}}}\\ & & \widetilde{A}\end{array}\qquad\begin{array}{ccc}A & \xrightarrow{k_A} & \widetilde{A}\\ & {\scriptstyle j_A}\searrow & \big\downarrow{\scriptstyle \psi_A=(j_A)^{\widetilde{}}}\\ & & \widehat{A}\end{array}$$

As for the commutativity of the preceding square diagram, both $\widetilde{f} \circ \varphi_A$ and $\varphi_B \circ \widehat{f}$ are easily shown to coincide with the homogenisation of $k_B \circ f\colon A \to \widetilde{B}$ with respect to the vector hull $j_A\colon A \to \widehat{A}$, therefore they coincide:



We can give a more abstract picture of this natural equivalence —see for instance [Mac 71]. Consider any choice of the vector hull functor $V\colon \mathbf{Aff} \to \mathbf{Vect}$, and also the forgetful functor $U\colon \mathbf{Vect} \to \mathbf{Aff}$, which sends every vector space to its underlying affine space. They are *adjoint functors* —V is left-adjoint to U and U is right-adjoint to V— since there is an isomorphism $\mathrm{Hom}_{\mathbf{Vect}}(V(\cdot), \cdot) \cong \mathrm{Hom}_{\mathbf{Aff}}(\cdot, U(\cdot))$. In more concrete terms, this is a consequence of the isomorphisms given by proposition 5.9, $\mathscr{L}(\widehat{A}, E) \cong \mathscr{A}(A, E)$, together with the two naturality conditions on A and E, which read $(h \circ f)\hat{} = h\hat{} \circ \widehat{f}$ for affine maps $A \xrightarrow{f} B \xrightarrow{h} E$, and $(T \circ S)|_A = T \circ S|_A$ for linear maps $\widehat{A} \xrightarrow{S} E \xrightarrow{T} F$. Since any two left-adjoints of U are naturally isomorphic, this shows the essential uniqueness of the vector hull functor.

Equivalence between homothetic fields and the dual of affine functions

Let A be an afine space of finite dimension. Write $\hat{A}$ for the homothetic vector fields and $\tilde{A} = \mathscr{A}(A, K)^*$. As we have seen, both can be considered as models for the vector hull of A. We want to give explicitly the map that gives the equivalence between both functors.

Let $X \in \hat{A}$. For an affine function $h\colon A \to K$, define

$$\bar{X}(h) = \vec{h} \circ X + w(X)h,$$

which is another function $\bar{X}(h)\colon A \to K$. This function is constant, and therefore can be identified with a scalar. Moreover, the expression $\bar{X}(h)$ is clearly linear in h. In this way, we have defined a linear form $\bar{X}\colon \mathscr{A}(A, K) \to K$. It turns out that the map

$$\hat{A} \to \tilde{A}, \qquad X \mapsto \bar{X},$$

is an isomorphism of vector spaces.

5.6 The vector hull of an affine bundle

Until now we are talked about vector spaces and affine spaces. We will see that the previous discussion can be naturally extended to vector bundles and affine bundles over a manifold. Recall section 2.4 for the definitions and basic properties of vector bundles and affine bundles.

Let $\pi: A \to M$ be an affine bundle modelled on the vector bundle $\vec{\pi}: \vec{A} \to M$. For each $m \in M$, A_m is an affine space modelled on $\vec{A}_m$. Consider its vector hull

$$j_m: A_m \to \widehat{A}_m$$

and define

$$\widehat{A} := \bigsqcup_{m \in M} \widehat{A}_m,$$

$$\begin{array}{rccc} \widehat{\pi}: & \widehat{A} & \longrightarrow & M \\ & \widehat{a}_m & \longmapsto & m \end{array},$$

and

$$\begin{array}{rccc} j: & A & \longrightarrow & \widehat{A} \\ & a_m & \longmapsto & j_m(a_m) \end{array}.$$

Proposition 5.28 *The set $\widehat{A}$ has a differentiable structure such that $\widehat{\pi}: \widehat{A} \to M$ is a vector bundle and j is an injective affine morphism.*

Proof. For each $m \in M$, its fibre $\widehat{\pi}^{-1}(m)$ is the vector space $\widehat{A}_m$.

Each local trivialisation of π will yield a local trivialisation of $\widehat{\pi}$ in the following way. Consider a local trivialisation $\Phi: \pi^{-1}(U) \to U \times \mathbf{R}^n$ of the affine bundle π over an open subset $U \subset M$. At each point $m \in U$, $\Phi_m = \mathrm{pr} \circ \Phi|_{A_m}: A_m \to \mathbf{R}^n$ is an affine isomorphism, so its vector prolongation $\widehat{\Phi}_m: \widehat{A}_m \to \widehat{\mathbf{R}}^n = \mathbf{R}^{n+1}$ is a linear isomorphism. Therefore,

$$\begin{array}{rccc} \widehat{\Phi}: & \widehat{\pi}^{-1}(U) & \longrightarrow & U \times \mathbf{R}^{n+1} \\ & \widehat{a}_m & \longmapsto & (m, \widehat{\Phi}_m(a_m)) \end{array}$$

is a trivialisation of $\widehat{\pi}$ on U.

Let Φ and Ψ be two trivialisations of π around m. The transition function $\Phi \circ \Psi^{-1}(m): \mathbf{R}^n \to \mathbf{R}^n$ is an affine isomorphism with expression $s^\alpha \mapsto b^\alpha(m) + T^\alpha_\beta(m)s^\beta$. The transition function $\widehat{\Phi} \circ \widehat{\Psi}^{-1}(m): \mathbf{R}^{n+1} \to \mathbf{R}^{n+1}$ between the corresponding trivialisations of $\widehat{\pi}$ is equal to the vector prolongation of $\Phi \circ \Psi^{-1}(m)$, so it has the expression $(s^0, s^\alpha) \mapsto (s^0, b^\alpha(m)s^0 + T^\alpha_\beta(m)s^\beta)$. Thus we see that the transition functions between two trivialisations of $\widehat{\pi}$ take values in linear isomorphisms of $\mathbf{R}^{n+1}$ and are smooth.

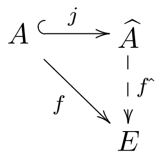
Once we have these trivialisations of $\widehat{\pi}$, it is a standard fact (see for instance [Sau 89]) that they provide $\widehat{A}$ with a differentiable structure such that $\widehat{\pi}$ is a vector bundle.

Finally, since each j_m is an injective affine map, it is clear that j is an injective affine morphism. ∎

The key fact is that the vector hull functor $A \rightsquigarrow \widehat{A}$ from the category of affine spaces to the category of vector spaces is a smooth functor, that is, the assignment $f \rightsquigarrow \widehat{f}$ is a smooth map $\mathscr{A}(A, B) \to \mathscr{L}(\widehat{A}, \widehat{B})$. Therefore, the vector hull functor can be extended to the categories of affine bundles and vector bundles. This is discussed in [KMS 93] in the context of functors between vector spaces (like the duality functor or the tensor product functor), but can be easily adapted to include affine spaces.

Of course, the vector bundle $\widehat{\pi}\colon \widehat{A} \to M$ is called the *vector hull* of the affine bundle $\pi\colon A \to M$. Inherited from the affine space case, the basic properties also hold for affine bundles and its vector hulls. We are going to explicitly state some of them.

Let $E \to M$ be a vector bundle. For each affine bundle morphism $f\colon A \to E$ there exists an unique vector bundle morphism $f\hat{}\colon \widehat{A} \to E$ such that $f\hat{} \circ j = f$:

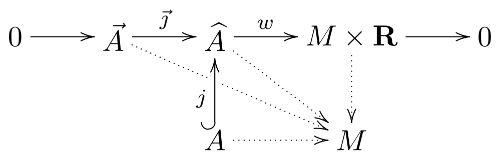


This vector bundle morphism $f\hat{}$ is given, in relation with the affine maps f_m on the fibers, by $f\hat{}(\widehat{a}) = (f_{\widehat{\pi}(\widehat{a})})\hat{}(\widehat{a})$.

We can define the form $w\colon \begin{array}{ccc} \widehat{A} & \longrightarrow & M \times \mathbf{R} \\ \widehat{a}_m & \longmapsto & (m, w_m(\widehat{a}_m)) \end{array}$, which gives the identifications

$$w^{-1}(M \times \{0\}) = \vec{A}, \qquad w^{-1}(M \times \{1\}) = A$$

and the exact sequence



There is a kind of converse statement which allows us to prove that a fixed vector bundle is the vector hull of a given affine bundle. It is the extension of corollary 5.8 to the context of bundles.

Proposition 5.29 *Let $A \to M$ be an affine bundle and $W \to M$ be a vector bundle. Suppose that exists an exact sequence of vector bundle morphisms*

$$0 \longrightarrow \vec{A} \xrightarrow{\alpha} W \xrightarrow{w} M \times \mathbf{R} \longrightarrow 0$$

Then $w^{-1}(1)$ is an affine bundle modelled on $\vec{A}$ (so it is isomorphic to A), and W is canonically isomorphic to the vector hull of $w^{-1}(1)$. ■

In sections 5.7 and 5.8 we will regularly use this proposition to identify a vector bundle as the vector hull of a given affine bundle.

Let $B \to M$ be an affine bundle. For each affine bundle morphism $f\colon A \to B$ there exists an unique vector bundle morphism $\widehat{f}\colon \widehat{A} \to \widehat{B}$ such that $\widehat{f} \circ j_A = j_B \circ f$:

$$\begin{array}{ccc} A & \xrightarrow{j_A} & \widehat{A} \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \widehat{f}} \\ B & \xrightarrow[j_B]{} & \widehat{B} \end{array}$$

We will call $\widehat{f}$ the vector extension of f. As it would be expected, $\widehat{f}$ restricted to a fiber is the vector extension of the restriction of f to that fiber. That is: $\widehat{f}(\widehat{a}) = \widehat{(f_{\widehat{\pi}(\widehat{a})})}(\widehat{a})$.

An affine bundle $\pi\colon A \to M$ has special coordinate systems called affine bundle coordinate systems, each of them defined by a vector bundle coordinate system (x^i, u^α) on $\vec{\pi}$ and a local section $o\colon M \to A$ of π (recall section 2.4).

Every affine bundle coordinate system on A induces a vector bundle coordinate system on its vector hull $\widehat{A}$ in the following way. Consider the local frame $(\vec{e}_\alpha)$ of $\vec{\pi}$ associated with the coordinates (x^i, u^α). We can define the local sections

$$e_0 := j \circ o\colon M \to \widehat{A},$$

and

$$e_\alpha := \vec{j} \circ \vec{e}_\alpha\colon M \to \widehat{A}.$$

Then, (e_0, e_α) is a basis of local sections of $\widehat{\pi}$ that, jointly with the local coordinates $(x^1, \ldots, x^m)$ on M, induce a system (x^i, y^0, y^α) of local coordinates on $\widehat{A}$.

It is worth noting that the coordinate y^0 is always the same, no matter from which affine coordinate system on A is induced. The images of A and $\vec{A}$ in $\widehat{A}$ are respectively $j(A) = \{y^0 = 1\}$ and $\vec{j}(\vec{A}) = \{y^0 = 0\}$.

Let us see which are the transformations between coordinates. If (x^i, y^α) and (r^j, s^β) are two affine bundle coordinate systems on A, related by

$$\begin{cases} x^i(r^j, s^\beta) = x^i(r^j) \\ y^\alpha(r^j, s^\beta) = T_0^\alpha(r^j) + s^\beta T_\beta^\alpha(r^j) \end{cases},$$

then the respectively induced coordinate systems on $\widehat{A}$, (x^i, y^0, y^α) and (r^j, s^0, s^β), are related by

$$\begin{cases} x^i(r^j, s^0, s^\beta) = x^i(r^j) \\ y^0(r^j, s^0, s^\beta) = s^0 \\ y^\alpha(r^j, s^0, s^\beta) = s^0 T_0^\alpha(r^j) + s^\beta T_\beta^\alpha(r^j) \end{cases}.$$

5.7 Vector hulls of jet bundles over R

In this and the following section we will deal with jet bundles. Our aim is to study the vector hull of those jet bundles which are affine bundles. Recall section 2.4 for a basic account of the theory of jet bundles; in any case we will repeat the most relevant facts when they are needed.

First order

Let $\rho\colon M \to \mathbf{R}$ be a bundle with base the real numbers. Consider its first order jet manifold $\mathrm{J}^1\rho$, which is fibred over M and $\mathbf{R}$ with the canonical projections

$$\begin{array}{ccc} \mathrm{J}^1\rho & \xrightarrow{\rho_{1,0}} & M \\ & \searrow^{\rho_1} & \downarrow \rho \\ & & \mathbf{R} \end{array}$$

We take local coordinates (t, q^i) on M, with t the canonical coordinate of $\mathbf{R}$, i.e., the identity. The induced coordinates on $\mathrm{J}^1\rho$ are denoted by (t, q^i, v^i).

We are interested in the affine bundle $\rho_{1,0}$, which is modelled on the vertical vector bundle of ρ, $\mathrm{V}\rho = \operatorname{Ker} \mathrm{T}\rho \subset \mathrm{T}M$. The vertical bundle $\mathrm{V}\rho$ is locally generated by the vector fields $\{\frac{\partial}{\partial q^i}\}$.

There is a canonical embedding (equation (2.4)) of $\mathrm{J}^1\rho$ into $\mathrm{T}M$:

$$\begin{array}{rccc} \iota\colon & \mathrm{J}^1\rho & \to & \mathrm{T}M \\ & \mathrm{j}^1_t\xi & \mapsto & \dot{\xi}(t) \end{array} .$$

In local coordinates, this embedding reads as $\iota(t, q^i, v^i) = (t, q^i; 1, v^i)$, which shows that ι is an affine bundle morphism. We have an exact sequence

$$0 \to \mathrm{V}\rho \xrightarrow{\vec{\iota}} \mathrm{T}M \xrightarrow{\mathrm{d}t} M \times \mathbf{R} \to 0,$$

where $\vec{\iota}$ is the vector bundle morphism associated with the affine bundle morphism ι and, by abuse of notation, $\mathrm{d}t$ denotes the contraction of tangent vectors with the differential form $\mathrm{d}t$.

Now, by proposition 5.29,

$$\widehat{\mathrm{J}^1\rho} = \mathrm{T}M.$$

It is worth noting that the adapted coordinates (t, q^i, v^i) on $\mathrm{J}^1\rho$ and $(t, q^i, \dot{t}, \dot{q}^i)$ on $\mathrm{T}M$, induced in the usual way by the coordinates (t, q^i) on M, are the affine and linear coordinates related with these coordinates (t, q^i), the section $v^i(t, q^i) = 0$ of $\rho_{1,0}$ and the basis of local sections $\{\frac{\partial}{\partial q^i}\}$ of $\mathrm{V}\rho$ in the way described in the preceding section.

Higher order

Consider now the k-th order jet manifold $\mathrm{J}^k\rho$ of the bundle $\rho\colon M \to \mathbf{R}$. Fibred coordinates (t, q^i) of ρ induce natural local coordinates on $\mathrm{J}^k\rho$, which we denote by $(t, q^i_{(0)}, q^i_{(1)}, \ldots, q^i_{(k-1)}, q^i_{(k)})$

The manifold $\mathrm{J}^k\rho$ has canonical projections to the lower-order jet manifolds, but we are especially interested in

$$\begin{array}{c} \mathrm{J}^k\rho \\ \downarrow \rho_{k,k-1} \\ \mathrm{J}^{k-1}\rho \end{array}$$

The bundle $\rho_{k,k-1}$ is affine and it is modelled on the vertical vector bundle of $\rho_{k-1,k-2}$, which is $\mathrm{V}\rho_{k-1,k-2} = \operatorname{Ker} \mathrm{T}\rho_{k-1,k-2} \subset \mathrm{TJ}^{k-1}\rho$. This bundle is locally generated by the vector fields $\{\frac{\partial}{\partial q^i_{(k-1)}}\}$.

There exists a canonical embedding of $\mathrm{J}^k\rho$ into $\mathrm{TJ}^{k-1}\rho$:

$$\begin{array}{rccc} \iota_k \colon & \mathrm{J}^k\rho & \longrightarrow & \mathrm{TJ}^{k-1}\rho \\ & \mathrm{j}^k_t\xi & \longmapsto & (\mathrm{j}^{k-1}\xi)^{\cdot}(t) \end{array}$$

In local coordinates:

$$\iota_k(t, q^i_{(0)}, q^i_{(1)}, \ldots, q^i_{(k-1)}, q^i_{(k)}) = (t, q^i_{(0)}, q^i_{(1)}, \ldots, q^i_{(k-1)}; 1, q^i_{(1)}, \ldots, q^i_{(k)}).$$

Note that $\mathrm{TJ}^{k-1}\rho$ can be identified as the vector hull of $\rho_{k,k-1}$ only if $k = 1$ (the first-order case discussed above), since the dimension of the fibers are, respectively, $nk+1$ and n. Our aim is to identify the vector hull of $\rho_{k,k-1}$ with a suitable subbundle of $\mathrm{T}(\mathrm{J}^k\rho)$.

The Cartan distribution on $\mathrm{J}^{k-1}\rho$ is the distribution generated by the vectors tangent to $(k-1)$-jet prolongations of sections of ρ. We denote it by $C\rho_{k-1,k-2}$, and it is locally generated by the $n+1$ vector fields $\{\frac{\partial}{\partial t} + \sum_{l=0}^{k-2} q^i_{(l+1)} \frac{\partial}{\partial q^i_{(l)}}, \frac{\partial}{\partial q^1_{(k-1)}}, \ldots, \frac{\partial}{\partial q^n_{(k-1)}}\}$.

It is easily seen that $\operatorname{Im}(\iota_k) \subset C\rho_{k-1,k-2}$, and that

$$0 \longrightarrow \mathrm{V}\rho_{k-1,k-2} \xrightarrow{\vec{\iota}_k} C\rho_{k-1,k-2} \xrightarrow{\mathrm{d}t} \mathrm{J}^{k-1}\rho \times \mathbf{R} \longrightarrow 0$$

is an exact sequence.

Therefore,

$$\widehat{\mathrm{J}^k\rho} = C\rho_{k-1,k-2}.$$

5.8 Vector hull of jet bundles over an arbitrary base

First order

Let $\pi \colon M \to B$ be a bundle. We consider here its first-order jet bundle $\mathrm{J}^1\pi$:

$$\begin{array}{ccc} \mathrm{J}^1\pi & \xrightarrow{\pi_{1,0}} & M \\ & \searrow^{\pi_1} & \downarrow \pi \\ & & B \end{array}$$

We take local coordinates (x^i) on B, (x^i, u^α) on M and $(x^i, u^\alpha, u^\alpha_i)$ the induced coordinates on $\mathrm{J}^1\pi$.

The bundle $\pi_{1,0}$ is affine and it is modelled on the vector bundle $\mathrm{V}\pi \otimes \pi^*(\mathrm{T}^*B) \to M$. This can be seen by immersing both bundles into the vector bundle $\operatorname{Hom}(\pi^*\mathrm{T}B, \mathrm{T}M) \to M$ in the following way:

- First, there is an inclusion

$$\mathrm{V}\pi \otimes \pi^*(\mathrm{T}^*B) \simeq \mathrm{Hom}(\pi^*\mathrm{T}B, \mathrm{V}\pi) \subset \mathrm{Hom}(\pi^*\mathrm{T}B, \mathrm{T}M).$$

  Note that this inclusion is characterized by

$$\mathrm{V}\pi \otimes \pi^*(\mathrm{T}^*B) = \{A \in \mathrm{Hom}(\pi^*\mathrm{T}B, \mathrm{T}M) \mid \mathrm{T}\pi \circ A = 0\} \tag{5.1}$$

- On the other hand, each $\mathrm{j}^1_b\phi \in \mathrm{J}^1\pi$ induces a homomorphism

$$\begin{array}{ccc} (\pi^*\mathrm{T}B)_m & \longrightarrow & \mathrm{T}_m M \\ (m, v_b) & \longmapsto & \mathrm{T}_b\phi(v_b) \end{array},$$

  where $m = \phi(b)$. Therefore, we have the inclusion

$$\mathrm{J}^1\pi \subset \mathrm{Hom}(\pi^*\mathrm{T}B, \mathrm{T}M),$$

  which is characterized by

$$\mathrm{J}^1\pi = \{A \in \mathrm{Hom}(\pi^*\mathrm{T}B, \mathrm{T}M) \mid \mathrm{T}\pi \circ A = \mathrm{Id}_{\mathrm{T}B}\}. \tag{5.2}$$

By means of these inclusions, we can add an element of $(\mathrm{V}\pi \otimes \pi^*(\mathrm{T}^*B))_m$ to a jet $\mathrm{j}^1_b\phi$, obtaining another jet in $(\mathrm{J}^1\pi)_m$. It is clear that with this operation, $\mathrm{J}^1\pi$ has the structure of affine bundle with associated vector bundle $\mathrm{Hom}(\pi^*\mathrm{T}B, \mathrm{T}M)$.

Taking into account equations (5.1) and (5.2), it is natural to choose the vector bundle

$$\{A \in \mathrm{Hom}(\pi^*\mathrm{T}B, \mathrm{T}M) \mid \exists\lambda \in \mathbf{R} \text{ such that } \mathrm{T}\pi \circ A = \lambda\,\mathrm{Id}_{\mathrm{T}B}\}$$

as candidate for vector hull of $\mathrm{J}^1\pi$. We temporarily denote this vector bundle by W.

It turns out that this sequence of vector bundle morphisms is exact:

$$0 \longrightarrow \mathrm{V}\pi \otimes \pi^*(\mathrm{T}^*B) \longrightarrow W \stackrel{w}{\longrightarrow} M \times \mathbf{R} \longrightarrow 0\,,$$

where $w(A) = (m, \lambda)$ if A belongs to the fiber over m and $\mathrm{T}\pi \circ A = \lambda\,\mathrm{Id}_{\mathrm{T}B}$. Since we then have that $\mathrm{J}^1\pi = w^{-1}(1)$, the vector hull of $\mathrm{J}^1\pi$ is W, that is:

$$\widehat{\mathrm{J}^1\pi} = \{A \in \mathrm{Hom}(\pi^*\mathrm{T}B, \mathrm{T}M) \mid \exists\lambda \in \mathbf{R} \text{ such that } \mathrm{T}\pi \circ A = \lambda\,\mathrm{Id}_{\mathrm{T}B}\}.$$

Another construction of the vector hull of $\mathrm{J}^1\pi$ can be performed, obtaining the identification of $\widehat{\mathrm{J}^1\pi}$ with the dual of the so-called extended multimomentum bundle $\mathcal{M}\pi$ [CCI 91] [EMR 00] as follows.

The multimomentum bundle $\mathcal{M}\pi \to M$ is the bundle of m-forms on M vanishing by contraction with two vertical vector fields. A form $\omega \in \mathcal{M}\pi$ has the local expression $\omega = p\,\mathrm{d}^m x + p^i_\alpha \mathrm{d}u^\alpha \wedge \mathrm{d}^{m-1}x_i$, where $\mathrm{d}^m x = \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^m$ and $\mathrm{d}^{m-1}x_i = \mathrm{d}x^1 \wedge \ldots \wedge \widehat{\mathrm{d}x^i} \wedge \ldots \wedge \mathrm{d}x^m$. Therefore, we can take $(x^i, u^\alpha, p, p^i_\alpha)$ as a system of local coordinates on $\mathcal{M}\pi$. Denote the dual coordinates on the dual $\mathcal{M}^*\pi$ of the multimomentum bundle by $(x^i, u^\alpha, s, s^\alpha_i)$.

Now we show canonical inclusions of $\mathrm{J}^1\pi$ and its associated vector bundle into $\mathcal{M}^*\pi$:

- We define the inclusion $i\colon \mathrm{J}^1\pi \hookrightarrow \mathcal{M}^*\pi$ by, for each jet $\mathrm{j}^1_b\phi \in \mathrm{J}^1\pi$ and form $\omega \in \mathcal{M}\pi$,
$$\langle i(\mathrm{j}^1_b\phi), \omega\rangle = \alpha \text{ if } (\phi^*\omega)(b) = \alpha\,(\mathrm{d}^m x)_b.$$
In local coordinates, this amounts to $\langle i(x^i, u^\alpha, u^\alpha_i), (x^i, u^\alpha, p, p^i_\alpha)\rangle = p + u^\alpha_i\, p^i_\alpha$ and, subsequently,
$$i(x^i, u^\alpha, u^\alpha_i) = (x^i, u^\alpha, 1, u^\alpha_i). \tag{5.3}$$

- The inclusion $\vec{\imath}\colon \mathrm{V}\pi \otimes \pi^*(\mathrm{T}^*B) \to \mathcal{M}^*\pi$ is defined by
$$\langle \vec{\imath}(\xi), \omega\rangle = \alpha \text{ if } i_\xi\omega = \alpha\,(\mathrm{d}^m x)_b,$$
where b is the base point in B. This locally reads as $\langle \vec{\imath}(x^i, u^\alpha, \xi^\alpha_i), (x^i, u^\alpha, p, p^i_\alpha)\rangle = \xi^\alpha_i\, p^i_\alpha$ so
$$\vec{\imath}(x^i, u^\alpha, \xi^\alpha_i) = (x^i, u^\alpha, 0, \xi^\alpha_i). \tag{5.4}$$

In view of expressions (5.3) and (5.4), it is clear that
$$\widehat{\mathrm{J}^1\pi} \simeq \mathcal{M}^*\pi.$$

Higher order

Now we will study the most general case: the k-th order jet bundle $\pi_{k,k-1}\colon \mathrm{J}^k\pi \to \mathrm{J}^{k-1}\pi$ of a bundle $\pi\colon M \to B$. Note that this case includes the cases studied in section 5.8 (when $k = 1$) and section 5.7 (when $B = \mathbf{R}$).

We will use the multi-index notation $(x^i, u^\alpha, u^\alpha_I)$ (see section 2.7) for the coordinates on $\mathrm{J}^k\rho$.

It is known that $\pi_{k,k-1}$ is an affine bundle (but the remaining $\pi_{k,l}$ are not). The corresponding vector bundle is the bundle whose fiber over a jet $\mathrm{j}^{k-1}_b\phi$ is isomorphic to $\mathcal{S}^k\mathrm{T}^*_bB \otimes \mathrm{V}_{\phi(b)}\pi$, where $\mathcal{S}^k\mathrm{T}^*_bB$ denotes the symmetric product of k copies of T^*_bB. That is, the vector bundle associated with $\pi_{k,k-1}$ is the pullback of $\mathcal{S}^k\mathrm{T}^*B \otimes \mathrm{V}\pi$ over $\mathrm{J}^{k-1}\pi$: $\pi^*_{k-1}(\mathcal{S}^k\mathrm{T}^*B) \otimes \pi^*_{k-1,0}(\mathrm{V}\pi) \to \mathrm{J}^{k-1}\pi$. A proof of this fact using coordinates can be found in [Sau 89] and a more intrinsic one in [KMS 93], where jets of maps in general, not only maps that are sections of a given bundle, are studied. Alternatively, the process that we will follow to find a vector hull also proves it, because the affine and the vector bundles are embedded into a bigger vector bundle where we can add points (jets of $\mathrm{J}^k\pi$) and vectors (elements of $\mathcal{S}^k\mathrm{T}^*B \otimes \mathrm{V}\pi$).

We will use two steps to embed the bundle $\pi_{k,k-1}$ into a vector bundle. First, we use the map $\iota_{1,k-1}\colon \mathrm{J}^k\pi \to \mathrm{J}^1\pi_{k-1}$, defined by
$$\iota_{1,k-1}(\mathrm{j}^k_b\phi) = \mathrm{j}^1_b(\mathrm{j}^{k-1}\phi).$$

The notation $\iota_{1,k-1}$ is that used in [Sau 89] (in fact, there exist maps $\iota_{r,s}\colon \mathrm{J}^{r+s}\pi \to \mathrm{J}^r\pi_s$ for each pair (r, s)). The map $\iota_{1,k-1}$ is an embedding and it is fibered over $\mathrm{J}^{k-1}\pi$, so we can consider $\pi_{k,k-1}\colon \mathrm{J}^k\pi \to \mathrm{J}^{k-1}\pi$ as a subbundle of $(\pi_{k-1})_{1,0}\colon \mathrm{J}^1\pi_{k-1} \to \mathrm{J}^{k-1}\pi$. Taking coordinates $(x^i, u^\alpha, u^\alpha_I; u^\alpha_{;j}, u^\alpha_{I;j})$ on $\mathrm{J}^1\pi_{k-1}$ defined by

$$u^\alpha_{;j}(\mathrm{j}^1_b\psi) = \left.\frac{\partial \psi^\alpha}{\partial x^j}\right|_b,$$

$$u^\alpha_{I;j}(\mathrm{j}^1_b\psi) = \left.\frac{\partial \psi^\alpha_I}{\partial x^j}\right|_b,$$

the embedding $\iota_{1,k-1}$ is written in coordinates as

$$u^\alpha_{;j} \circ \iota_{1,k-1} = u^\alpha_j,$$

$$u^\alpha_{I;j} \circ \iota_{1,k-1} = u^\alpha_{I+1_j}.$$

Thus we see that $\iota_{1,k-1}(\mathrm{J}^k\pi)$ is the submanifold of $\mathrm{J}^1\pi_{k-1}$ where the derivative coordinates are totally symmetric, that is,

$$\text{if } I + 1_j = I' + 1_{j'} \text{ then } u^\alpha_{I;j} = u^\alpha_{I';j'}. \tag{5.5}$$

The second step lies in using the results of the previous subsection, in this case with the bundle π_{k-1} as a starting point. As we saw, there is an inclusion $\mathrm{J}^1\pi_{k-1} \subset \mathrm{Hom}(\pi^*_{k-1}\mathrm{T}B, \mathrm{T}(\mathrm{J}^{k-1}\pi))$, which turns $(\pi_{k-1})_{1,0}$ into an affine subbundle of the vector bundle $\mathrm{Hom}(\pi^*_{k-1}\mathrm{T}B, \mathrm{T}(\mathrm{J}^{k-1}\pi)) \to \mathrm{J}^{k-1}\pi$. We will denote the projection of this vector bundle simply by τ_{k-1}. Combining the two embeddings, we obtain that $\mathrm{J}^k\pi$ can be viewed as an affine subbundle of $\mathrm{Hom}(\pi^*_{k-1}\mathrm{T}B, \mathrm{T}(\mathrm{J}^{k-1}\pi))$. Let us denote this embedding by $\mathrm{hom}^k\colon \mathrm{J}^k\pi \hookrightarrow \mathrm{Hom}(\pi^*_{k-1}\mathrm{T}B, \mathrm{T}(\mathrm{J}^{k-1}\pi))$. The linear map corresponding to a jet $\mathrm{j}^k_b\phi$ is

$$\mathrm{hom}^k(\mathrm{j}^k_b\phi) = \mathrm{T}_b(\mathrm{j}^{k-1}\phi)\colon \mathrm{T}_bB \to \mathrm{T}_{(\mathrm{j}^{k-1}_b\phi)}\mathrm{J}^{k-1}\pi.$$

If the jet $\mathrm{j}^k_b\phi$ has coordinates $(\phi^i, \phi^\alpha, \phi^\alpha_I)$, then the corresponding linear map $\mathrm{hom}^k(\mathrm{j}^k_b\phi)$ is described by

$$\left.\frac{\partial}{\partial x^j}\right|_b \mapsto \left.\frac{\partial}{\partial x^j}\right|_{\mathrm{j}^{k-1}_b\phi} + \sum_{|I|=0}^{k-1} \phi^\alpha_{I+1_j} \left.\frac{\partial}{\partial u^\alpha_I}\right|_{\mathrm{j}^{k-1}_b\phi}. \tag{5.6}$$

Let us denote the components of the matrix of a linear map $A\colon \mathrm{T}_bB \to \mathrm{T}_{(\mathrm{j}^{k-1}_b\phi)}\mathrm{J}^{k-1}\pi$ by $(A^i_{,j}, A^\alpha_{,j}, A^\alpha_{I,j})$, defined by $A(\frac{\partial}{\partial x^j}\big|_b) = A^i_{,j} \frac{\partial}{\partial x^i}\big|_{\mathrm{j}^{k-1}_b\phi} + A^\alpha_{,j} \frac{\partial}{\partial u^\alpha}\big|_{\mathrm{j}^{k-1}_b\phi} + \sum_{|I|=1}^{k-1} A^\alpha_{I,j} \frac{\partial}{\partial u^\alpha_I}\Big|_{\mathrm{j}^{k-1}_b\phi}$.

Taking into account equation (5.6), we see that the fibre $\mathrm{J}^k_{(j^{k-1}_b\phi)}\pi$ of $\pi_{k,k-1}$ over a jet $j^{k-1}_b\phi$ is identified with the affine subspace of linear maps $A\colon \mathrm{T}_bB \to \mathrm{T}_{(\mathrm{j}^{k-1}_b\phi)}\mathrm{J}^{k-1}\pi$ whose matrix has components

$$A^i_{,j} = \delta^i_j, \quad A^\alpha_{,j} = \phi^\alpha_j, \quad A^\alpha_{I,j} = \phi^\alpha_{I+1_j} \text{ for } |I| < k-1, \tag{5.7}$$

and, as a consequence of condition (5.5) on the coordinates of a jet of $\mathrm{J}^k\pi$, such that the remaining components $A^\alpha_{I,j}$ (for $|I| = k-1$) are symmetric on the lower indices, in the sense that

$$A^\alpha_{I,j} = A^\alpha_{I',j'} \text{ if } I + 1_j = I' + 1_{j'}. \tag{5.8}$$

For the sake of shortness, we will say that a linear map $A \in \mathrm{Hom}(\mathrm{T}_b B, \mathrm{T}_{(\mathrm{j}_b^{k-1}\phi)} \mathrm{J}^{k-1}\pi)$ is symmetric if condition (5.8) holds.

Note that the set of equations (5.7) is equivalent to $\mathrm{hom}^{k-1}(\tau_{k-1}(A)) = \mathrm{T}\pi_{k-1,k-2} \circ A$. Therefore, we can describe $\mathrm{hom}^k(\mathrm{J}^k\rho)$ as the subbundle of τ_{k-1} whose fiber over a jet $\mathrm{j}_b^{k-1}\phi$ is

$$\{A \in \mathrm{Hom}(\mathrm{T}_b B, \mathrm{T}_{(\mathrm{j}_b^{k-1}\phi)} \mathrm{J}^{k-1}\pi) \mid \mathrm{T}\pi_{k-1,k-2} \circ A = \mathrm{hom}^{k-1}(\mathrm{j}_b^{k-1}\phi) \text{ and } A \text{ is symmetric}\}. \tag{5.9}$$

Let us consider now the vector bundle $\pi_{k-1}^*(\mathcal{S}^k \mathrm{T}^* B) \otimes \pi_{k-1,0}^*(\mathrm{V}\pi) \to \mathrm{J}^{k-1}\pi$. Its fiber over a jet $\mathrm{j}_b^{k-1}\phi$ is isomorphic to $\mathcal{S}^k \mathrm{T}_b^* B \otimes \mathrm{V}_{\phi(b)}\pi$. We know that $\mathrm{J}^{k-1}_{(\mathrm{j}_b^{k-2}\phi)}\pi = (\pi_{k-1,k-2})^{-1}(\mathrm{j}_b^{k-2}\phi)$ is an affine space with associated vector space $\mathcal{S}^{(k-1)}\mathrm{T}_b^* B \otimes \mathrm{V}_{\phi(b)}\pi$. Therefore, since $\mathrm{j}_b^{k-1}\phi \in \mathrm{J}^{k-1}_{(\mathrm{j}_b^{k-2}\phi)}\pi \subset \mathrm{J}^{k-1}\pi$, we have that

$$\mathcal{S}^{(k-1)}\mathrm{T}_b^* B \otimes \mathrm{V}_{\phi(b)}\pi \backsimeq \mathrm{T}_{(\mathrm{j}_b^{k-1}\phi)}(\mathrm{J}^{k-1}_{(\mathrm{j}_b^{k-2}\phi)}\pi) \subset \mathrm{T}_{(\mathrm{j}_b^{k-1}\phi)}\mathrm{J}^{k-1}\pi.$$

Given that $\mathsf{S}^k \mathrm{T}_b^* B \otimes \mathrm{V}_{\phi(b)}\pi \subset \mathrm{Hom}(\mathrm{T}_b B, \mathsf{S}^{(k-1)}\mathrm{T}_b^* B \otimes \mathrm{V}_{\phi(b)}\pi)$, there is an inclusion

$$\mathsf{S}^k \mathrm{T}_b^* B \otimes \mathrm{V}_{\phi(b)}\pi \subset \mathrm{Hom}(\mathrm{T}_b B, \mathrm{T}_{(\mathrm{j}_b^{k-1}\phi)}\mathrm{J}^{k-1}\pi)$$

for each $\mathrm{j}_b^{k-1}\phi \in \mathrm{J}^{k-1}\pi$. And these inclusions in each fiber sum up to an inclusion of $\pi_{k-1}^*(\mathcal{S}^k \mathrm{T}^* B) \otimes \pi_{k-1,0}^*(\mathrm{V}\pi) \to \mathrm{J}^{k-1}\pi$ as a vector subbundle of τ_{k-1}.

The previous process shows that an element of $\mathcal{S}^k \mathrm{T}_b^* B \otimes \mathrm{V}_{\phi(b)}\pi$ can be considered as a linear map $A\colon \mathrm{T}_b B \to \mathrm{T}_{(\mathrm{j}_b^{k-1}\phi)}\mathrm{J}^{k-1}\pi$ whose image are $\pi_{k-1,k-2}$-vertical vectors. In coordinates, this amounts to

$$A^i_{,j} = 0,$$

$$A^\alpha_{,j} = 0,$$

$$A^\alpha_{I,j} = 0,$$

for $|I| < k-1$. Furthermore, as for the jets, the matrix components $A^\alpha_{I,j}$ for $|I| = k-1$ are symmetric on the lower indices, see equation (5.8). Therefore, $\mathcal{S}^k \mathrm{T}_b^* B \otimes \mathrm{V}_{\phi(b)}\pi$ is identified with the vector subbundle of τ_{k-1} whose fiber over a jet $\mathrm{j}_b^{k-1}\phi$ is

$$\{A \in \mathrm{Hom}(\mathrm{T}_b B, \mathrm{T}_{(\mathrm{j}_b^{k-1}\phi)}\mathrm{J}^{k-1}\pi) \mid \mathrm{T}\pi_{k-1,k-2} \circ A = 0 \text{ and } A \text{ is symmetric}\}. \tag{5.10}$$

Considering (5.9) and (5.10), it is clear that the vector hull $\widehat{\mathrm{J}^k\pi}$ can be identified with the vector subbundle of τ_{k-1}

$$\left\{ A \in \mathrm{Hom}(\pi_{k-1}^*\mathrm{T}B, \mathrm{T}\mathrm{J}^{k-1}\pi) \mid \begin{array}{l} A \text{ symmetric and } \exists \lambda \in \mathbf{R} \text{ such that} \\ \mathrm{T}\pi_{k-1,k-2} \circ A = \lambda \cdot \mathrm{hom}^{k-1}(\tau_{k-1}(A)) \end{array} \right\}.$$

5.9 The vector hull of a second-order tangent bundle

In this section we will study the second order tangent bundle T^2M of a manifold M (recall section 2.8). This case, aside from being interesting by itself, has a particular feature: as in the previous cases, the second order tangent bundle, as a fibration over $\mathrm{T}M$, is an affine bundle naturally included into a vector bundle, but in this case the vector hull can not be embedded into the same vector bundle. As we will see, this fact is related to remark 5.16.

We will use the common notation $(x^i, \dot{x}^i, \ddot{x}^i)$ for the natural coordinates on T^2M induced by the local coordinates (x^i) on M.

The manifold T^2M is fibred over M and $\mathrm{T}M$. The second fibration, $\tau_{2,1}\colon \mathrm{T}^2M \to \mathrm{T}M$, is an affine bundle. Its associated vector bundle is the vertical bundle of $\mathrm{T}M$:

$$\mathrm{V}(\mathrm{T}M) = \operatorname{Ker} \mathrm{T}\tau_M = \left\langle \frac{\partial}{\partial \dot{x}^i} \right\rangle \subset \mathrm{T}(\mathrm{T}M).$$

There exists a canonical inclusion $\iota_{1,1}\colon \mathrm{T}^2M \hookrightarrow \mathrm{T}(\mathrm{T}M)$, defined by $\iota_{1,1}(\ddot{\gamma}(t)) = (\dot{\gamma})^{\cdot}(t)$; in coordinates, $\iota_{1,1}(x^i, \dot{x}^i, \ddot{x}^i) = (x^i, \dot{x}^i; \dot{x}^i, \ddot{x}^i)$.

In spite of having included T^2M into a vector bundle, we can not continue in the same way as with the affine bundles that we have previously discussed. This is because $\widehat{\mathrm{T}^2M}$ is not a subbundle of $\mathrm{T}(\mathrm{T}M)$, since the homogenisation $(\iota_{1,1})\hat{}\colon \widehat{\mathrm{T}^2M} \to \mathrm{T}(\mathrm{T}M)$ is not injective along the fibres where $0 \in \operatorname{Im}(\iota_{1,1})$. These are the fibres whose base point is a zero tangent vector of M. Therefore, we are forced to construct the vector hull by means of the vector prolongation of $\iota_{1,1}$, obtaining

$$\widehat{\mathrm{T}^2M} \subset \mathrm{T}\widehat{(\mathrm{T}M)} \simeq \mathrm{T}(\mathrm{T}M) \oplus \mathbf{R}.$$

Denoting by $(x^i, \dot{x}^i; v^i, \dot{v}^i; s)$ the coordinates of the vector bundle $\mathrm{T}(\mathrm{T}M) \oplus \mathbf{R}$, $\widehat{\mathrm{T}^2M}$ is described as a vector subbundle by the relation $v^i = s\,\dot{x}^i$. In geometric terms,

$$\widehat{\mathrm{T}^2M} = \{(w, s) \in \mathrm{T}(\mathrm{T}M) \oplus \mathbf{R} \mid \mathrm{T}\tau_M(w) = s\,\tau_{\mathrm{T}M}(w)\}.$$

Chapter 6

Time-dependent systems

The purpose of this chapter is to extend the concepts and results of chapter 3 to the time-dependent case.

We study the geometric framework of time-dependent first-order implicit differential equations,

$$F(t, x, \dot{x}) = 0,$$

and the linearly singular case, when F is affine in the velocities,

$$\mathbf{A}(t, x)\dot{x} = b(t, x),$$

where $\mathbf{A}$ is a matrix that is generically singular. We see the geometric formulation of these equations and propose constraint algorithms to find their solutions. We also discuss their relation to time-independent equations, especially in the linearly singular case, where the known constraint algorithm (see section 6.2) for autonomous systems will be useful to the find solutions for time-dependent systems. Let us describe these items in more detail.

First we should point out that our model for the time-dependent configuration space, rather than a trivial product $M = \mathbf{R} \times Q$, is a fibre bundle $\rho: M \to \mathbf{R}$, where the base $\mathbf{R}$ contains the time variable. Such an M is isomorphic to a product $\mathbf{R} \times Q$, but in practical applications there may not be a privileged trivialization, and a possible extension to deal with field theory of course should not be based on a trivial bundle. Some references about time-dependent lagrangian systems are [EMR 91, CF 93, CLM 94] in the product case and [Kru 97, MPV 03, MS 98, LMMMR 02] in the fibre bundle case; see also references therein. Time-dependent systems in general are studied in many books, as for instance [AM 78, Olv 93].

The basic difference between the formulation of the autonomous and the non-autonomous case is the use of tangent bundles and jet bundles respectively. To describe an autonomous differential equation on a configuration space Q we use the tangent bundle $\mathrm{T}Q$, which is a vector bundle. On the other hand, to describe a non-autonomous differential equation on a time-dependent configuration space M, we use its jet bundle

$J^1\rho$, which is an affine bundle over M. Naturally, we will use affine morphisms defined on this affine bundle to describe a linearly singular equation on M.

We also propose a constraint algorithm for time-dependent singular systems. This algorithm is the natural generalization of the algorithm for the autonomous case —both in the general implicit [RR 94, MMT 95] and linearly singular cases [GP 91, GP 92a]. The case of an implicit equation in a product $M = \mathbf{R} \times Q$ has already been discussed in [Del 04]. It is worth noting that constraint algorithms for some particular time-dependent systems have been described in several recent works, as for instance [CF 93, CLM 94, ILMM 99, LMM 96, LMMMR 02, Vig 00]. Since all these systems are of linearly singular type, they are included within our framework. Their various algorithms are also particular instances of the general constraint algorithms that we will study here. So, there is a general procedure that can be applied to these several systems, and their particular details are secondary with respect to the algorithm followed to obtain their solutions.

When studying a time-dependent differential equation, sometimes it is useful to convert it into an equivalent time-independent one. This is even more interesting for implicit equations; for instance, the constraint algorithm for the autonomous case is easier to implement than for the non-autonomous case, because of the fact that vector fields instead of jet fields are used to obtain the constraint functions.

Therefore, we will examine the possibility of associating an autonomous linearly singular system with a time-dependent one, so that the solutions of both systems will be in correspondence. Essentially, we use the canonical inclusion of $J^1\rho$ into TM. In order to perform this association, we propose two different strategies. One possibility is to choose a connection on the jet bundle to induce a splitting of the tangent bundle. The other possibility, which does not make use of any choice, is based on the notion of vector hull that we studied in the previous chapter The main idea is that any affine space A can be canonically embedded as a hyperplane in a vector space $\hat{A}$ —the vector hull of A; with this immersion affine maps can be homogenized, that is, converted into linear maps.

Since our main motivation for studying these systems comes from Euler–Lagrange equations and mechanical systems, where equations of motion are of second order, it is natural to extend the preceding study to second-order implicit and linearly singular equations:

$$F(t, x, \dot{x}, \ddot{x}) = 0, \qquad \mathbf{A}(t, x, \dot{x})\ddot{x} = b(t, x, \dot{x}).$$

It is also of interest to study time-dependent differential equations with constraints and multipliers, of the form

$$\mathbf{A}(t, x)\dot{x} = b(t, x) + \sum_{\mu} u^{\mu}\, h_{\mu}(t, x), \qquad \phi_{\alpha}(t, x) = 0,$$

which are the time-dependent version of the differential equations studied in chapter 4, the kind of equations that arise from a nonholonomic system. Therefore, we will introduce the concept of time-dependent generalized nonholonomic system, which generalizes the construction presented in chapter 4 to the time-dependent case.

As applications of the formalism, we give two descriptions of time-dependent mechanical systems, in the form of time-dependent singular lagrangian systems and in the mixed velocity-momentum description (sometimes called Skinner–Rusk formulation [Ski 83, SR 83]) of time-dependent mechanics [CMC 02]. As a concrete example, we also study a pendulum of variable length.

The chapter is organized as follows. In section 6.1 we study the geometric formulation of time-dependent differential equations, either in the implicit and in the linearly singular case. In section 6.2 we describe constraint algorithms for both cases. In section 6.3 we present two constructions of an autonomous system associated with a given time-dependent system, and an example to illustrate the procedure is given in section 6.4. The extension to second-order equations is presented in section 6.5. Applications to singular lagrangian mechanics are presented in section 6.6. Finally, in section 6.7 we introduce the time-dependent generalized nonholonomic systems.

6.1 Time-dependent systems

In this section we discuss first-order time-dependent singular differential equations. As a model of time-dependent configuration space, we take a fibre bundle $\rho\colon M \to \mathbf{R}$ over the real line (though more general settings could also be considered).

The appropriate geometric framework to deal with derivatives is that of jet bundles, recall the facts and notation given in section 2.7.

Implicit systems

In general, a (time-dependent) *implicit differential equation* is defined by a submanifold $\mathcal{D} \subset \mathrm{J}^1\rho$. A local section $\xi\colon I \to M$ of ρ is a *solution* of the differential equation if

$$\mathrm{j}^1\xi(t) \in \mathcal{D} \tag{6.1}$$

for each t. If the subset $\mathcal{D}$ is locally described in coordinates by some equations $F^\alpha(t, q^i, v^i) = 0$, then the differential equation reads $F^\alpha(t, \xi^i(t), \dot{\xi}^i(t)) = 0$.

Suppose that $\mathcal{D}$ is the image of a jet field, that is, of a section $X\colon M \to \mathrm{J}^1\rho$. Then the solutions of the differential equation are the integral sections of X, which are the solutions of the explicit differential equation

$$\mathrm{j}^1\xi = X \circ \xi.$$

In coordinates, if $X(t, q^i) = (t, q^i, X^j(t, q^i))$, the differential equation reads $\dot{\xi}^i(t) = X^i(t, \xi^j(t))$.

Consider again $\mathcal{D} \subset \mathrm{J}^1\rho$. Given a jet field X, its integral sections are solutions of the implicit equation defined by $\mathcal{D}$ iff

$$X(M) \subset \mathcal{D}. \tag{6.2}$$

So, in a certain sense, solving this equation is equivalent to solving the implicit equation (6.1).

For an explicit differential equation there always exist solutions, and each initial condition $x \in M$ defines a unique maximal solution. For an implicit differential equation existence and uniqueness may fail; in this case one is lead to study the subset of points covered by solutions, and the multiplicity of solutions.

Linearly singular systems

A (time-dependent) *linearly singular system* on M is defined by a vector bundle $\pi\colon E \to M$ and an affine bundle morphism $\mathcal{A}\colon \mathrm{J}^1\rho \to E$:

$$\begin{array}{ccc} \mathrm{J}^1\rho & \xrightarrow{\ \mathcal{A}\ } & E \\ {\scriptstyle \rho_{1,0}}\downarrow & \swarrow{\scriptstyle \pi} & \\ M & & \end{array} \tag{6.3}$$

For the sake of brevity, we will refer to this linearly singular system simply as $\mathcal{A}$.

The system $\mathcal{A}$ has an associated implicit system given by

$$\mathcal{D} = \mathcal{A}^{-1}(0) \subset \mathrm{J}^1\rho, \tag{6.4}$$

provided that this subset is a submanifold —this can be assured, for instance, when $\mathcal{A}$ has constant rank. A local section $\xi\colon I \to M$ is a solution of $\mathcal{D}$, equation (6.1), iff it is a solution of the linearly singular differential equation

$$\mathcal{A} \circ \mathrm{j}^1\xi = 0. \tag{6.5}$$

In local coordinates, the bundle morphisms are given by

$$\pi(t, q^i, u^\alpha) = (t, q^i), \qquad \mathcal{A}(t, q^i, v^i) = (t, q^i, \mathcal{A}^\alpha_j(t, q^i)v^j + c^\alpha(t, q^i)),$$

thus the differential equation reads

$$\mathcal{A}^\alpha_j(t, \xi^i(t))\, \dot\xi^j(t) + c^\alpha(t, \xi^i(t)) = 0.$$

As before, it may be convenient to describe the solutions of the differential equation as integral sections of jet fields. A jet field $X\colon M \to \mathrm{J}^1\rho$ is a solution jet field of $\mathcal{D}$, equation (6.2), iff

$$\mathcal{A} \circ X = 0. \tag{6.6}$$

Then its integral sections are solutions of the differential equation defined by $\mathcal{A}$.

Locally, $X(t, q^i) = (t, q^i, X^i(t, q^i))$ is a solution jet field of $\mathcal{A}$ when

$$\mathcal{A}^\alpha_j(t, q^i) X^j(t, q^i) + c^\alpha(t, q^i) = 0.$$

Let us remark that, instead of a vector bundle, we could have considered an affine bundle E with a section b and an affine bundle morphism $\mathcal{A}\colon \mathrm{J}^1\rho \to E$. This slight generalization does not seem too relevant for applications, and indeed the section b endows E with a vector bundle structure.

6.2 Constraint algorithm

In general, an implicit system does not have solution jet fields, and does not have solution sections passing through every point in M. We want to find a maximal subbundle $\rho'\colon M' \to \mathbf{R}$ of ρ (over $\mathbf{R}$ for simplicity, but more general situations could occur) where there exist solution jet fields $X\colon M' \to \mathrm{J}^1\rho'$ and solution sections $\xi\colon I \to M'$ through every point in M'.

To this end, we can adapt the constraint algorithms of the time-independent case, both for implicit systems [RR 94] [MMT 95] and linearly singular systems [GP 91, GP 92a], to the time-dependent case. A constraint algorithm for a time-dependent implicit equation in a product $M = \mathbf{R} \times Q$ has been recently discussed in [Del 04].

Implicit systems

Let $\mathcal{D} \subset \mathrm{J}^1\rho$ be an implicit system. We say that a 1-jet $y \in \mathcal{D}$ is *integrable* (or *locally solvable*) if there exists a solution $\xi\colon I \to M$ of $\mathcal{D}$ such that $\mathrm{j}^1\xi$ passes through y. One of the purposes of the constraint algorithm is to find the set $\mathcal{D}_{\mathrm{int}}$ of all integrable 1-jets.

If a solution passes through a point $x \in M$, necessarily x belongs to the subset

$$M_{(1)} := \rho_{1,0}(\mathcal{D}). \tag{6.7}$$

Denote by $\rho_{(1)}\colon M_{(1)} \to \mathbf{R}$ the restriction of ρ to $M_{(1)}$. To proceed with the algorithm, we will assume that $\rho_{(1)}$ is a subbundle of ρ. In this case, the inclusion $i_{(1)}\colon M_{(1)} \hookrightarrow M$ has a 1-jet prolongation, $\mathrm{j}^1 i_{(1)}\colon \mathrm{J}^1\rho_{(1)} \hookrightarrow \mathrm{J}^1\rho$. By means of this inclusion, we can define

$$\mathcal{D}_{(1)} := \mathrm{J}^1\rho_{(1)} \cap \mathcal{D}. \tag{6.8}$$

Since the solutions of $\mathcal{D}$ lay on $M_{(1)}$, the integrable jets of $\mathcal{D}$ must be contained in $\mathcal{D}_{(1)}$. If this is a submanifold, we have obtained a new implicit system, now on $M_{(1)}$.

This procedure can be iterated: from $M_{(0)} = M$ and $\mathcal{D}_{(0)} = \mathcal{D}$, and assuming that at each step one obtains subbundles and submanifolds, one may define $M_{(i)} := \rho_{1,0}(\mathcal{D}_{(i-1)})$ and $\mathcal{D}_{(i)} := \mathrm{J}^1\rho_{(i)} \cap \mathcal{D}_{(i-1)}$. The algorithm finishes when, for some k, we have $M_{(k+1)} = M_{(k)}$. In this case, since $\rho_{1,0}(\mathcal{D}_{(k)}) = M_{(k)}$, if we suppose for instance that the projection $\mathcal{D}_{(k)} \to M_{(k)}$ is a submersion, we have that $\mathcal{D}_{\mathrm{int}} = \mathcal{D}_{(k)}$.

Linearly singular systems

Let $\mathcal{A}\colon \mathrm{J}^1\rho \to E$ be a time-dependent linearly singular system as described in section 6.1. We can proceed by applying the preceding algorithm for implicit systems, and also by adapting the algorithm for time-independent linearly singular systems.

So we begin with $\mathcal{D} = \mathcal{A}^{-1}(0)$, which we assume to be a submanifold. As before, the configuration space must be restricted to $M_{(1)} := \rho_{1,0}(\mathcal{D})$, which can also be described as

$$M_{(1)} = \{x \in M \mid 0_x \in \operatorname{Im} \mathcal{A}_x\};$$

note that $0_x \in \operatorname{Im} \mathcal{A}_x$ is the necessary consistency condition for (6.6) to hold on a given point $x \in M$.

As above, we assume that $\rho_{(1)}\colon M_{(1)} \to \mathbf{R}$ is a subbundle of M.

Let us restrict all the data to $M_{(1)}$: $\mathcal{A}_{(1)} := \mathcal{A}|_{\mathrm{J}^1\rho_{(1)}}$, $E_{(1)} := E|_{M_{(1)}}$, and $\pi_{(1)} := \pi|_{M_{(1)}}$. So we obtain a linearly singular system on $M_{(1)}$:

$$\begin{array}{ccc} \mathrm{J}^1\rho_{(1)} & \xrightarrow{\mathcal{A}_{(1)}} & E_{(1)} \\ {\scriptstyle (\rho_{(1)})_{1,0}}\downarrow & \swarrow {\scriptstyle \pi_{(1)}} & \\ M_{(1)} & & \end{array} \tag{6.9}$$

It is clear that the implicit system defined by $\mathcal{A}_{(1)}$ coincides with the implicit system $\mathcal{D}_{(1)}$ obtained above, (6.8), that is,

$$\mathcal{A}_{(1)}^{-1}(0) = \mathrm{J}^1\rho_{(1)} \cap \mathcal{D}. \tag{6.10}$$

Thus, if we assume for instance that each $M_{(i)}$ is a subbundle of $M_{(i-1)}$, we obtain a constraint algorithm for the linearly singular case:

$$M_{(i)} := \{x \in M_{(i-1)} \mid 0_x \in \operatorname{Im}(\mathcal{A}_{(i-1)})_x\},$$

$$\mathcal{A}_{(i)} := \mathcal{A}_{(i-1)}\big|_{\mathrm{J}^1\rho_{(i)}},$$

$$E_{(i)} := E_{(i-1)}\big|_{M_{(i)}},$$

$$\pi_{(i)} := \pi_{(i-1)}\big|_{M_{(i)}}.$$

When the algorithm finishes, we arrive to a final system which is integrable everywhere.

$$\begin{array}{ccc} \mathrm{J}^1\rho_{(k)} & \xrightarrow{\mathcal{A}_{(k)}} & E_{(k)} \\ {\scriptstyle (\rho_{(k)})_{1,0}}\downarrow & \swarrow {\scriptstyle \pi_{(k)}} & \\ M_{(k)} & & \end{array}$$

6.3 From non-autonomous to autonomous systems

It is usual to convert a time-dependent system into a time-independent one by considering the evolution parameter as an new dependent variable. From a geometric viewpoint, this is easily done with an implicit system $D \subset \mathrm{J}^1\rho$ by means of the canonical inclusion $\iota\colon \mathrm{J}^1\rho \to \mathrm{T}M$: its image $E = \iota(D)$ is a submanifold of $\mathrm{T}M$, so it defines an autonomous implicit equation. The equivalence between both equations is immediate:

Proposition 6.1 *A map $\xi\colon \mathbf{R} \to M$ is a solution section of the time-dependent system D iff it is a solution path of the autonomous system E such that $\rho(\xi(t_0)) = t_0$ for any arbitrarily given t_0.* ∎

Now let us focus on the case of a time-dependent linearly singular system $\mathcal{A}$:

$$\begin{array}{ccc} \mathrm{J}^1\rho & \xrightarrow{\ \mathcal{A}\ } & E \\ {\scriptstyle \rho_{1,0}}\downarrow & \swarrow{\scriptstyle \pi} & \\ M & & \end{array} \tag{6.11}$$

We will also relate this system to an autonomous one. The main motivation for finding such a relation is that the constraint algorithm described in the preceding section is easier to implement in the autonomous case. The reason is that, instead of jet fields, vector fields can be used to obtain constraint functions that define the submanifolds in the constraint algorithm, as will be shown later on.

Two constructions to achieve our goal will be proposed. In the first one, we use a connection to define a complement of $\mathrm{V}\rho$ in $\mathrm{T}M$. In the second construction, we use the vector hull functor described in the previous chapter to define vector bundles and morphisms from affine bundles and morphisms.

Jet field construction

Consider the linearly singular system given by (6.11). Let us choose an arbitrary jet field $\Gamma\colon M \to \mathrm{J}^1\rho$. This jet field Γ induces in a natural way a connection $\tilde{\Gamma}$ on ρ (see section 2.7). We define the section of π

$$b_\Gamma := -\mathcal{A} \circ \Gamma\colon M \to E,$$

and the vector bundle morphism

$$A_\Gamma := \vec{\mathcal{A}} \circ v_\Gamma\colon \mathrm{T}M \to E,$$

where $\vec{\mathcal{A}}\colon \mathrm{V}\rho \to E$ is the vector bundle morphism associated with the affine map $\mathcal{A}$.

With these objects we can construct a time-independent linearly singular system:

$$\begin{array}{ccc} \mathrm{T}M & \xrightarrow{\ A_\Gamma \oplus \mathrm{d}t\ } & E \oplus \mathbf{R} \\ {\scriptstyle \tau_M}\downarrow & \nearrow{\scriptstyle b_\Gamma \oplus 1} & \\ M & & \end{array} \tag{6.12}$$

where $E \oplus \mathbf{R}$ denotes the Whitney sum of the vector bundle $\pi\colon E \to M$ and the trivial bundle $M \times \mathbf{R} \to M$.

This system is equivalent to the time-dependent system (6.11) in the sense of the following proposition.

Proposition 6.2 *Consider the time-dependent system given by (6.11). Given any jet field* $\Gamma\colon M \to \mathrm{J}^1\rho$*, we have:*

i) A map $\xi\colon \mathbf{R} \to M$ *is a solution section of the time-dependent system (6.11) if, and only if, it is a solution path of the autonomous system (6.12) such that* $\rho(\xi(t_0)) = t_0$ *for any arbitrarily given* t_0*.*

ii) A map $X\colon M \to \mathrm{J}^1\rho$ *is a solution jet field of the time-dependent system (6.11) if, and only if, considered as a vector field in* M*, it is a solution vector field of the autonomous system (6.12).*

(Note that we are using the embedding $\iota\colon \mathrm{J}^1\rho \to \mathrm{T}M$ defined in section 2.7 to identify jet fields as vector fields.)

Proof. It is immediate, taking into account the local expressions of the equations defined respectively by both systems:

- $\mathcal{A}^\alpha_j(t, q^i)\, v^j + c^\alpha(t, q^i) = 0$
- $\begin{cases} \mathcal{A}^\alpha_j(t, q^i)\, (\dot q^j - \dot t\, \Gamma^j(t, q^i)) = -\mathcal{A}^\alpha_j(t, q^i)\, \Gamma^j(t, q^i) - c^\alpha(t, q^i) \\ \dot t = 1 \end{cases}$ ∎

Vector hull construction

Here we will apply the vector hull functor described in the previous chapter. The affine bundle morphism $\mathcal{A}$ in (6.11) induces a vector bundle morphism $\widehat{\mathcal{A}}$ between the vector hulls of $\mathrm{J}^1\rho$ and E:

$$\begin{array}{ccc} \mathrm{J}^1\rho & \hookrightarrow & \widehat{\mathrm{J}^1\rho} \\ {\scriptstyle\mathcal{A}}\downarrow & & \downarrow{\scriptstyle\widehat{\mathcal{A}}} \\ E & \underset{i}{\hookrightarrow} & \widehat{E} \end{array}$$

The 0 section of $\pi\colon E \to M$ also induces a section $\hat 0$ of $\widehat\pi\colon \widehat E \to M$, defined by $\hat 0 = i \circ 0$. Recall that, since π is a vector bundle, we have the identification $\widehat E = \mathbf{R} \oplus E$, with $\hat 0 = (1, 0)$.

Using the canonical identification of $\widehat{\mathrm{J}^1\rho}$ with $\mathrm{T}M$, we can construct the following linearly singular system:

$$\begin{array}{ccc} \mathrm{T}M & \xrightarrow{\widehat{\mathcal{A}}} & \widehat{E} \\ {\scriptstyle\tau_M}\downarrow & \nearrow{\scriptstyle\hat 0} & \\ M & & \end{array} \tag{6.13}$$

This system is equivalent to the time-dependent system (6.11) in the sense of the following proposition.

Proposition 6.3 *Consider the time-dependent system given by (6.11).*

i) A map $\xi\colon \mathbf{R} \to M$ is a solution section of the time-dependent system (6.11) if, and only if, it is a solution path of the associated autonomous system (6.13) such that $\rho(\xi(t_0)) = t_0$ for any arbitrarily given t_0.

ii) A map $X\colon M \to \mathrm{J}^1\rho$ is a solution jet field of the time-dependent system (6.11) if, and only if, considered as a vector field in M, it is a solution vector field of the associated autonomous system (6.13).

Proof. Again we can prove the result in local coordinates, where the equations of the systems (6.11) and (6.13) read, respectively,

- $c^\alpha(t, q^i) + \mathcal{A}^\alpha_j(t, q^i)v^j = 0$
- $\begin{cases} \dot{t} = 1 \\ c^\alpha(t, q^i)\dot{t} + \mathcal{A}^\alpha_j(t, q^i)\dot{q}^j = 0 \end{cases}$

We have used the local coordinates (t, q^i, u^0, u^α) on $\widehat{E}$ induced by the coordinates (t, q^i, u^α) and the 0 section of π (see section 5.6).

■

6.4 An example: a simple pendulum of given variable length

In this section we present an example to illustrate the conversion of a time-dependent system into a time-independent one (in this case we will use the jet field construction) with the intention of applying the time-independent constraint algorithm to it.

Consider a simple pendulum whose length is given by a time-dependent function $R(t)$. Its equation of motion can be written as [GP 92a]

$$\begin{cases} \dot{x} &= v_x \\ \dot{y} &= v_y \\ \dot{v_x} &= -\tau x \\ \dot{v_y} &= -\tau y - g \\ x^2 + y^2 &= R^2(t) \end{cases},$$

where g is the gravitational acceleration and $\tau R(t)$ is the string tension per unit mass.

This system can be described as a time-dependent linearly singular system in the following way. Take $M := \mathbf{R}^6$, with coordinates $(t, x, y, v_x, v_y, \tau)$, as a configuration manifold fibred over $\mathbf{R}$, with coordinate t. The product $M \times \mathbf{R}^5$ is a trivial vector bundle over M, and the affine bundle morphism $\mathcal{A}\colon \mathrm{J}^1\rho \to M \times \mathbf{R}^5$ defined by

$$\mathcal{A}(\dot{x}, \dot{y}, \dot{v_x}, \dot{v_y}, \dot{\tau})_p = (\dot{x} - v_x, \dot{y} - v_y, \dot{v_x} + \tau x, \dot{v_y} + \tau y + g, -(x^2 + y^2 - R^2(t)))_p,$$

where $p = (t, x, y, v_x, v_y, \tau) \in M$, models the system.

Choosing the connection $\tilde{\Gamma} = \mathrm{d}t \oplus \frac{\partial}{\partial t}$, as described in section 6.3, we can convert this system into an autonomous linearly singular system on M, which can be written as

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \dot{t} \\ \dot{x} \\ \dot{y} \\ \dot{v_x} \\ \dot{v_y} \\ \dot{\tau} \end{pmatrix} = \begin{pmatrix} v_x \\ v_y \\ -\tau x \\ -\tau y - g \\ x^2 + y^2 - R^2(t) \\ 1 \end{pmatrix}$$

We solve this autonomous linearly singular system by means of the constraint algorithm for the autonomous case that we explained in section 3.1. In this case, three steps are needed to solve the system. We give here only the constraint functions ϕ^i and manifolds M_i obtained at each step:

1. $\phi^1 = x^2 + y^2 - R^2(t)$,
   $M_1 = \{\phi^1 = 0\}$,
   and the possible solution vector fields are of the form
   $$X \underset{M_1}{\simeq} \frac{\partial}{\partial t} + v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} - \tau x \frac{\partial}{\partial v_x} - (\tau y + g) \frac{\partial}{\partial v_y} + f \frac{\partial}{\partial \tau},$$
   where $f \in \mathrm{C}^\infty(M_1)$ is a function to be determined.

2. $\phi^2 = X \cdot \phi^1 \underset{M_1}{\simeq} x v_x + y v_y - R R'$,
   $M_2 = \{\phi^1 = \phi^2 = 0\}$.

3. $\phi^3 = X \cdot \phi^2 \underset{M_2}{\simeq} v_x^2 + v_y^2 - \tau R^2 - (R R'' + (R')^2)$,
   $M_3 = \{\phi^1 = \phi^2 = \phi^3 = 0\}$.

4. $\phi^4 = X \cdot \phi^3 \underset{M_3}{\simeq} -4\tau R R' - 3 v_y g - (R R''' + 3 R' R'') - f R^2$.
   The equation $\phi^4 = 0$ determines the function $f \in \mathrm{C}^\infty(M_3)$, so it is not a new constraint and the system is solved.

Using polar coordinates $(t, r, \varphi, v_r, v_\varphi, \tau)$, defined by

$$\left\{ \begin{array}{rcl} x & = & r \cos\varphi \\ y & = & r \sin\varphi \\ v_x & = & v_r \cos\varphi - v_\varphi r \sin\varphi \\ v_y & = & v_r \sin\varphi + v_\varphi r \cos\varphi \end{array} \right. ,$$

we see that the final submanifold M_3 is diffeomorphic to $\mathbf{R} \times \mathrm{T}\mathbf{S}^1$, and it is embedded in M by $(t, \varphi, v_\varphi) \to (t, R(t), \varphi, R'(t), (v_\varphi^2 R(t) - g \sin\varphi - R''(t))/R(t))$. The (unique) solution vector field is

$$X = \frac{\partial}{\partial t} + v_\varphi \frac{\partial}{\partial \varphi} - \frac{2R'(t) v_\varphi + g \cos\varphi}{R(t)} \frac{\partial}{\partial v_\varphi}.$$

6.5 The second-order case

The results of section 6.3 can be extended to consider higher-order implicit and linearly singular differential equations. Of course, the most important case, due to its applications to mechanics, is that of second-order equations, to which we devote this section.

Second-order implicit and linearly singular systems

Similar to the first-order case, a *second-order implicit differential equation* is defined by a submanifold $\mathcal{D} \subset \mathrm{J}^2\rho$. A local section $\xi\colon I \to M$ is a *solution* of the differential equation if $\mathrm{j}^2\xi(t) \in \mathcal{D}$ for each t. In coordinates, this equation can be expressed as $F^\alpha(t, \xi^i(t), \dot\xi^i(t), \ddot\xi^i(t)) = 0$. Like in the first-order case, if $\mathcal{D}$ is the image of a section $X\colon \mathrm{J}^1\rho \to \mathrm{J}^2\rho$, the equation can be written in normal form.

Now let us consider the linearly singular case. A *time-dependent second-order linearly singular system* is defined by a vector bundle $\pi\colon E \to \mathrm{J}^1\rho$ and an affine bundle morphism $\mathcal{A}\colon \mathrm{J}^2\rho \to E$:

$$\begin{array}{ccc} \mathrm{J}^2\rho & \xrightarrow{\mathcal{A}} & E \\ {\scriptstyle \rho_{2,1}}\downarrow & \swarrow{\scriptstyle \pi} & \\ \mathrm{J}^1\rho & & \end{array} \qquad (6.14)$$

Its *solution sections* are sections ξ of ρ such that

$$\mathcal{A} \circ \mathrm{j}^2\xi = 0.$$

Locally this reads

$$\mathcal{A}^\alpha_j(t, \xi^i(t), \dot\xi^i(t))\, \ddot\xi^j(t) + c^\alpha(t, \xi^i(t), \dot\xi^i(t)) = 0. \qquad (6.15)$$

A second-order jet field, that is, a section X of $\rho_{(2,1)}\colon \mathrm{J}^2\rho \to \mathrm{J}^1\rho$, is a *solution jet field* if

$$\mathcal{A} \circ X = 0;$$

locally this reads $\mathcal{A}^\alpha_j(t, q^i, v^i) X^j(t, q^i, v^i) + c^\alpha(t, q^i, v^i) = 0$.

In a similar way of what we did in the previous section, it is interesting to convert the singular system given by (6.14) into a first-order autonomous linearly singular system. As before, we present two constructions of this.

Jet field construction

Choose an arbitrary second-order jet field $\Gamma\colon \mathrm{J}^1\rho \to \mathrm{J}^2\rho$. Again this determines a splitting of the tangent bundle of $\mathrm{J}^1\rho$ as a direct sum $\mathrm{TJ}^1\rho = \mathrm{V}\rho_{1,0} \oplus H_\Gamma$, with associated projections v_Γ and h_Γ. If the local expression of the jet field is $\Gamma(t, q^i, v^i) = (t, q^i, v^i, \Gamma^i(t, q^i, v^i))$ then the two projections are locally given by

$$v_\Gamma(t, q^i, v^i; \dot t, \dot q^i, \dot v^i) = \left(t, q^i, v^i; 0, 0, \dot v^i - \dot t\,\Gamma^i(t, q^i, v^i) - \frac{1}{2}(\dot q^j - \dot t v^j)\frac{\partial \Gamma^i}{\partial v^j}\right),$$

$$h_\Gamma(t, q^i, v^i; \dot t, \dot q^i, \dot v^i) = \left(t, q^i, v^i; \dot t, \dot q^i, \dot t\,\Gamma^i(t, q^i, v^i) + \frac{1}{2}(\dot q^j - \dot t v^j)\frac{\partial \Gamma^i}{\partial v^j}\right).$$

Now we define a section of π

$$b_\Gamma := -\mathcal{A} \circ \Gamma\colon \mathrm{J}^1\rho \to E$$

and a vector bundle morphism

$$A_\Gamma := \vec{\mathcal{A}} \circ v_\Gamma\colon \mathrm{TJ}^1\rho \to E.$$

With these definitions, we obtain an autonomous first-order linearly singular system on the manifold $\mathrm{J}^1\rho$:

$$\begin{array}{ccc} \mathrm{TJ}^1\rho & \xrightarrow{A_\Gamma \oplus S \oplus \mathrm{d}t} & E \oplus \mathrm{V}\rho_{1,0} \oplus \mathbf{R} \\ \downarrow & \nearrow_{b_\Gamma \oplus 0 \oplus 1} & \\ \mathrm{J}^1\rho & & \end{array} \tag{6.16}$$

A result quite similar to proposition 6.2 can be formulated, in the sense that this system is equivalent to the original time-dependent second-order system (6.14). This can be readily seen by comparing the local expression (6.15) with that of the equation defined by system (6.16). We skip the details.

Vector hull construction

As opposite to the first-order case, the vector hull of $\mathrm{J}^2\rho$ can not be identified with a tangent bundle, but rather with a tangent subbundle. As it was shown in the previous chapter, $\widehat{\mathrm{J}^2\rho}$ can be identified with the Cartan distribution $C\rho_{1,0}$ on $\mathrm{J}^1\rho$. Then, as in section 6.3, if we homogenize the system (6.14) we obtain the following:

$$\begin{array}{ccc} C\rho_{1,0} & \xrightarrow{\widehat{\mathcal{A}}} & \widehat{E} \\ {\scriptstyle \tau_{\mathrm{J}^1\rho}}\downarrow & \nearrow_{\hat 0} & \\ \mathrm{J}^1\rho & & \end{array} \tag{6.17}$$

This is an autonomous linearly singular system on $\mathrm{J}^1\rho$, except for the fact that there is only a subbundle $C\rho_{1,0} \subset \mathrm{TJ}^1\rho$ instead of the whole tangent bundle. The interpretation of this system is the same as in the ordinary case, but with the additional requirement that, for a path $\eta\colon I \to \mathrm{J}^1\rho$, its derivative $\dot\eta$ must lie in $C\rho_{1,0}$ —which is a natural condition if η has to be the lift $\mathrm{j}^1\xi$ of a section of ρ. In coordinates, if $\eta = (t, q^i, v^i)$, this requirement amounts to

$$\dot q^i = \dot t\, v^i.$$

Assuming that $\dot\eta \subset C\rho_{1,0}$, the local equations for the path η to be a solution of the system (6.17) are

$$\begin{cases} \dot t = 1 \\ \dot t\, c^\alpha(t, q^i, v^i) + \dot v^j A^\alpha_j(t, q^i, v^i) = 0 \end{cases}$$

Comparing these three equations with equation (6.15), we see that systems (6.14) and (6.17) are equivalent.

Finally, we will show that the system (6.17) can be expressed as a linearly singular system, provided we have an appropriate extension of $\mathcal{A}$. Since E is a vector bundle, we have a canonical identification $\widehat{E} = \mathbf{R} \oplus E$, and the vector extension $\widehat{\mathcal{A}}$ can be written $\widehat{\mathcal{A}} = \mathrm{d}t \oplus \mathcal{A}\hat{}$. Suppose that we have an extension $\bar A\colon \mathrm{T}(\mathrm{J}^1\rho) \to E$ of the map $\mathcal{A}\hat{}\colon C\rho_{1,0} \to E$; in some applications (see next section) there exists a natural extension $\bar A$. Then the system (6.17) can be described as the following linearly singular system:

$$\begin{array}{ccc} \mathrm{TJ}^1\rho & \xrightarrow{\mathrm{d}t\oplus\bar A\oplus S} & \mathbf{R}\oplus E\oplus \mathrm{V}\rho_{1,0} \\ \big\downarrow & \nearrow_{1\oplus 0\oplus 0} & \\ \mathrm{J}^1\rho & & \end{array} \tag{6.18}$$

The only thing to be noted is that $C\rho_{1,0}$ is the kernel of S, see (2.7).

6.6 Some applications to mechanics

Time-dependent Lagrangian systems

Recall the concept of autonomous Lagrangian system studied in section 3.2. Those systems were appropriate to model physical systems whose dynamics are independent of time. However, there are physical systems whose characteristics depend on time. We can model these time-dependent physical systems with a time-dependent version of the Lagrangian systems. The following formulation is standard and it can be found for instance in [CPT 84, LR 89, CF 93].

Definition 6.4 *A* time-dependent Lagrangian system *consists of a fibre bundle $\rho\colon M \to \mathbf{R}$ over the real numbers and a function on the first order jet bundle of ρ, $L\colon \mathrm{J}^1\rho \to \mathbf{R}$. The total space M is called the* time-dependent configuration space *and L is the* lagrangian *of the system.*

Naturally, the real line $\mathbf{R}$ represents the time and a fibre M_t of ρ represents the possible positions of the system at time t. Thus, the Lagrangian L is a function that depends on the positions, the velocities and the time. We note that sometimes (see [Sau 89, EMR 96, GMS 97, LMMMR 02]) the dynamical information is given by a ρ_1-semibasic 1-form $\mathcal{L} \in \Omega^1(\mathrm{J}^1\rho)$, called the lagrangian density. This is done because the formulation of time-dependent mechanics can then be easily extended to field theories. In this case, the lagrangian function is the function $L \in \mathrm{C}^\infty(\mathrm{J}^1\rho)$ such that $\mathcal{L} = L\mathrm{d}t$.

Now, the Hamilton's principle for time-independent systems states that the motions of the system are the extremal sections $\xi\colon I \to M$ of the functional action given by

$$J_L(\xi) = \int_I (\mathrm{j}^1\xi)^*(L\mathrm{d}t).$$

It can be seen that an extremal section of J_L must be a solution of the Euler–Lagrange equations, which in local coordinates (t, q^i, v^i) of $\mathrm{J}^1\rho$ read as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^i} \circ \mathrm{j}^1\xi\right) - \frac{\partial L}{\partial q^i} \circ \mathrm{j}^1\xi = 0. \tag{6.19}$$

The Euler–Lagrange equations can be written in an intrinsic way, using the Poincaré–Cartan forms, defined as

$$\Theta_L = {}^t S \circ \mathrm{d}L + L\mathrm{d}t \,\in\, \Omega^1(\mathrm{J}^1\rho), \tag{6.20}$$

$$\Omega_L = -\mathrm{d}\Theta_L \,\in\, \Omega^2(\mathrm{J}^1\rho),$$

where S is the vertical endomorphism of $\mathrm{J}^1\rho$ (recall equation (2.6)). By contraction, Ω_L defines a morphism $\hat{\Omega}_L\colon \mathrm{T}(\mathrm{J}^1\rho) \to \mathrm{T}^*(\mathrm{J}^1\rho)$. With the notation $\hat{p}_i = \partial L/\partial v^i$, the Poincaré–Cartan forms locally read as

$$\Theta_L = \hat{p}_i(\mathrm{d}q^i - v^i\mathrm{d}t) + L\mathrm{d}t,$$

$$\Omega_L = (\mathrm{d}q^i - v^i\mathrm{d}t) \wedge \mathrm{d}\hat{p}_i - \frac{\partial L}{\partial q^i}\mathrm{d}q^i \wedge \mathrm{d}t.$$

Now, given a vector field $X \in \mathfrak{X}(\mathrm{J}^1\rho)$, we can compute

$$i_X\Omega_L = \Big((X\cdot q^i) - v^i(X\cdot t)\Big)\,\mathrm{d}\hat{p}_i - \left((X\cdot\hat{p}_i) - \frac{\partial L}{\partial q^i}(X\cdot t)\right)\mathrm{d}q^i + \left((X\cdot\hat{p}_i)v^i - \frac{\partial L}{\partial q^i}(X\cdot q^i)\right)\mathrm{d}t. \tag{6.21}$$

Consider the case where X is a second-order vector field,

$$X = \frac{\partial}{\partial t} + v^i\frac{\partial}{\partial q^i} + A^i(t, q, v)\frac{\partial}{\partial v^i}.$$

Recall that every integral curve $\eta\colon I \to \mathrm{J}^1\rho$ of a second order vector field is the prolongation $\mathrm{j}^1\xi$ of a curve $\xi\colon I \to M$. The second-order vector fields can be globally characterized as the vector fields X verifying the two equations:

$$i_X\mathrm{d}t = 1, \qquad S \circ X = 0. \tag{6.22}$$

For a second-order vector field, expression (6.21) simplifies to

$$i_X \Omega_L = \left((X \cdot \hat{p}_i) - \frac{\partial L}{\partial q^i} \right) (v^i \mathrm{d}t - \mathrm{d}q^i).$$

Since the Euler–Lagrange equations (6.19) can be written as $\mathrm{d}\hat{p}_i/\mathrm{d}t = \partial L/\partial q^i$, the integral curves of a second order vector field X are prolongations of a solution of the Euler–Lagrange equations iff

$$i_X \Omega_L = 0.$$

Thus, the three equations

$$\begin{cases} i_X \mathrm{d}t = 1 \\ S \circ X = 0 \\ i_X \Omega_L = 0 \end{cases} \tag{6.23}$$

constitute an intrinsic formulation of the Euler–Lagrange equations (6.19).

A lagrangian $L \in \mathrm{C}^\infty(\mathrm{J}^1\rho)$ (and the associated time-dependent Lagrangian system) is said to be *regular* if any of the following equivalent statements hold:

- The $(2n+1)$-form $\mathrm{d}t \wedge \Omega_L^{\wedge n}$ is a volume form.
- $\operatorname{Ker} \Omega_L$ is a 1-dimensional distribution generated by a second-order vector field Γ_L.

In coordinates, these conditions are equivalent to the regularity of the Hessian matrix

$$\left(\frac{\partial^2 L}{\partial v^i \, \partial v^j} \right).$$

The lagrangians (and Lagrangian systems) that are not regular are called *singular* or *degenerate*.

It is clear that if the Lagrangian is regular, the second-order vector field Γ_L that generates $\operatorname{Ker} \Omega_L$, called the Euler–Lagrange vector field, is the unique solution of (6.23). Therefore, a regular Lagrangian system has solution in every point and the dynamical motions are the curves $\xi = \rho_{1,0}(\eta)$, where η is an integral curve of Γ_L. Aside with the formulation given here, several geometric formulations of time-dependent regular systems have been proposed, all of them equivalent in the sense that they lead to the Euler–Lagrange equations. A review of these alternative formulations can be found in [EMR 91].

On the other hand, if the Lagrangian is singular, equations (6.23) have no solution in general, and, if it exists it will be generally not unique. Following the pattern of Gotay and Nester algorithm for autonomous systems, some procedures have been developed to obtain a constraint submanifold where solutions exist. Cariñena and Fernández-Núñez [CF 93] have studied the case when the configuration space is a trivial product $M = \mathbf{R} \times Q$ and ρ is the trivial bundle $\mathbf{R} \times Q \to \mathbf{R}$. This case is also discussed by de León and coworkers in [CLM 94], where the authors use the notion of cosymplectic structure. The results of this paper are extended to fibre bundles $\rho\colon M \to \mathbf{R}$ (not trivial in general) in [LMM 96], and a more detailed review of singular lagrangian system on jet bundles can be found in [LMMMR 02]. In particular, it is used an auxiliary connection to give an exhaustive description of the constraint functions. O. Krupková [Kru 94] uses another framework, based on the notion of Lepagean form, for time-dependent Lagrangian systems on a trivial bundle $\mathbf{R} \times Q \to \mathbf{R}$. Higher order dynamics are also considered and the constraint algorithm developed in this paper allows general constraint spaces (not necessarily submanifolds). Finally, Vignolo [Vig 00] has studied the constraint algorithm for the gauge-invariant formulation of time-dependent mechanics developed by Massa and coworkers [MPL 00].

Now we will see that equations (6.23) are those of a linearly singular system of the kind studied in the previous sections. Recall (see equation (2.5)) the affine inclusion $\iota_2\colon \mathrm{J}^2\rho \hookrightarrow \mathrm{TJ}^1\rho$, which identifies jet fields $\mathrm{J}^1\rho \to \mathrm{J}^2\rho$ with second-order vector fields on $\mathrm{J}^1\rho$. It is clear that the lagrangian dynamics may be described by the following second-order linearly singular system on Q:

$$\begin{array}{ccc} \mathrm{J}^2\rho & \xrightarrow{\hat{\Omega}_L\circ\iota_2} & \mathrm{T}^*\mathrm{J}^1\rho \\ {\scriptstyle\rho_{2,1}}\downarrow & \swarrow{\scriptstyle\tau^*_{\mathrm{J}^1\rho}} & \\ \mathrm{J}^1\rho & & \end{array} \tag{6.24}$$

Using the vector hull construction described in the preceding section, we can convert this system into a first-order autonomous system on $\mathrm{J}^1\rho$:

$$\begin{array}{ccc} \mathrm{TJ}^1\rho & \xrightarrow{\mathrm{d}t\oplus\hat{\Omega}_L\oplus S} & \mathbf{R}\oplus\mathrm{T}^*\mathrm{J}^1\rho\oplus\mathrm{V}\rho_{1,0} \\ \downarrow & \nearrow{\scriptstyle 1\oplus 0\oplus 0} & \\ \mathrm{J}^1\rho & & \end{array}$$

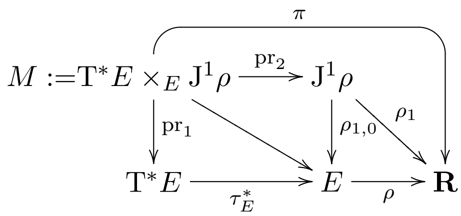
and note that its equations of motion are precisely (6.23).

If the lagrangian is regular, then this linearly singular system is regular; otherwise, the system is singular and the constraint algorithm for linearly singular systems can be applied to obtain the dynamics.

First-order formulation of time-dependent mechanics

As we saw in section 3.2, a mixed lagrangian-hamiltonian formulation of time-independent mechanics was studied geometrically in a series of papers by Skinner and Rusk [Ski 83, SR 83]. Recently, the time-dependent case has been studied in [CMC 02]. We will show how this can be described in our formalism.

Consider a time-dependent Lagragian system defined on a bundle $\rho\colon E \to \mathbf{R}$ with lagrangian $L\colon \mathrm{J}^1\rho \to \mathbf{R}$. In this formulation, the dynamics is represented by a first-order system on the manifold $M := \mathrm{T}^*E \times_E \mathrm{J}^1\rho$. Denote the several projections as in the following diagram:



We can define the following function on M:

$$\mathcal{H} = \langle \mathrm{pr}_1, \mathrm{pr}_2 \rangle - \mathrm{pr}_2^* L,$$

where $\langle\, , \rangle$ denotes the natural pairing between vectors and covectors on E, and the 2-form on M

$$\Omega_{\mathcal{H}} = \mathrm{pr}_1^* \omega_E - \mathrm{d}\mathcal{H} \wedge \mathrm{d}t,$$

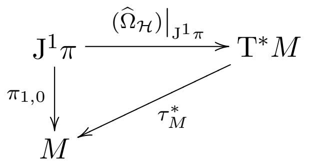
where ω_E is the canonical symplectic form on T^*E.

With these definitions we can write the equations of the dynamics in the Skinner-Rusk formulation, which are

$$\begin{cases} i_Z \Omega_{\mathcal{H}} = 0 \\ i_Z \mathrm{d}t = 1 \end{cases} \tag{6.25}$$

for a vector field $Z \in \mathfrak{X}(M)$. In [CMC 02] it is shown that this first-order formulation is equivalent to the standard one seen in the previous section. More precisely, a solution Z of (6.25) projects under pr_2 onto a vector field X solution of (6.23). The solution of (6.25), if exists, it is not unique because $\ker \Omega_{\mathcal{H}} \cap \ker \mathrm{d}t$ is not equal to 0. Nevertheless, if the lagrangian L is regular, (6.25) has solutions and all of them projects under pr_2 onto the Euler–Lagrange vector field Γ_L. If the lagrangian es singular, a constraint algorithm is applied and, in the best case, there is a final constraint submanifold M_f where solutions of (6.25) exist that projects under pr_2 onto solutions of (6.23) defined on the submanifold $\mathrm{pr}_2(M_f)$.

It is clear that equations (6.25) are those of the time-dependent linearly singular system on M defined by the following diagram:



6.7 Time-dependent generalized nonholonomic systems

In chapter 4 we studied the generalized nonholonomic systems in the time-independent case. Recall that a generalized nonholonomic system is a particular of linearly singular system, obtained from another linearly singular system by performing two operations:

restriction to a submanifold and projection to a quotient. We showed that nonholonomic mechanical systems can be included in this framework.

In this section we present the time-independent version of the generalized nonholonomic systems introduced in section 4.3.

Let $\mathcal{B}\colon \mathrm{J}^1\rho \to G$ be a time-dependent linearly singular system on the bundle $\rho\colon N \to \mathbf{R}$.

A subsystem of $\mathcal{B}$ is defined by a subbundle $\rho|_M\colon M \to \mathbf{R}$ of ρ:

$$\begin{array}{ccc} \mathrm{J}^1(\rho|_M) & \xrightarrow{\mathcal{B}\circ \mathrm{j}^1 i} & G|_M \\ \downarrow & \swarrow & \\ M & & \end{array}$$

where i denotes the inclusion $M \subset N$, which is a bundle morphism over $\mathrm{id}_{\mathbf{R}}$ whose prolongation $\mathrm{j}^1 i$ defines an inclusion $\mathrm{J}^1(\rho|_M) \subset \mathrm{J}^1\rho$. If the system $\mathcal{B}$ has an associated implicit system given by $\mathcal{D} = \mathcal{B}^{-1}(0)$, then the subsystem on M has an associated implicit system given by $(\mathcal{B}\circ \mathrm{j}^1 i)^{-1}(0) = \mathrm{J}^1(\rho|_M)\bigcap \mathcal{D}$. Then, the solution sections of the subsystem are the sections $\gamma\colon I \to M$ of $\rho|_M$ that are solutions of the system $\mathcal{B}$.

On the other hand, a quotient system of the system $\mathcal{B}$ is given by a vector subbundle $G' \subset G$:

$$\begin{array}{ccc} \mathrm{J}^1\rho & \xrightarrow{p\circ\mathcal{B}} & G/G' \\ \downarrow & \swarrow & \\ N & & \end{array}$$

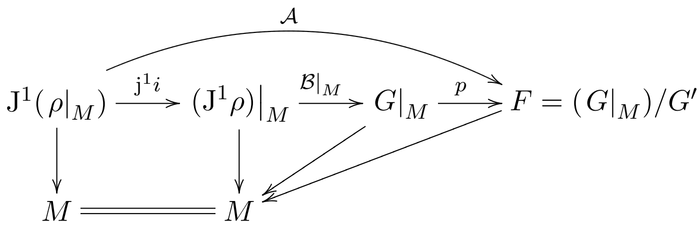
where p is the projection $G \twoheadrightarrow G/G'$. Now, the implicit system associated with this quotient system is given by the submanifold $\mathcal{B}^{-1}(G')$. This implies that if Y is a solution jet field of the system $\mathcal{B}$ defined on certain submanifold N_{f} and Z is any section of $G'|_{N_{\mathrm{f}}}$, then $Y + Z$ is a solution jet field of the quotient system.

These two operations can be combined to obtain another system that we will call time-dependent generalized nonholonomic system. The reason of this name is clear since this construction is a time-dependent version of the generalized nonholonomic systems studied in chapter 4.

Definition 6.5 *Let $\mathcal{B}\colon \mathrm{J}^1\rho \to G$ be a time-dependent linearly singular system on the bundle $\rho\colon N \to \mathbf{R}$, $\rho|_M\colon M \to \mathbf{R}$ a subbundle of ρ and $G' \subset G|_M$ a vector subbundle. The* time-dependent generalized nonholonomic system *defined from $\mathcal{B}$ by the subbundle $\rho|_M$ and the vector subbundle G' is the time-dependent linearly singular system $\mathcal{A} := p\circ \mathcal{B}|_M \circ \mathrm{j}^1 i\colon \mathrm{J}^1(\rho|_M) \to (G|_M)/G'$, where i denotes the inclusion $M \subset N$.*

The manifold M is called the constraint submanifold *and the vector subbundle G' is called the* subbundle of constraint forces

The following diagram shows the construction of a time-dependent generalized nonholonomic system



The implicit system associated with the time-dependent generalized nonholonomic system is $\mathcal{A}^{-1}(0) = \mathrm{J}^1(\rho|_M) \cap \mathcal{B}^{-1}(G')$. Therefore, it is clear that a section $\gamma\colon I \to N$ is a solution of $\mathcal{A}$ if and only if it is contained in the constraint submanifold M and

$$\mathcal{B} \circ \mathrm{j}^1\gamma \in G'. \tag{6.26}$$

Let ϕ^α be a set of functions linearly independent at each point, that locally define M, and Δ_ν a frame for G'. If the affine bundle morphism $\mathcal{B}$ has local expression $\mathcal{B}(t, q^i, v^i) = (t, q^i, \mathcal{B}^l_j(t, q^i)v^j + c^l(t, q^i))$ then equation (6.26) in local coordinates reads as

$$\begin{cases} \phi^\alpha(\gamma(t)) = 0 \\ \mathcal{B}^l_i(\gamma(t))\dot{\gamma}^i(t) + c^l(\gamma(t)) = \lambda^\nu(t)\Delta^l_\nu(\gamma(t)) \end{cases},$$

for a section $\gamma(t) = (t, \gamma^i(t))$ of ρ and multipliers $\lambda^\nu(t)$.

A jet field X of $\rho|_M$ is a solution jet field of the time-dependent generalized nonholonomic system $\mathcal{A}$ if

$$\mathcal{B} \circ X \subset G';$$

in local coordinates,

$$\mathcal{B}^l_i(t, q^j)X^i(t, q^j) + c^l(t, q^j) = v^\nu(t, q^j)\Delta^l_\nu(t, q^j),$$

for some multipliers $v^\nu(t, q^j)$.

In order to find the solutions of the time-dependent generalized nonholonomic system we can use the procedures of sections 6.2 and 6.3.

Now we present an example to illustrate the modelling of time-dependent nonholonomic systems in this framework.

Example 6.6 Homogeneous sphere on a rotating table
This example of time-dependent nonholonomic system is studied in the book by Neimark and Fufaev [NF 72]. Consider an homogeneous sphere, of mass m and radius r, rolling without slipping on a horizontal plane which rotates about a fixed vertical axis with non constant angular velocity $\Omega(t)$. We assume that there are no external forces apart from the constant gravitational force.

We choose a fixed cartesian frame such that the z-axis is the rotation axis. The configuration of the sphere is specified by the coordinates (x, y) of the point of contact of the sphere with the plane and the Eulerian angles (ϕ, θ, ψ). Therefore, the configuration space of the system is the bundle $\rho\colon Q \to \mathbf{R}$, where $Q = \mathbf{R}^2 \times \mathbf{SO}(3) \times \mathbf{R}$ and ρ is the projection on the third factor, which parametrizes the time.

The lagrangian is the kinetic energy of the sphere,

$$L = \frac{1}{2}\left(m(\dot{x}^2 + \dot{y}^2) + I(\dot{\phi}^2 + \dot{\theta}^2 + \dot{\psi}^2 + 2\dot{\phi}\dot{\psi}\cos\theta)\right),$$

where I is the moment of inertia of the sphere about a diameter. The associated Poincaré–Cartan 2-form is

$$\begin{aligned} \Omega_L = \;& m(\mathrm{d}x - \dot{x}\mathrm{d}t) \wedge \mathrm{d}\dot{x} + m(\mathrm{d}y - \dot{y}\mathrm{d}t) \wedge \mathrm{d}\dot{y} + I(\mathrm{d}\phi - \dot{\phi}\mathrm{d}t) \wedge \mathrm{d}(\dot{\phi} + \dot{\psi}\cos\theta) + \\ & I(\mathrm{d}\theta - \dot{\theta}\mathrm{d}t) \wedge \mathrm{d}\dot{\theta} - I\dot{\phi}\dot{\psi}\sin\theta\mathrm{d}\theta \wedge \mathrm{d}t + I(\mathrm{d}\psi - \dot{\psi}\mathrm{d}t) \wedge \mathrm{d}(\dot{\psi} + \dot{\phi}\cos\theta) \end{aligned}$$

The condition of rolling without slipping is expressed by the two constraints

$$\begin{aligned} \Phi_1 &= \dot{x} - (r\sin\phi)\dot{\theta} + (r\sin\theta\cos\phi)\dot{\psi} + \Omega(t)y = 0, \\ \Phi_2 &= \dot{y} + (r\cos\phi)\dot{\theta} + (r\sin\theta\sin\phi)\dot{\psi} - \Omega(t)x = 0. \end{aligned}$$

These affine constraints define a subbundle M of $\rho_1\colon \mathrm{J}^1\rho \to \mathbf{R}$. According to d'Alembert's principle, the constraint force belongs to the subbundle $G' := {}^t S((\mathrm{T}M)^{\vdash})$ of $\mathrm{T}^*(\mathrm{J}^1\rho)\big|_M$, generated by the 1-forms

$$\begin{aligned} \alpha_1 &= (\mathrm{d}x - \dot{x}\mathrm{d}t) - (r\sin\phi)(\mathrm{d}\theta - \dot{\theta}\mathrm{d}t) + (r\sin\theta\cos\phi)(\mathrm{d}\psi - \dot{\psi}\mathrm{d}t), \\ \alpha_2 &= (\mathrm{d}y - \dot{y}\mathrm{d}t) + (r\cos\phi)(\mathrm{d}\theta - \dot{\theta}\mathrm{d}t) + (r\sin\theta\sin\phi)(\mathrm{d}\psi - \dot{\psi}\mathrm{d}t). \end{aligned}$$

Now we have all the elements to construct the time-dependent generalized nonholonomic system

$$\begin{array}{ccc} \mathrm{J}^1(\rho_1|_M) & \xrightarrow{p\circ\hat{\Omega}_L\circ\iota} & \mathrm{T}^*(\mathrm{J}^1\rho)\big|_M / G', \\ \downarrow & \swarrow & \\ M & & \end{array}$$

where p denotes the projection of $\mathrm{T}^*(\mathrm{J}^1\rho)\big|_M$ to the quotient and ι the composition of inclusions $\mathrm{J}^1(\rho_1|_M) \subset \mathrm{J}^1\rho_1 \subset \mathrm{T}(\mathrm{J}^1\rho)$. The solution sections $\gamma\colon I \to M$ of this linearly singular system that are holonomic (that is, $\gamma = \mathrm{j}^1(\rho_{1,0} \circ \gamma)$) are prolongations of solutions of the physical system.

Bibliography

[AM 78] R. Abraham and J. E. Marsden, *Foundations of Mechanics* (2nd ed.), Addison–Wesley, Reading, 1978.

[AMR 83] R. Abraham, J.E. Marsden and T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, Addison-Wesley, Reading, 1983.

[Arn 83] V. I. Arnol'd (ed.), *Dynamical Systems III*, Encyclopædia of Math. Sciences 3, Springer, Berlin, 1988.

[BB 75] Y. Bamberger et J.-P. Bourguignon, "Torseurs sur un espace affine", pp. 151–202 in L. Schwartz, *Les tenseurs*, Hermann, Paris, 1975.

[BCP 96] K. E. Brenan, S. L. Campbell and L. R. Petzold, *Numerical solution of initial-value problems in differential-algebraic equations*, Classics in Applied Mathematics 14, SIAM, Philadelphia, 1996.

[Ber 04] W. Bertram, "From linear algebra via affine algebra to projective algebra", *Lin. Alg. Appl.* **378** (2004) 109–134.

[Ber 77] M. Berger, *Géométrie I*, Nathan, Paris, 1977.

[BG 55] P.G. Bergmann and I. Goldberg, "Dirac bracket transformations in phase space", *Phys. Rev.* **98** (1955) 531–538.

[BGM 96] L. Bates, H. Graumann and C. MacDonnell, "Examples of gauge conservation laws in nonholonomic systems", *Rep. Math. Phys.* **37** (1996) 295–308.

[BGPR 86] C. Batlle, J. Gomis, J. M. Pons and N. Román-Roy, "Equivalence between the lagrangian and hamiltonian formalisms for constrained systems", *J. Math. Phys.* **27** (1986) 2953–2962.

[BKMM 96] A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden, R. M. Murray, "Nonholonomic mechanical systems with symmetry", *Arch. Rational Mech. Anal.* **36** (1996) 21–99.

[Bla 00] G. Blankenstein *Implicit Hamiltonian systems: Symmetry and Interconnection*, PhD Thesis, University of Twente, 2000.

[Bou 62] N. Bourbaki, *Algèbre, châpitre 2*, 3ème édition, Actualités Scientifiques et Industrielles 1236, Hermann, Paris, 1962.

[Bou 70] N. Bourbaki, *Algèbre, châpitres 1–3*, Hermann, Paris, 1970.

[BS 01] G. Blankenstein and A.J. van der Schaft, "Symmetry and reduction in implicit generalized Hamiltonian systems", Rep. Math. Phys. **47** (2001) 57–100.

[BŚ 93] L. Bates, J. Śniatycki, "Nonholonomic reduction", *Rep. Math. Phys.* **32** (1993) 99–115.

[Car 90] J. F. Cariñena, "Theory of singular lagrangians", *Fortschrit. Phys.* **38** (1990) 641–679.

[Car 92] M.P. do Carmo, *Riemannian geometry*, Birkhäuser, Boston–Basel–Berlin, 1992.

[CCI 91] J.F. Cariñena, M. Crampin and L.A. Ibort, "On the multisymplectic formalism for first order field theories", *Diff. Geom. Appl.* **1** (1991) 345–374.

[CF 93] J. F. Cariñena and J. Fernández-Núñez, "Geometric theory of time-dependent singular lagrangians", *Fortschr. Phys.* **41** (1993) 517–552.

[Che 34] N.G. Chetaev, "On Gauss principle", *Izv. Fiz-Mat. Obsc. Kazan Univ.* **7** (1934) 68-71.

[CILM 04] H. Cendra, A. Ibort, M. de León and D. Martín de Diego, "A generalization of Chetaev's principle for a class of higher order nonholonomic constraints", *J. Math. Phys.* **45**(7) (2004) 2785–2801.

[CL 87] J. F. Cariñena and C. López, "The time-evolution operator for singular lagrangians", *Lett. Math. Phys.* **19** (1987) 203–210.

[CL 99] J. Cortés and M. de León, "Reduction and reconstruction of the dynamics of nonholonomic systems", *J. Phys. A: Math. Gen.* **32** (1999) 8615–8645.

[CLM 94] D. Chinea, M. de León and J. C. Marrero, "The constraint algorithm for time-dependent lagrangians", *J. Math. Phys.* **35** (1994) 3410–3447.

[CLMM 02] J. Cortés, M. de León, D. Martín de Diego and S. Martínez, "Geometric description of vakonomic and nonholonomic dynamics. Comparison of solutions", *SIAM J. Control Optim.* **41**(5) (2002) 1389–1412 (electronic).

[CMC 02] J. Cortés, S. Martínez and F. Cantrijn, "Skinner-Rusk approach to time-dependent mechanics", *Phys. Lett. A* **300** (2002) 250–258.

[CMF 92] J. F. Cariñena, E. Martínez and J. Fernández-Núñez, "Noether's theorem in time-dependent Lagrangian mechanics", *Rep. Math. Phys.* **31**(2) (1992), 189–203.

[Cor 02] J. Cortés Monforte, *Geometric, control and numerical aspects of nonholonomic systems*, Lecture Notes in Mathematics 1793, Springer-Verlag, Berlin, 2002.

[Cou 90] T.J. Courant, "Dirac manifolds", *Trans. Amer. Math. Soc.* **319**(2) (1990) 631–661.

[CPT 84] M. Crampin, G. E. Prince and G. Thompson "A geometrical version of the Helmholtz conditions in time-dependent Lagrangian dynamics", *J. Phys. A* **17** (1984) 1437–1447.

[CR 93] J. F. Cariñena and M. F. Rañada, "Lagrangian systems with constraints: a geometric approach to the method of Lagrange multipliers". *J. Phys. A: Math. Gen.* **26** (1993) 1335–1351.

[Cra 83] M. Crampin, "Tangent bundle geometry for lagrangian dynamics", *Int. J. Theor. Phys.* **16** (1983), 3755–3772.

[Del 04] M. Delgado, *Métodos geométricos en problemas de control óptimo singulares: fundamentos y aplicaciones*, Ph. D. thesis, Univ. Carlos III de Madrid, 2004.

[Dir 64] P.A.M. Dirac, *Lectures on Quantum Mechanics*, Belfer Graduate School of Science, Yeshiva University, New York, 1964.

[Dor 93] I. Dorfman, *Dirac structures and integrability of nonlinear evolution equations*, John Wiley, Chichester, 1993.

[DS 99] M. Dalsmo and A. van der Schaft, "On representations and integrability of mathematical structures in energy-conserving physical systems", SIAM J. Control Optim. **37**(1) (1999) 54–91.

[EMR 00] A. Echeverría-Enríquez, M. C. Muñoz-Lecanda and N. Román-Roy, "On the multimomentum bundles and the Legendre maps in field theories", *Rep. Math. Phys.* **45** (2000) 85–105.

[EMR 91] A. Echeverría Enríquez, M. C. Muñoz Lecanda and N. Román Roy, "Geometrical setting of time-dependent regular systems. Alternative models", *Rev. Math. Phys.* **3** (1991) 301–330.

[EMR 96] A. Echeverría-Enríquez, M.C. Muñoz-Lecanda and N. Román-Roy, "Geometry of Lagrangian first-order classical field theories", *Fortschr. Phys.* **44** (1996) 235–280.

[FP 90] C. Ferrario, A. Passerini, "Symmetries and constants of motion for constrained Lagrangian systems: a presymplectic version of the Noether theorem", *J. Phys. A* **23** (1990) 5061–5081.

[Fre 73] J. Frenkel, *Géométrie pour l'éève-professeur*, Hermann, Paris, 1973.

[Gal 01] J. Gallier, *Geometric methods and applications for computer science and engineering*, Springer-Verlag, New York, 2001.

[GGU 03] K. Grabowska, J. Grabowski and P. Urbański, "AV-differential geometry: Poisson and Jacobi structures", preprint, 2004.

[GM 04a] X. Gràcia and R. Martín, "Nonholonomic systems and the relativistic particle", pp. 223–228 in *Proceedings of the XI Fall Workshop on Geometry and Physics, Oviedo, 2002*, Publicaciones de la RSME **6**, Real Sociedad Matemática Española, Madrid, 2004.

[GM 04b] X. Gràcia and R. Martín, "Time-dependent singular systems", pp. 155–160 in *Proceedings of the XII Fall Workshop on Geometry and Physics, Coimbra, 2003*, Publicaciones de la RSME **7**, Real Sociedad Matemática Española, Madrid, 2004.

[GM 05a] X. Gràcia and R. Martín, "Regularity and symmetries of nonholonomic systems", *J. Phys. A: Math. Gen.* **38** (2005) 1071–1087.

[GM 05b] X. Gràcia and R. Martín, "Geometric aspects of time-dependent singular differential equations", *Int. J. Geom. Methods Mod. Phys.* **2** (2005), 597–618.

[GM 05c] X. Gràcia and R. Martín, "Vector hulls and jet bundles", pp. 215–220 in *Proceedings of the XIII Fall Workshop on Geometry and Physics, Murcia, 2004*, Publicaciones de la RSME **9**, Real Sociedad Matemática Española, Madrid, 2005.

[GM 06] X. Gràcia and R. Martín, "Vector hulls of affine spaces and affine bundles", submitted for publication, 2006.

[GMM 03] X. Gràcia, J. Marín-Solano and M.C. Muñoz-Lecanda, "Some geometric aspects of variational calculus in constrained systems" *Rep. Math. Phys.* **51** (2003) 127–148.

[GMR 04] X. Gràcia, M. C. Muñoz-Lecanda and N. Román-Roy, "On some aspects of the geometry of differential equations in physics", *Int. J. Geometric Methods Mod. Phys.* **1** (2004) 265–284.

[GMS 97] G. Giachetta, L. Mangiarotti and G. Sardanashvili, "Differential geometry of time-dependent mechanics", dg-ga/9702020.

[GN 79] M.J. Gotay and J.M. Nester, "Presymplectic lagrangian systems I: the constraint algorithm and the equivalence theorem", *Ann. Inst. H. Poincaré A* **30** (1979) 129–142.

[GNH 78] M. J. Gotay, J. M. Nester and G. Hinds, "Presymplectic manifolds and the Dirac-Bergmann theory of constraints", *J. Math. Phys.* **19** (1978) 2388–2399.

[God 69] C. Godbillon, *Géometrie différentielle et mécanique analytique*, Hermann, Paris, 1969.

[GP 00] X. Gràcia and J.M. Pons, "Canonical Noether symmetries and commutativity properties for gauge systems", *J. Math. Phys.* **41**(11) (2000) 7333–7351.

[GP 01] X. Gràcia and J. M. Pons, "Singular lagrangians: some geometric structures along the Legendre map", *J. Phys. A: Math. Gen.* **34** (2001) 3047–3070.

[GP 02] X. Gràcia and J. M. Pons, "Symmetries and infinitesimal symmetries of singular differential equations", *J. Phys. A: Math. Gen.* **35** (2002) 5059–5077.

[GP 88] X. Gràcia and J. M. Pons, "Gauge generators, Dirac's conjecture, and degrees of freedom for constrained systems", *Ann. Physics* **187** (1988) 355–368.

[GP 89] X. Gràcia and J. M. Pons, "On an evolution operator connecting lagrangian and hamiltonian formalisms", *Lett. Math. Phys.* **17** (1989) 175–180.

[GP 91] X. Gràcia and J. M. Pons, "Constrained systems: a unified geometric approach", *Int. J. Theor. Phys.* **30** (1991) 511–516.

[GP 92a] X. Gràcia and J. M. Pons, "A generalized geometric framework for constrained systems", *Diff. Geom. Appl.* **2** (1992) 223–247.

[ILMM 96] L. A. Ibort, M. de León, G. Marmo amd D. Martín de Diego, "Nonholonomic constrained systems as implicit differential equations", *Rend. Sem. Mat. Univ. Politec. Torino* **54** (1996) 295–317.

[ILMM 99] A. Ibort, M. de León, J. C. Marrero and D. Martín de Diego, "Dirac brackets in constrained dynamics", *Fortsch. Phys.* **47** (1999) 459–492.

[IREM 75] IREM de Strasbourg, *Le livre du probème, vol. 5. Calcul barycentrique*, CEDIC, Paris, 1975.

[JS 98] J. V. José and E. J. Saletan, *Classical dynamics*, Cambridge University Press, Cambridge, 1998.

[KM 98] W. S. Koon and J. E. Marsden, "Poisson reduction for nonholonomic mechanical systems with symmetry", *Rep. Math. Phys.* **42** (1998) 101–134.

[KM 01] O. Krupková and J. Musilová, "The relativistic particle as a mechanical system with non-holonomic constraints", *J. Phys. A: Math. Gen.* **34** (2001) 3859–3875.

[KMS 93] I. Kolář, P.W. Michor and J. Slovák, *Natural operations in differential geometry*, Springer-Verlag, Berlin, 1993.

[KN 63] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Interscience Publishers, Wiley, New York, 1963.

[Koi 92] J. Koiller, "Reduction of some classical non-holonomic systems with symmetry", *Arch. Rational Mech. Anal.* **118** (1992) 113–48.

[Koz 83] V.V. Kozlov, "Realization of nonintegrable constraints in classical mechanics", (Russian), *Dokl. Akad. Nauk SSSR* **272** (1983) 550–554; English translation: *Soviet Phys. Dokl.* **28** (1983) 735–737.

[Kru 02] O. Krupková, "Recent results in the geometry of constrained systems", *Rep. Math. Phys.* **49** (2002) 269–278.

[Kru 94] O. Krupková, "A geometric setting for higher-order Dirac-Bergmann theory of constraints", J. Math. Phys. **35** (1994) 6557–6576.

[Kru 97] O. Krupková, *The geometry of ordinary variational equations*, LNM 1678, Springer, Berlin, 1997.

[Lan 85] S. Lang, *Differential manifolds* (2nd ed.), Springer-Verlag, New York, 1985.

[Lee 03] J. M. Lee, *Introduction to smooth manifolds*, Springer-Verlag, New York, 2003.

[Lee 97] J. M. Lee, *Riemannian manifolds*, Springer-Verlag, New York, 1997.

[LM 87] P. Libermann and Ch.-M. Marle, *Symplectic geometry and analytical mechanics*, D. Reidel Publishing Co., Dordrecht, 1987.

[LM 95] A. D. Lewis and R. M. Murray "Variational principles for constrained systems: theory and experiment", *Internat. J. Non-Linear Mech.* **30** (1995) 793–815.

[LM 96] M. de León, D. Martín de Diego, "On the geometry of non-holonomic lagrangian systems", *J. Math. Phys.* **37** (1996) 3389–414.

[LMM 00] M. de León, J.C. Marrero and D. Martín de Diego, "Vakonomic mechanics versus non-holonomic mechanics: a unified geometrical approach", *J. Geom. Phys.* **35** (2000) 126–144.

[LMM 96] M. de León, J. Marín-Solano and J. C. Marrero, "The constraint algorithm in the jet formalism", *Diff. Geom. Appl.* **6** (1996) 275–300.

[LMMMR 02] M. de León, J. Marín-Solano, J. C. Marrero, M. C. Muñoz-Lecanda and N. Román-Roy, "Singular lagrangian systems on jet bundles", *Fortsch. Phys.* **50** (2002) 105–169.

[LR 89] M. de León and P.R. Rodrigues, *Methods of Differential Geometry in Analytical Mechanics*, North-Holland Math. Studies 158, Elsevier, Amsterdam, 1989.

[Mac 71] S. Mac Lane, *Categories for the working mathematician*, GTM 5, Springer-Verlag, New York, 1971.

[Mar 03] C.-M. Marle "On symmetries and constants of motion in hamiltonian systems with nonholonomic constraints", *Classical and quantum integrability (Warsaw, 2001)*, 223–242, Banach Center Publ., 59, Polish Acad. Sci., Warsaw, 2003.

[Mar 98] C.-M. Marle "Various approaches to conservative and nonconservative nonholonomic systems", *Rep. Math. Phys.* **42** (1998) 211–229.

[Mes 03] T. Mestdag, *Berwald-type connections in time-dependent mechanics and dynamics on affine Lie algebroids*, PhD Thesis, U. Gent, 2003.

[MMS 02] E. Martínez, T. Mestdag and W. Sarlet, "Lie algebroid structures and lagrangian systems on affine bundles", *J. Geom. Phys.* **44** (2002) 70–95.

[MMT 95] G. Mendella, G. Marmo and W. M. Tulczijew, "Integrability of implicit differential equations", *J. Phys. A: Math. Gen.* **28** (1995) 149–163.

[MPL 00] E. Massa, E. Pagani and P. Lorenzoni "On the gauge structure of classical mechanics", *Transport Theory Statist. Phys.* **29**(1-2) (2000) 69–91.

[MPV 03] E. Massa, E. Pagani and S. Vignolo, "Legendre transformation and analytical mechanics: a geometric approach", *J. Math. Phys.* **44** (2003) 1709–1722.

[MR 99] J.E. Marsden and T.S. Ratiu, *Introduction to mechanics and symmetry. A basic exposition of classical mechanical systems* (2nd ed.), Springer-Verlag, New York, 1999.

[MS 98] L. Mangiarotti and G. Sardanashvily, *Gauge mechanics*, World Scientific, Singapore, 1998.

[MT 78] M. R. Menzio and W. M. Tulczyjew, "Infinitesimal symplectic relations and generalized hamiltonian dynamics", *Ann. Inst. Henri Poincaré A* **28** (1978) 349–367.

[MVB 02] E. Massa, S. Vignolo and D. Bruno, "Non-holonomic lagrangian and hamiltonian mechanics: an intrinsic approach", J. Phys. A: Math. Gen. **35** (2002) 6713–6742.

[NF 72] J. Neimark and N. Fufaev, *Dynamics of Nonholonomic Systems*, Transactions of Mathematical Monographs **33**, AMS, Providence, RI, 1972.

[Olv 93] P. J. Olver, *Applications of Lie groups to differential equations*, 2nd ed., GTM 107, Springer, New York, 1993.

[Pon 88] J.M. Pons, "New relations between Hamiltonian and Lagrangian constraints", *J. Phys. A* **21** (1988) 2705–2715.

[PV 00] F. Pugliese and A. M. Vinogradov, "On the geometry of singular Lagrangians", *J. Geom. Phys.* **35** (2000) 35–55.

[Ram 89] L. Ramshaw "Blossoms are polar forms", Digital Equipment Corporation, research report, 1989.

[Ros 77] R. Rosenberg, *Analytical Dynamics*, Plenum, New York, 1977.

[RR 94] P. J. Rabier and W. C. Rheinboldt, "A geometric treatment of implicit differential-algebraic equations", *J. Differential Equations* **109** (1994) 110–146.

[Sau 89] D. J. Saunders, *The Geometry of Jet Bundles*, London Mat. Soc. Lecture Note Series 142, Cambridge University Press, Cambridge, 1989.

[Sch 75] L. Schwartz, *Les tenseurs*, Hermann, Paris, 1975.

[Ski 83] R. Skinner, "First-order equations of motion for classical mechanics", *J. Math. Phys.* **24** (1983) 2581–2588.

[SM 95a] A. J. van der Schaft and B. M. Maschke, "Mathematical modelling of constrained Hamiltonian systems", In *Preprints of the 3rd Nonlinear Control Systems Design Symposium, NOLCOS* (1995) 678–683.

[SM 95b] A. J. van der Schaft and B. M. Maschke, "The Hamiltonian formulation of energy conserving physical systems with external ports", *Archiv für Elektronik und Übertragungstechnik* **49** (1995) 362–371.

[Sou 97] J. M. Souriau, *Structure of dynamical systems. A symplectic view of physics*, Translated from the French by C. H. Cushman-de Vries. Birkhäuser Boston, 1997.

[SR 83] R. Skinner and R. Rusk, "Generalized hamiltonian dynamics. I. Formulation on $T^*Q \oplus TQ$", *J. Math. Phys.* **24** (1983) 2589–2594.

[Tis 83] C. Tisseron, *Géométrie affine, projective et euclidienne*, Hermann, Paris, 1983.

[Tul 86] W. M. Tulczyjew, "Differential geometry of mechanical systems with constraints", *Atti Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur.* **120** (1986) 211–216.

[Vai 94] I. Vaisman, *Lectures on the geometry of Poisson manifolds*, Progress in Mathematics, 118. Birkhäuser Verlag, Basel, 1994.

[vdS 98] A. J. van der Schaft, "Implicit Hamiltonian systems with symmetry", *Rep. Math. Phys.* **41**(2) (1998) 203–221.

[VF 72] A. M. Vershik, L. D. Faddeev, "Differential geometry and lagrangian mechanics with constraints", *Sov. Phys. Dokl.* **17** (1972) 34–36.

[Vig 00] S. Vignolo, "A new presymplectic geometrical framework for time-dependent lagrangian systems: the constraint algorithm and the second-order differential equation problem", *J. Phys. A: Math. Gen.* **33** (2000) 5117–5135.