Control of Non Holonomic Systems by Active Constraints

Franco Rampazzo
University of Padova
rampazzo@math.unipd.it

SADCO Summer School
Imperial College, London
September 5-9, 2011

This version. Due to time constraints these notes are in a preliminary form. They surely need much revision. An online copy will be soon available on www.math.unipd.it/∼rampazzo. Updated versions will be uploaded at the same address. The author wishes to express his gratitude to Elena Bossolini and Dario Paccagnan who have helped him in the technical preparation of these notes in due time. Of course, the blame for still existing misprints or errors is to be put entirely on the author.

Contents

Chapter 1

Introduction

1.1 Coordinates (=constraints) as controls

A mechanical system can be controlled in two fundamentally different ways. In a commonly adopted framework [18, 39], the controller modifies the time evolution of the system by applying additional forces. This leads to a control problem in standard form, where the time derivatives of the state variables depend continuously on the control function.

In other situations, also physically realistic, the controller acts on the system by directly assigning (as *controls*) the values of some of the coordinates. [1] The evolution of the remaining coordinates can then be determined by solving a control system where the vector field is a quadratic polynomial of the time derivatives of the control-coordinates.

This alternative point of view was introduced, independently, in [**?**] and in [29]. An akin approach can also be found within in the literature of underactuated system (see e.g. [3]).

Pre-assigning some coordinates' evolution have a global counterpart in the following scheme: consider a system whose configuration space is a differential manifold $\mathcal{Q}$, and assume that a surjective submersion[2]

$$\pi : \mathcal{Q} \to \mathcal{U}$$

is given (see Figure 1.1). A *control* $t \mapsto \mathbf{u}(t) \in \mathcal{U}$ acts on the system as a moving bilateral constraint, meaning that a trajectory $\mathbf{q}(\cdot)$ agrees with this constraint if $\pi(\mathbf{q}(t)) = \mathbf{u}(t)$ for every time t. As it is well-known from

[1] In a more intrinsic language one says that the controls are additional time-dependent (bilateral) constraints.

[2] See Definition A.2.1 in Appendix A.

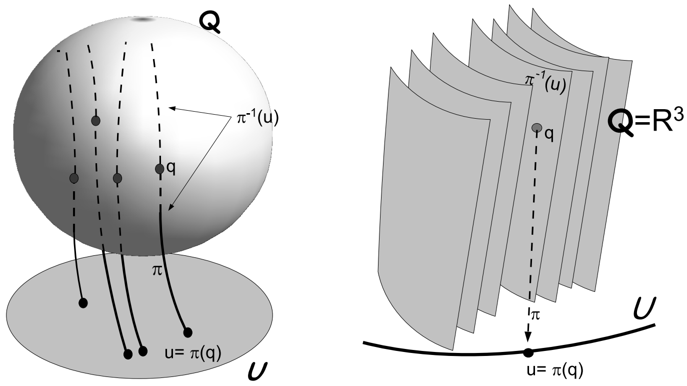

Figure 1.1: Surjective submersions, in 2- and 1-dimensional control spaces.

elementary Mechanics, the problem is not well-posed unless we specify the set of reaction forces one utilizes to implement u: we simply assume that this reaction forces are *orthogonal* (with respect to the Riemannian metric induced by the inertial tensor) to the submanifolds $\pi(q) = u = cost$ (the so-called *frozen constraint*) . This is, in fact, nothing but the d'Alembert hypothesis.

Within this second approach, a number of classical control problems can be investigated.

For instance one can study *stabilizability* or *optimal control* problems. Actually, these notes are mainly concerned with stabilizability, and, more specifically, with *vibrational stabilizability.* However we will touch an important aspect connected with optimization in the chapter dealing with impulsive control systems (see Chapter 4).

1.2 Vibrational stabilizability

By "vibrational stabilizability" we shall mean the possibility of stabilizing the system at a given state $\bar{\mathbf{q}}$ by means of *some control* $\mathbf{u}(\cdot)$ *that oscillates rapidly around* $\bar{\mathbf{u}} = \pi(\bar{\mathbf{q}})$.

A well known example where stability is obtained by oscillation of a parameter is provided by a pendulum whose suspension point can oscillate on a vertical guide, as in Figure 1.2. In this case $\mathcal{Q} = S^1 \times I, \mathcal{U} = I$, where S^1 is the circle and is an open interval, and π is simply the projection on the second factor I. Calling q^1 the angle and q^2 the height of the pivot, one has $u = q^2$. If we take $\bar{q}^1 = 0$ as the (unstable) upper vertical position of the pendulum, it is well-known (see for example [1, 25, 26] and references therein) that this configuration can be made stable by rapidly moving the pivot up and down a fixed value $\bar{\mathbf{u}}$. (This is commonly refereed as the "Kapitza pendulum").

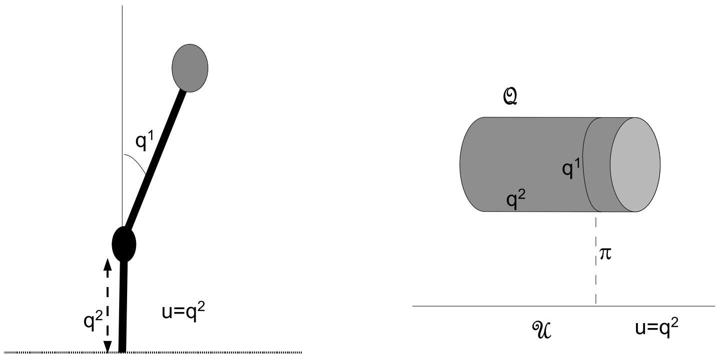

Figure 1.2: A pendulum with vertically moving pivot.

More generally, we will see that this system can be asymptotically stabilized at any angle $\bar{q}^1$ with $-\pi/2 < \bar{q}^1 < \pi/2$, by a suitable choice of an oscillating control function.

On the other hand, consider the variable length pendulum, where the pivot is fixed at the origin, but the radius of oscillation r can be assigned as a function of time, see Figure 1.2. The system is again described by two coordinates (q^1, q^2), the angle and the length. However, in this case, the upright equilibrium position is *not* stabilizable by any oscillatory motion of the *control* $u = q^2(t)$ (the radius) around a fixed value $\bar{q}^2$.

The crucial difference between the above systems is that the equation of motion of the first one contains a *quadratic* term in the time derivative $\dot{\mathbf{u}} \doteq d\mathbf{u}/dt$, while the equation for the variable-length pendulum is *affine* w.r.t. the variable $\dot{\mathbf{u}}$. An akin case where *vibrational stabilizability is not achievable* occurs if, instead of the length of the pendulum, the control

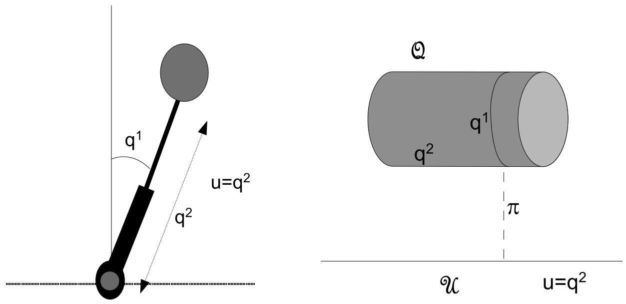

Figure 1.3: Length as control.

u represents a second pendulum attached at the free end of the primary pendulum (see Figure 1.2).

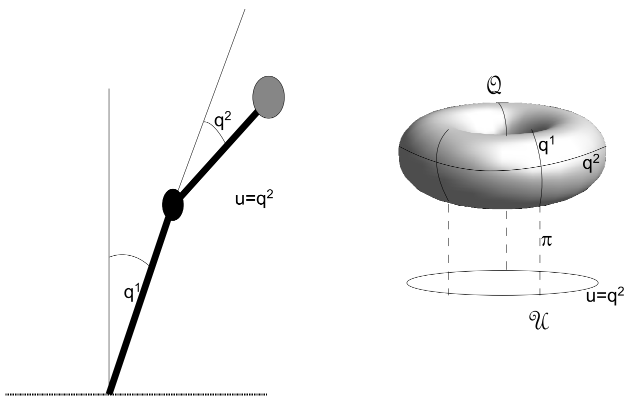

Figure 1.4: Second pendulum as control.

In order to understand the general problem, one has to consider two main issues:

- **The geometric issue.** It involves the *orthogonal curvature* of the foliation made of the fibers of the projection π, namely the family $\Lambda \doteq \left\{\pi^{-1}(\mathbf{q}) \quad \mathbf{q} \in \mathcal{Q}\right\}$. *Orthogonality* is here meant with respect to the Riemannian metric associated with the kinetic energy. The orthogonal curvature is a measure of how a geodesic, which is orthogonal to the leaf $\pi^{-1}(\mathbf{q})$ at a given point $\mathbf{q}$, fails to remain perpendicular to the other leaves it meets. If this curvature is non-zero, then the resulting control equations (which are second order for $\mathbf{q}$, or, equivalently, first order for the corresponding Hamiltonian-type system) contain a quadratic term in the time derivative $\dot{\mathbf{u}}$ of the control function. This will be analyzed in detail in Chapter 4.

The above geometrical considerations are valid when the original system on $\mathcal{Q}$ is *holonomic*, i.e. it derives from a Newtonian system subject to space (*ideal*) constraints. However, if also *non holonomic constraints*[3] act on the original system, the relation between quadratic dependence and geometry is much more involved, as it will be illustrated in Chapter 5.

- **The analytical issue.** In brief, the question consists in how to exploit the quadratic terms in $\dot{\mathbf{u}}$, in order to achieve stabilization. In particular, we shall study the set of solutions for a system with quadratic, unbounded, controls, making essential use of reparameterization techniques. These, in turn, will be combined with arguments involving local controllability or Lyapunov functions for the convexification of the reparameterized system.

  Incidentally, a chapter (Chapter 3) will be devoted to the particular case when the quadratic term is zero (which corresponds, on the mechanical side, to the vanishing of the orthogonal curvature). Actually, this subject is more crucial in optimization than in stabilization, even though it provides a case where *vibrational* stabilization is not attainable.

1.3 These notes' organization

This notes are organized as follows: There are two Parts, the former dealing with control systems depending on the (unbounded) derivative of the controls, the latter concerning applications to mechanical systems. In particular

[3] i.e. constraints on the velocities which cannot be *integrated*, namely obtained from holonomic constraints by diffrentiation.

Part 1 consists of Chapter 2 and Chapter 3. In Chapter 2 one investigates the case when the derivative of the control appear quadratically in the equations. The closure of solutions' set is studied together with stabilizability issues. Chapter 3 concerns the particular case when the dependence on the derivative of the control is affine. The question of the closure of the solutions' set is briefly mentioned, together with its strict connection with the interaction between *impulses* —namely, discontinuities of the control— and Lie brackets of the involved vector fields.

Part 2 is made of Chapters 4-6. Chapter 4 is devoted to the dynamics of holonomic systems driven by active constraints. In particular the control equations are presented. Moreover, some sections are devoted to curvature-like aspects and their close relation with the functional dependence of the equations on the control's derivative. In Chapter 5 the more involved dynamics of non holonomic systems is investigated. Equation in coordinates are deduced. In addition, intrinsic interpretation of the quadratic dependence on the control's derivative extend the results found for non-holonomic systems. In particular, besides the curvature-term already present in the holonomic case, one sees that the "lack of holonomy" brings in the equation a new quadratic term. This is essential in many issues, e.g. in vibrational stabilization, as it is also illustrated by an example.

Two Appendices conclude these notes. The former is a quite basic and rapid exposition of fundamental notions on differential manifolds. The latter, which is not directly connected with the other parts of the notes, consists of a medley of elementary considerations on the invariant structure of Lagrangian and Hamiltonian equations.

Part I

Nonlinear Systems with Unbounded Controls

Chapter 2

Quadratic control systems

We investigate general control systems of the form:

$$\dot{x} = f(x) + \sum_{\alpha=1}^{m} g_\alpha(x)\,\dot{u}^\alpha + \sum_{\alpha,\beta=1}^{m} h_{\alpha,\beta}(x)\,\dot{u}^\alpha \dot{u}^\beta\,. \tag{2.1}$$

The state variable x and the control variable u take values in $I\!R^n$ and in $I\!R^m$, respectively. We remark that no a priori bounds are imposed on the derivative $\dot{u}$. The important degenerate case where all the $h_{\alpha,\beta}$ vanish identically –namely, the affine control case– will be discussed in Chapter 3 .

Remark 2.0.1 To avoid confusion with other issues that go under the same name in literature, let us point out that

1. The vector fields g, h (2.1) *are not assumed to be constant.* In particular, the unboundedness of $\dot{u}$ interferes with the nonlinearity of these fields;

2. The controls $v \doteq \dot{u}$ are point-wise *unbounded* . (Actually, this is true also in the standard case of quadratic systems).

Our main goal is to understand under which conditions the system can be *stabilized to a given point* $\bar{x}$. In particular, relying on the quadratic dependence on $\dot{u}$ of the right-hand side of (2.1), in Section 2.3 we shall investigate what can be called *vibrational stabilization*, that is a stabilization achieved by means of small rapid oscillations of the control function. In Chapter 6 these results will be applied to the stabilization of the mechanical systems.

We assume that the functions f, g_α, and $h_{\alpha,\beta} = h_{\beta,\alpha}$ are at least twice continuously differentiable. We remark that the more general system

$$\dot{x} = \tilde{f}(t,x,u) + \sum_{\alpha=1}^{m} \tilde{g}_\alpha(t,x,u)\,\dot{u}^\alpha + \sum_{\alpha,\beta=1}^{m} \tilde{h}_{\alpha,\beta}(t,x,u)\,\dot{u}^\alpha \dot{u}^\beta\,,$$

where the vector fields depend also on time and on the control u, can be easily rewritten in the form (2.1). Indeed, it suffices to work in the extended state space $x \in I\!R^{1+n+m}$, introducing the additional state variables $x^0 = t$ and $x_{n+\alpha} = u^\alpha$, with equations

$$\dot{x}^0 = 1\,, \qquad \dot{x}_{n+\alpha} = \dot{u}^\alpha \qquad \alpha = 1,\dots,m\,.$$

2.1 Quadratic control systems

Given the initial condition

$$x(0) = \check{x}\,, \tag{2.2}$$

for every smooth control function $u : [0,T] \mapsto I\!R^m$ one obtains a unique solution $t \mapsto x(t,\,u)$ of the Cauchy problem (2.1)-(2.2). More generally, since the equation (2.1) is quadratic w.r.t. the derivative $\dot{u}$, it is natural to consider admissible controls in a set of absolutely continuous functions $u(\cdot)$ with derivatives in $I\!L^2$. For example, for a given $K > 0$, one could allow the controls to belong to

$$\left\{ u : [0,T] \mapsto I\!R^m\,; \qquad \int_0^T \left|\dot{u}(t)\right|^2 dt \leq K \right\}. \tag{2.3}$$

Since our aim is stabilization, it is conceivable to investigate the limits of this set of trajectories. In fact, the main goal of the following analysis is to provide a characterization of the closure (in appropriate topologies) of this set of trajectories. This will be achieved in terms of an auxiliary differential inclusion.

It is expectable that three main factors interplay in this program:

(I) the (pointwise) unboundedness of the controls derivatives $\dot{u}$;

(II) the quadratic dependence on $\dot{u}$;

(III) the usual chattering phenomena.

Let us comment on these factors:

(I) Notice that the pointwise unboundedness of $\dot{u}$ cannot be approached with mere measure-theoretical tools. This has been widely recognized in the case of affine dependence (namely $h_{\alpha,\beta} \equiv 0$) —see Chapter 3 — and has a basic-theoretical explication in the fact that the vector fields g_α and $h_{\alpha,\beta}$ are *not constant*. Incidentally, we remark that the fact of g_α and $h_{\alpha,\beta}$ being *not constant*, despite appearance, *is an intrinsic property*, easily expressed by the condition that all Lie brackets are equal to zero (See Theorem A.4.2 and the following comments). One can say that unboundedness of the controls makes the *non-commutativity* of the vector fields crucial in the very definition of solution.

(II) If the system where affine in $\dot{u}$ we would allow a wider class of controls then the one in (2.3). An instance is given by a family of controls whose derivatives are uniformly L^1-bounded (see Chapter 3). In the general case, it is crucial that for rapidly one dimensional oscillations of the controls u (smaller and smaller in the C^0 norm) the linear term is practically negligible so the dynamics is asymptotically governed only by the quadratic term.

(III) The chattering phenomena are already present in the case of L^∞-bounded controls, so there is no surprise in finding them in this more general situation. In particular, once reduced the system to a bounded-control system (via suitable reparameterization) the limits of trajectories are represented as trajectories of the convexified dynamics.

2.1.1 A graph-differential inclusion

Let us notice that the system (2.1) is naturally connected with the differential inclusion

$$\dot{x} \in \mathcal{F}(x), \tag{2.4}$$

where, for every $x \in I\!R^n$,

$$\mathcal{F}(x) \doteq \overline{co} \left\{ f(x) + \sum_{\alpha=1}^{m} g_\alpha(x) w^\alpha + \sum_{\alpha,\beta=1}^{m} h_{\alpha,\beta}(x) w^\alpha w^\beta \,; \quad (w^1, \dots, w^m) \in I\!R^m \right\}. \tag{2.5}$$

Here and in the sequel, for any given subset A of a topological vector space, $\overline{co}A$ denotes the closed convex hull of A.

In addition, it will be convenient to work also in an extended state space, using the variable $\hat{x} = \begin{pmatrix} x^0 \\ x \end{pmatrix} \in I\!R^{1+n}$, where x^0 represents time. For every $\hat{x}$, consider the set

$$F(\hat{x}) \doteq \overline{co} \Bigg\{ \begin{pmatrix} 1 \\ f(x) \end{pmatrix} (a^0)^2 + \textstyle\sum_{\alpha=1}^m \begin{pmatrix} 0 \\ g_\alpha(x) \end{pmatrix} a^0 a^\alpha + \\ + \textstyle\sum_{\alpha,\beta=1}^m \begin{pmatrix} 0 \\ h_{\alpha,\beta}(x) \end{pmatrix} a^\alpha a^\beta \,; \quad a^0 \in [0,1], \quad \textstyle\sum_{\alpha=0}^m (a^\alpha)^2 = 1 \Bigg\}. \tag{2.6}$$

Notice that F is a convex, compact valued multifunction on $I\!R^{1+n}$, Lipschitz continuous w.r.t. the Hausdorff metric [2]. (Instead, $\mathcal{F}$ is not bounded).

For a given interval $[0, S]$, the set of trajectories of the *graph differential inclusion*

$$\frac{d}{ds}\hat{x}(s) \in F(\hat{x}(s)) \,, \qquad \hat{x}(0) = \begin{pmatrix} 0 \\ x^\sharp \end{pmatrix} \tag{2.7}$$

is a non-empty, closed, bounded subset of $\mathcal{C}\big([0,S]\,;\ I\!R^{1+n}\big)$. Consider one particular solution, say $s \mapsto \hat{x}(s) = \begin{pmatrix} x^0(s) \\ x(s) \end{pmatrix}$, defined for $s \in [0, S]$. Assume that $T \doteq x^0(S) > 0$. Since the map $s \mapsto x^0(s)$ is non-decreasing, it admits a generalized inverse

$$s = s(t) \qquad \text{iff} \qquad x^0(s) = t \,. \tag{2.8}$$

Indeed, for all but countably many times $t \in [0, T]$ there exists a unique value of the parameter s such that the identity on the right of (2.8) holds. We can thus define a corresponding trajectory

$$t \mapsto x(t) = x\big(s(t)\big) \in I\!R^n. \tag{2.9}$$

This map is well defined for almost all times $t \in [0, T]$.

2.1.2 L^2-reparameterizazion

To establish a connection between the original control system (2.1) and the differential inclusion (2.7), consider first a smooth control function $u(\cdot)$. Let us define a reparameterized time variable by setting[1]

$$s(t) \doteq \int_0^t \Big(1 + \sum_{\alpha=1}^m (\dot{u}^\alpha)^2(\tau)\Big)\, d\tau \,. \tag{2.10}$$

[1]See [46] for a more general version of reparameterization including the polynomial dependence on $\dot{u}$.

Notice that the map $t \mapsto s(t)$ is strictly increasing. The inverse map $s \mapsto t(s)$ is uniformly Lipschitz continuous and satisfies

$$\frac{dt}{ds} = \left(1 + \sum_{\alpha=1}^{m} (\dot{u}^\alpha)^2(t)\right)^{-1}.$$

Let now $x : [0,T] \mapsto I\!R^n$ be a solution of (2.1) corresponding to the smooth control $u : [0,T] \mapsto I\!R^m$. We claim that the map $s \mapsto \hat{x}(s) \doteq \begin{pmatrix} t(s) \\ x(t(s)) \end{pmatrix}$ is a solution to the differential inclusion (2.7). Indeed, setting

$$a^0(s) \doteq \frac{1}{\sqrt{1 + \sum_{\beta=1}^m (\dot{u}^\beta)^2\big(t(s)\big)}}\,, \qquad a^\alpha(s) \doteq \frac{\dot{u}^\alpha\big(t(s)\big)}{\sqrt{1 + \sum_{\beta=1}^m (\dot{u}^\beta)^2\big(t(s)\big)}}\,, \tag{2.11}$$

$\alpha = 1, \ldots, m$, one has

$$\begin{cases} \frac{dt}{ds} = (a^0)^2(s) \\ \frac{dx}{ds} = f\Big(x(s)\Big)\,(a^0)^2(s) + \sum_{\alpha=1}^m g_\alpha\Big(x(s)\Big)\,a^0(s)a^\alpha(s) + \\ \qquad + \sum_{\alpha,\beta=1}^m h_{\alpha,\beta}\Big(x(s)\Big)\,a^\alpha(s)a^\beta(s)\,. \end{cases} \tag{2.12}$$

Hence $\hat{x}(\cdot) = (t(\cdot), x(\cdot))$ verifies (2.7), because, by (2.11),

$$a^0(s) \in [0,1]\,, \qquad \sum_{\alpha=0}^m (a^\alpha)^2(s) \equiv 1\,.$$

Notice that the derivatives $\dot{u}^\alpha$ can now be recovered as

$$\dot{u}^\alpha(t) = \frac{a^\alpha(s(t))}{a^0(s(t))} \qquad \alpha = 1, \ldots, m\,. \tag{2.13}$$

The following theorem shows that every solution of the differential inclusion (2.7) can be approximated by smooth solutions of the original control system (2.1).

Theorem 2.1.1 *Let $\hat{x} = (x^0, x) : [0,S] \mapsto I\!R^{1+n}$ be a solution to the multivalued Cauchy problem (2.7) such that $x^0(S) = T > 0$. Then there exists a sequence of smooth control functions $u_\nu : [0,T] \mapsto I\!R^M$ such that the corresponding solutions*

$$s \mapsto \hat{x}_\nu(s) = \begin{pmatrix} t_\nu(s) \\ x_\nu(s) \end{pmatrix}$$

of the equations (2.11)-(2.12) *converge to the map* $s \mapsto \hat{x}(s)$ *uniformly on* $[0, S]$. *Moreover, defining the function* $x(t) = x(s(t))$ *as in* (2.9), *we have*

$$\lim_{\nu\to\infty} \int_0^T \big|x(t) - x_\nu(t)\big|\, dt = 0\,. \tag{2.14}$$

Proof. By the assumption, the extended vector fields

$$\hat{f} = \begin{pmatrix} 1 \\ f \end{pmatrix}, \qquad \hat{g}_\alpha = \begin{pmatrix} 0 \\ g_\alpha \end{pmatrix}, \qquad \hat{h}_{\alpha,\beta} = \begin{pmatrix} 0 \\ h_{\alpha,\beta} \end{pmatrix}$$

are Lipschitz continuous. Consider the set of trajectories of the control system

$$\frac{d}{ds}\,\hat{x} = \hat{f}\cdot(a^0)^2 + \sum_{\alpha=1}^m \hat{g}_\alpha\, a^0 a^\alpha + \sum_{\alpha,\beta=1}^m \hat{h}_{\alpha,\beta}\, a^\alpha a^\beta\,, \qquad \hat{x}(0) = \begin{pmatrix} 0 \\ x^\sharp \end{pmatrix}, \tag{2.15}$$

where the controls $a = (a^0, a^1, \ldots, a^m)$ satisfy the pointwise constraints

$$a^0(s) \in [0,1]\,, \qquad \sum_{\alpha=0}^m (a^\alpha)^2(s) = 1 \qquad s \in [0, S]\,. \tag{2.16}$$

In the above setting, it is well known [2] that the set of trajectories

$$s \mapsto \hat{x}(s) = (x^0, x^1, \ldots, x^n)(s)$$

of (2.15)-(2.16) is dense on the set of solutions to the differential inclusion (2.7). Hence there exists a sequence of control functions $s \mapsto a_\nu(s) = \big(a^0_\nu, \ldots, a^m_\nu\big)(s)$, $\nu \geq 1$, such that the corresponding solutions $s \mapsto \hat{x}_\nu(s)$ of (2.15) converge to $\hat{x}(\cdot)$ uniformly for $s \in [0, S]$. In particular, this implies the convergence of the first components:

$$x^0{}_\nu(S) \;=\; \int_0^S \big[a^0_\nu(s)\big]^2\, ds \;\to\; x^0(S) \;=\; T\,. \tag{2.17}$$

We now observe that the "input-output map" $a(\cdot) \mapsto \hat{x}(\cdot, a)$ from controls to trajectories is uniformly continuous as a map from $\mathbf{L}^1\big([0, S]\,;\, \mathbb{R}^{1+m}\big)$ into $\mathcal{C}\big([0, S]\,;\, \mathbb{R}^{1+n}\big)$. By slightly modifying the controls a_ν in $\mathbf{L}^1$, we can replace the sequence a_ν by a new sequence of smooth control functions $\tilde{a}_\nu : [0, S] \mapsto \mathbb{R}^{1+m}$ with the following properties:

$$\tilde{a}^0_\nu(s) > 0 \qquad \text{for all } \; s \in [0, S]\,, \quad \nu \geq 1\,. \tag{2.18}$$

$$\int_0^S \big[\tilde a^0_\nu(s)\big]^2\,ds = T \qquad\qquad \text{for all } \nu\geq 1\,, \tag{2.19}$$

$$\lim_{\nu\to\infty}\int_0^S \big|\tilde a_\nu(s)-a_\nu(s)\big|\,ds \;=\; 0\,. \tag{2.20}$$

This implies the uniform convergence

$$\lim_{\nu\to\infty}\big\|\hat x(\cdot,\tilde a_\nu)-\hat x(\cdot)\big\|_{\mathcal{C}([0,S];\,I\!R^{1+n})} \;=\; 0\,. \tag{2.21}$$

By (2.18), for each $\nu\geq 1$ the map

$$s\mapsto x^0_\nu(s)\doteq\int_0^s \big[\tilde a^0_\nu(s)\big]^2\,ds$$

is strictly increasing. Therefore it has a smooth inverse $s=s_\nu(t)$. Recalling (2.13), we now define the sequence of smooth control functions $u_\nu:[0,T]\mapsto I\!R^m$ by setting $u_\nu(t)=\big(u^1_\nu,\ldots,u^m_\nu\big)(t)$, with

$$u^\alpha_\nu(t)=\int_0^t \frac{\tilde a^\alpha_\nu(s_\nu(\tau))}{\tilde a^0_\nu(s_\nu(\tau))}\;d\tau\,. \tag{2.22}$$

By construction, the solutions $t\mapsto x_\nu(t\,,\,u_\nu)$ of the original system (2.1) corresponding to the controls u_ν coincide with the trajectories $t\mapsto(x^1_\nu,\ldots,x^n_\nu)(s_\nu(t))$, where $\hat x_\nu=(x^0_\nu,x^1_\nu,\ldots,x^n_\nu)$ is the solution of (2.15) with control $\tilde a_\nu=(\tilde a^0_\nu,\ldots,\tilde a^m_\nu)$.

To prove the last statement in the theorem, define the increasing functions

$$t(s)=\int_0^s\big[\tilde a^0(r)\big]^2\,dr\,,\qquad\qquad t_\nu(s)=\int_0^s\big[\tilde a^0_\nu(r)\big]^2\,dr\,,$$

and let $t\mapsto s(t)$, $t\mapsto s_\nu(t)$ be their inverses, respectively. Notice that each $s_\nu(\cdot)$ is smooth. Moreover,

$$\left|\frac{d}{ds}t(s)\right|\leq 1\,,\qquad\qquad \left|\frac{d}{ds}t_\nu(s)\right|\leq 1\,, \tag{2.23}$$

$$\lim_{\nu\to\infty}\int_0^T\big|s(t)-s_\nu(t)\big|\,dt \;=\; \lim_{\nu\to\infty}\int_0^S\big|t(s)-t_\nu(s)\big|\,ds=0\,. \tag{2.24}$$

Using (2.23), we obtain the estimate

$$\begin{aligned}\int_0^T\big|x(t)-x_\nu(t)\big|\,dt \;&=\; \int_0^T\Big|x(s(t))-x_\nu(s(t))\Big|\,dt+\int_0^T\Big|x_\nu(s(t))-x_\nu(s_\nu(t))\Big|\,dt\\ &\leq\; \int_0^S\big|x(s)-x_\nu(s)\big|\,ds+C\cdot\int_0^T\big|s(t)-s_\nu(t)\big|\,dt\,.\end{aligned} \tag{2.25}$$

Here the constant C denotes an upper bound for the derivative w.r.t. s, for example

$$C \doteq \sup_x \left\{ \big|f(x)\big| + \sum_i \big|g_\alpha(x)\big| + \sum_{\alpha,\beta} \big|h_{\alpha,\beta}(x)\big| \right\}, \tag{2.26}$$

where the supremum is taken over a compact set containing the graphs of all functions $x_\nu(\cdot)$. By (2.21) and (2.24), the right hand side of (2.25) vanishes in the limit $\nu \to \infty$. This completes the proof of the theorem. ◊

Remark 2.1.2 For a given time interval $[0, T]$, we are considering controls $u(\cdot)$ in the Sobolev space $W^{1,2}$. The corresponding solutions are absolutely continuous maps, namely they belong to $W^{1,1}$. Now consider a sequence of control functions u_ν, whose derivatives are uniformly bounded in L^2. Assume that the corresponding reparameterized trajectories $s \mapsto (t_\nu(s),\, x_\nu(s))$, constructed as in (2.11)-(2.12), converge to a path $s \mapsto (t(s),\, x(s))$, providing a solution to (2.7). We wish to point out that, in general, the projection on the state space $t \mapsto x(s(t))$ *may well be discontinuous.* Notice that, on the contrary, the uniform limit of the controls $t \mapsto u_\nu(t)$ must be Hölder continuous, because of the uniform $\mathbf{L}^2$ bound on the derivatives. A completely different situation arises when all the vector fields $h_{\alpha,\beta}$ vanish identically, so that (2.1) reduces to

$$\dot x = f(x) + \sum_{\alpha=1}^{m} g_\alpha(x)\, \dot u^\alpha \tag{2.27}$$

This case will be treated in Chapter 3.

2.2 Stabilization

In this section we examine various concepts of stability for the impulsive system (2.1) and relate them to the weak stability of the differential inclusion (2.6)-(2.7).

Definition 2.2.1 *We say that the control system (2.1) is* stabilizable *at the point $\bar x \in I\!R^n$ if, for every $\varepsilon > 0$ there exists $\delta > 0$ such that the following holds. For every initial state $x^\sharp$ with $|x^\sharp - \bar x| \le \delta$ there exists a smooth control function $t \mapsto u(t) = (u^1, \ldots, u^m)(t)$ such that the corresponding trajectory of (2.1)-(2.2) satisfies*

$$|x(t,u) - \bar x| \le \varepsilon \qquad \forall t \ge 0\,. \tag{2.28}$$

We say that the system (2.1) is asymptotically stabilizable *at the point* $\bar{x}$ *if a control* $u(\cdot)$ *can be found such that, in addition to (2.28), there holds*

$$\lim_{t\to\infty} x(t,u) = \bar{x}\,. \tag{2.29}$$

Remark 2.2.2 Notice that the point $\bar{x}$ needs not to be an equilibrium point for the vector field f.

Remark 2.2.3 We require here that the stabilizing controls be smooth. As it will become apparent in the sequel, this is hardly a restriction. Indeed, in all cases under consideration, if a stabilizing control $u \in W^{1,2}$ is found, by approximation one can construct a smooth control $\tilde{u}$ which is still stabilizing.

Remark 2.2.4 In the above definitions we are not putting any constraint on the control function $u : [0,\infty[\mapsto I\!R^m$. In principle, one may well have $|u(t)| \to \infty$ as $t \to \infty$. If one wishes to stabilize the system (2.1) and at the same time keep the control values within a small neighborhood of a given value $\bar{u}$, it suffices to consider the stabilization problem for an augmented system, adding the variables $x^{n+1}, \ldots, x^{n+m}$ together with the equations

$$\dot{x}^{n+\alpha} = \dot{u}^{\alpha} \qquad\qquad \alpha = 1,\ldots,m\,.$$

Similar stability concepts can be also defined for a differential inclusion

$$\dot{x} \in K(x)\,, \tag{2.30}$$

see for example [56]. We recall that a trajectory of (2.30) is an absolutely continuous function $t \mapsto x(t)$ which satisfies the differential inclusion at a.e. time t.

Definition 5.2. The point $\bar{x}$ is *weakly stable* for the differential inclusion (2.30) if, for every $\varepsilon > 0$ there exists $\delta > 0$ such that the following holds. For every initial state $x^\sharp$ with $|x^\sharp - \bar{x}| \le \delta$ there exists a trajectory $x(\cdot)$ of (2.30) such that

$$x(0) = x^\sharp\,, \qquad |x(t) - \bar{x}| \le \varepsilon \qquad \forall t \ge 0\,. \tag{2.31}$$

Moreover, $\bar{x}$ is *weakly asymptotically stable* if, there exists a trajectory which, in addition to (2.31), satisfies

$$\lim_{t\to\infty} x(t) = \bar{x}\,. \tag{2.32}$$

In connection with the multifunction F defined at (2.6), we consider a second multifunction $F^\diamondsuit$ obtained by projecting the sets $F(\hat{x}) \subset I\!R \times I\!R^n$ into the second factor $I\!R^n$. More precisely, we set

$$F^\diamondsuit(x) \doteq \overline{co}\Big\{ f(x)\,(a^0)^2 + \textstyle\sum_{\alpha=1}^m g_\alpha(x)\,a^0 a^\alpha + \sum_{\alpha,\beta=1}^m h_{\alpha,\beta}(x)\,a^\alpha a^\beta \;;$$
$$w^0 \in [0,1]\,, \quad \textstyle\sum_{\alpha=0}^m (w^\alpha)^2 = 1 \Big\}\,. \tag{2.33}$$

Observe that, if the vector fields f, g_α, and $h_{\alpha,\beta}$ are Lipschitz continuous, then the multifunction $F^\diamondsuit$ is Lipschitz continuous with compact, convex values. Our first result in this section is:

Theorem 2.2.5 *The impulsive system (2.1) is asymptotically stabilizable at the point $\bar{x}$ if and only if $\bar{x}$ is weakly asymptotically stable for the* projected graph differential inclusion

$$\frac{d}{ds}x(s) \in F^\diamondsuit(x(s))\,. \tag{2.34}$$

Proof. Let $\bar{x}$ be weakly asymptotically stable for (2.34). Without loss of generality, we can assume $\bar{x} = 0$.

Given $\varepsilon > 0$, choose $\delta > 0$ such that, if $|x^\sharp| \leq \delta$, then there exists a trajectory $t \mapsto x(s)$ of the differential inclusion (2.34) such that $x(0) = x^\sharp$, $|x(s)| \leq \varepsilon/2$ for all $t \geq 0$ and $x(s) \to 0$ as $t \to \infty$. Using the basic approximation property stated in Theorem 2.1.1, we will construct a smooth control $t \mapsto u(t) = (u^1, \ldots, u^m)(t)$ such that the corresponding trajectory $x(\cdot; u)$ of (2.1)-(2.2) satisfies

$$|x(t)| \leq \varepsilon \qquad \forall t \geq 0\,, \qquad\qquad \lim_{t\to\infty} x(t) = 0\,. \tag{2.35}$$

Define the decreasing sequence of positive numbers $\varepsilon_k \doteq \varepsilon\, 2^{-k}$. For each $k \geq 0$, choose $\delta_k > 0$ so that, whenever $|x^\sharp| \leq \delta_k$, there exists a solution to (2.34) with

$$x(0) = x^\sharp\,, \qquad \lim_{s\to\infty} x(s) = 0\,, \qquad |x(s)| < \frac{\varepsilon_k}{2} \quad \forall s \geq 0\,. \tag{2.36}$$

Choose a sequence of strictly positive integers $k(1) \leq k(2) \leq \cdots$, such that

$$\lim_{j\to\infty} k(j) = \infty\,, \qquad\qquad \sum_{j=1}^\infty \delta_{k(j)} = \infty\,. \tag{2.37}$$

Note that the second condition in (2.37) is certainly satisfied if the numbers $k(j)$ grow at a sufficiently slow rate.

Assume $|x^\sharp| \leq \delta_0$. A smooth control u steering the system (2.1) from $x^\sharp$ asymptotically toward the origin will be constructed by induction on j. For $j = 1$, let $x : [0, s_1] \mapsto I\!R^n$ be a trajectory of the differential inclusion (2.34) such that

$$x(0) = x^\sharp , \qquad |x(s_1)| < \frac{\delta_{k(1)}}{3} , \qquad |x(s)| < \frac{\varepsilon_0}{2} \quad \forall s \in [0, s_1] .$$

By the definition of $F^\diamondsuit$, there exists a trajectory of the differential inclusion (2.7) having the form $s \mapsto \hat{x}(s) = (x^0(s), x(s))$. Notice that, in order to apply Theorem 2.1.1 and approximate $x(\cdot)$ with a smooth solution of the control system (2.1) we would need $x^0(s_1) > 0$. This is not yet guaranteed by the above construction. To take care of this problem, we define $s_1' \doteq s_1 + \delta_{k(1)}/3C$, where C provides a local upper bound for the magnitude of the vector field f, as in (2.26). We then prolong the trajectory $\hat{x}(\cdot)$ to the larger interval $[0, s_1']$, by setting

$$\frac{d}{ds}\begin{pmatrix} x^0(s) \\ x(s) \end{pmatrix} = \begin{pmatrix} 1 \\ f(x) \end{pmatrix} \qquad\qquad s \in \,]s_1, \, s_1'] .$$

This construction achieves the inequalities

$$x^0(s_1') \geq s_1' - s_1 \geq \frac{\delta_{k(1)}}{3C} , \qquad\qquad |x(s_1')| < \frac{2}{3}\delta_{k(1)} .$$

Set $\tau^1 \doteq x^0(s_1')$. By Theorem 2.1.1, there exists a smooth control $u : [0, \tau^1] \mapsto I\!R^m$ such that the corresponding solution $s \mapsto (x^0(s, u), x(s, u))$ of (2.11)-(2.12) differs from the above trajectory by less than $\delta_{k(1)}/3$, namely

$$|x^0(s, u) - x^0(s)| < \frac{\delta_{k(1)}}{3} , \qquad |x(s, u) - x(s)| < \frac{\delta_{k(1)}}{3} \qquad \forall s \in [0, s_1'] .$$

In particular, setting $x(t, u) \doteq x(s(t), u)$ as in (2.9), this implies

$$|x(\tau_1, u)| < \delta_{k(1)} , \qquad\qquad |x(t, u)| < \frac{\varepsilon_0}{2} + \frac{\delta_{k(1)}}{3} \leq \varepsilon_0 \qquad \forall t \in [0, \tau_1] .$$

The construction now proceeds by induction on j. Assume that a smooth control $u(\cdot)$ has been constructed on the time interval $[0, \tau_j]$, in such a way that

$$|x(\tau_j, u)| < \delta_{k(j)} , \qquad\qquad |x(t, u)| < \varepsilon_{k(j-1)} \qquad \forall t \in [\tau_{j-1}, \, \tau_j] . \qquad (2.38)$$

By assumptions, there exists a trajectory $s \mapsto x(s)$ of the differential inclusion (2.34) such that

$$x(0) = x(\tau_j, u)\,, \qquad |x(s_j)| < \frac{\delta_{k(j+1)}}{3}\,, \qquad |x(s)| < \frac{\varepsilon_{k(j)}}{2} \quad \forall s \in [0, s_j]\,. \tag{2.39}$$

This trajectory is extended to the slightly larger interval $[0, s'_j]$, with $s'_j = s_j + \delta_{k(j)}/3C$, by setting

$$\frac{d}{ds}\begin{pmatrix} x^0(s) \\ x(s) \end{pmatrix} = \begin{pmatrix} 1 \\ f(x) \end{pmatrix} \qquad s \in]s_j,\, s'_j]\,. \tag{2.40}$$

Notice that, by (2.39), (2.40), and (2.26), we have

$$x^0(s'_j) \geq s'_j - s_j \geq \frac{\delta_{k(j)}}{3C}\,, \qquad |x(s'_j)| < \frac{2}{3}\delta_{k(j+1)}\,. \tag{2.41}$$

Set $\tau_{j+1} \doteq \tau_j + x^0(s'_j)$. Using again Theorem 2.1.1, we can extend the control $u : [0, \tau_j] \mapsto I\!R^m$ to a continuous, piecewise smooth control defined on the larger interval $[0, \tau_{j+1}]$, such that the corresponding solution $s \mapsto x(s, u)$ of (2.1)-(2.2) satisfies

$$|x(\tau_{j+1}, u)| < \delta_{k(j+1)}\,, \qquad |x(t, u)| < \varepsilon_{k(j)} \qquad \forall t \in [\tau_j,\, \tau_{j+1}]\,. \tag{2.42}$$

Notice that, at this stage, the control u is obtained by piecing together two smooth control functions, defined on the intervals $[0, \tau_j]$ and $[\tau_j, \tau_{j+1}]$ respectively. This makes u continuous but possibly not $\mathcal{C}^1$ in a neighborhood of the point τ_j. To fix this problem, we slightly modify the values of u in a small neighborhood of τ_j, so that u becomes smooth also at this point, while the strict inequalities (2.42) still hold.

Having completed the inductive steps for all $j \geq 1$ we observe that

$$\lim_{j\to\infty} \tau_j = \sum_j \frac{\delta_{k(j)}}{3C} = \infty$$

because of (2.37). As $t \to \infty$, by (2.42) we have $x(t, u) \to 0$. This shows that the impulsive system (2.1) is asymptotically stabilizable at the origin, proving one of the implications stated in the theorem.

The converse implication is obvious, because every solution of the system (2.1) corresponding to a smooth control yields a solution to the differential inclusion (2.34), after a suitable time rescaling.

Corollary 2.2.6 *Let a point $\bar{x}$ be weakly asymptotically stable for the differential inclusion (2.4), namely $\dot{x} \in \mathcal{F}(x)$. Then the system (2.1) is asymptotically stabilizable at $\bar{x}$.*

Proof. Since the point $\bar{x}$ is weakly asymptotically stable for (2.4), then it is asymptotically stable for the differential inclusion (2.34), which, in turn, implies that the impulsive system (2.1) can be stabilized at $\bar{x}$.

2.2.1 Lyapunov functions

There is an extensive literature, in the context of O.D.E's and of control systems or differential inclusions, relating the stability of an equilibrium state to the existence of a Lyapunov function. We recall below the basic definition, in a form suitable for our applications. For simplicity, we henceforth consider the case $\bar{x} = 0 \in I\!R^n$, which of course is not restrictive.

Definition 2.2.7 *A scalar function V defined on a neighborhood $\mathcal{N}$ of the origin is a* weak Lyapunov function *for the differential inclusion*

$$\dot{x} \in \mathcal{F}(x)$$

if the following holds.

(i) V is continuous on $\mathcal{N}$, and continuously differentiable on $\mathcal{N} \setminus \{0\}$.

(ii) $V(0) = 0$ while $V(x) > 0$ for all $x \neq 0$.

(iii) For each $\delta > 0$ sufficiently small, the sublevel set $\{x\,;\;\; V(x) \leq \delta\}$ is compact.

(iv) At each $x \neq 0$ one has

$$\inf_{y \in \mathcal{F}(x)} \nabla V(x) \cdot y \leq 0\,. \tag{2.43}$$

The following theorem relates the stability of the impulsive control system (2.1) to the existence of a Lyapunov function for the differential inclusion (2.4).

Theorem 2.2.8 *Consider the multifunction $\mathcal{F}$ defined at* (2.5)*. Assume that the differential inclusion* (2.4) *admits a Lyapunov function $V = V(x)$ defined on a neighborhood $\mathcal{N}$ of the origin. Then the control system* (2.1) *can be stabilized at the origin.*

Remark 2.2.9 Notice that the multifunction $\mathcal{F}$ in (2.5) has unbounded values. Yet we can rephrase condition (iv) in the definition 2.2.7 with the

following equivalent condition, which is formulated in terms of the bounded multifunction F governing (2.6):
(iv') For every $x \in \mathcal{N} \setminus \{0\}$, there exists $\hat{y} = (y_0, y) \in F(x)$ such that

$$\nabla V(x) \cdot y \leq 0 \qquad\qquad y_0 > 0\,. \tag{2.44}$$

Remark 2.2.10 The set of conditions (i)-(iii) and (iv') represents a slight strengthening of the notion of weak Lyapunov function when this is applied to the projected graph differential equation (2.34). Yet, let us point out that the weak stability of (2.34) is not enough to guarantee the stability of the control system (2.1), so the condition $y_0 > 0$ in (2.44) plays a crucial role. Indeed, on $I\!R^2$, consider the constant vector fields $f = (1,0)$, $h_{11} = (0,1)$, $h_{22} = (0,-1)$, $g_1 = g_2 = h_{12} = h_{21} = (0,0)$. Then, choosing $a^0 = 0$, $a^1 = a^2 = 1/\sqrt{2}$ we see that $(0,0,0) \in F(x)$ for every $x \in I\!R^2$. Hence condition

$$\inf_{y \in F(x)} \nabla V \cdot y \leq 0$$

is trivially satisfied by any function V. However, it is clear that in this case the system (2.1) is not stabilizable at the origin.

Remark 2.2.11 Theorem 2.29 is somewhat weaker than its counterpart, Theorem 2.2.5, dealing with asymptotic stability. Indeed, to prove that the impulsive control system (2.1) is stabilizable, we need to assume not only that the differential inclusion (2.34) is weakly stable, but also that there exists a Lyapunov function.

Proof of Theorem 2.2.8. Given $\varepsilon > 0$, choose $\delta > 0$ such that

$$V(x) \leq 2\delta \qquad \text{implies} \qquad |x| \leq \varepsilon.$$

Let an initial state $x^\sharp$ be given, with $V(x^\sharp) \leq \delta$.

According to Remark 2.2.9, for every $x \neq 0$ there exists $(y_0, y) \in F(x)$ such that (2.44) holds. We recall that the multifunction F in (2.6) is Lipschitz continuous, with compact, convex values. Since the set $\Omega \doteq \{x\,;\ \ \delta \leq V(x) \leq 3\delta\}$ is compact, by the continuity of ∇V we can find $\kappa > 0$ such that, for every $x \in \Omega$, there exists $\hat{y} = (y_0, y) \in F(x)$ with

$$\nabla V(x) \cdot y \leq 0\,, \qquad\qquad y_0 \geq \kappa.$$

The control u will be defined inductively on a sequence of the time intervals $[\tau_{j-1},\, \tau_j]$, with $\tau_j \geq j\kappa$. Set $\tau_0 = 0$. Consider the differential inclusion

$$\frac{d}{ds}\hat{x}(s) \in \begin{cases} F(x(s)) \cap \{(y_0, y)\,;\ \ \nabla V(x)\cdot y \leq 0\,,\ \ y_0 \geq \kappa\} & \text{if} \quad \delta < V(x) < 2\delta\,, \\ F(x(s)) \quad \text{if} \quad V(x) \leq \delta \quad \text{or} \quad V(x) \geq 2\delta\,, \end{cases} \tag{2.45}$$

with initial data $\hat{x}(0) = (0, x^\sharp)$. The right- hand side of (2.45) is an upper semicontinuous multifunction, with nonempty compact convex values. Therefore (see for example [2]), the Cauchy problem admits at least one solution $s \mapsto \hat{x}(s) = (x^0(s), x(s))$, defined for $s \in [0,1]$. We observe that this solution satisfies

$$x^0(1) \geq \kappa\,, \qquad\qquad V(x(s)) \leq \delta \qquad \forall s \in [0,1]\,.$$

Hence, by Theorem 2.1.1 there exists a smooth control $u : [0, \tau_1] \mapsto I\!R^m$, with $\tau_1 = x^0(1) \geq \kappa$, such that the corresponding trajectory of (2.1)-2.2) satisfies

$$V(x(t,u)) < \frac{3}{2}\delta = 2\delta - 2^{-1}\delta \qquad\qquad \forall t \in [0, \tau_1]\,.$$

By induction, assume now that a smooth control $u(\cdot)$ has been constructed on the interval $[0, \tau_j]$ with $\tau_j \geq \kappa\, j$, and that the corresponding trajectory $t \mapsto x(t,u)$ of the impulsive system (2.1)-(2.2) satisfies

$$V(x(t,u)) \leq 2\delta - 2^{-j}\delta \qquad\qquad t \in [0, \tau_j]\,. \tag{2.46}$$

We then construct a solution $s \mapsto \hat{x}(s) = (x^0(s),\, x(s))$ of the differential inclusion (2.45) for $s \in [0,1]$, with initial data $\hat{x}(0) = (0,\, x(\tau_j, u))$. This function will satisfy

$$x^0(1) \geq \kappa\,, \qquad\qquad V(x(s)) < 2\delta - 2^{-j}\delta \quad \forall s \in [0,1]\,.$$

Using again Theorem 2.1.1, we can prolong the control u to a larger time interval $[0,\, \tau_{j+1}]$, with $\tau_{j+1} - \tau_j = x^0(1) \geq \kappa$, in such a way that

$$V(x(t,u)) < 2\delta - 2^{-j-1}\delta \qquad\qquad t \in [0,\, \tau_{j+1}]\,. \tag{2.47}$$

At a first stage, this control u will be piecewise smooth, continuous but not $\mathcal{C}^1$ in a neighborhood of the point τ_j. By a local approximation, we can slightly change its values in a small neighborhood of the point τ_j, making it smooth also at the point τ_j, and preserving the strict inequalities (2.47).

Since $\tau_j \geq k\, j$ for all $j \geq 1$, as $j \to \infty$ the induction procedure generates a smooth control function $u(\cdot)$, defined for all $t \geq 0$, whose corresponding trajectory satisfies $V(x(t,u)) < 2\delta$ for all $t \geq 0$. This completes the proof of the theorem. $\diamondsuit$

Let us consider the 2-homogeneous term of $\mathcal{F}$:

$$\mathcal{F}_2 \doteq f(x) + \overline{co}\left\{\sum_{\alpha,\beta=1}^{m} h_{\alpha,\beta}(x)\, w^\alpha w^\beta\,; \quad (w^1,\ldots,w^m) \in I\!R^m\right\}$$

In Remark 2.3.2 one easily shows that

$$f(x) + \mathcal{F}_2 \subset \mathcal{F}.$$

Therefore, from Theorem 2.2.8 we obtain the following result.

Corollary 2.2.12 *Assume that the reduced differential inclusion*

$$\dot{x} \in f(x) + \mathcal{F}_2 \tag{2.48}$$

admits a Lyapunov function $V = V(x)$ defined on a neighborhood $\mathcal{N}$ of the origin. Then the control system (2.1) *can be stabilized at the origin.*

2.3 A selection technique

In the previous section we proved two general results, relating the stability of the control system (2.1) to the weak stability of the differential inclusion (2.4). A complete description of the sets $\mathcal{F}(x)$ in (2.5) may often be very difficult. However, as shown in [56], to establish a stability property it suffices to construct a suitable family of smooth selections. We shall briefly describe this approach.

Let a point $\bar{x} \in I\!R^n$ be given, and assume that there exists a $\mathcal{C}^1$ selection

$$\gamma(x,\xi) \in \mathcal{F}_1(x) \doteq \overline{co}\left\{\sum_{\alpha=1}^{m} g_\alpha(x)\, w^\alpha + \sum_{\alpha,\beta=1}^{m} h_{\alpha,\beta}(x)\, w^\alpha w^\beta\,; \quad (w^1,\ldots,w^m) \in I\!R^m\right\}$$

depending on an additional parameter $\xi \in I\!R^d$, such that

$$f(\bar{x}) + \gamma(\bar{x},\bar{\xi}) = 0\,. \tag{2.49}$$

for some $\bar{\xi} \in I\!R^d$. Assuming that γ is defined on an entire neighborhood of $(\bar{x},\bar{\xi})$, consider the Jacobian matrices of partial derivatives computed at $(\bar{x},\bar{\xi})$:

$$A \doteq \frac{\partial f}{\partial x} + \frac{\partial \gamma}{\partial x}\,, \qquad\qquad B \doteq \frac{\partial \gamma}{\partial \xi}\,.$$

Theorem 2.3.1 *In the above setting, if the linear system with constant coefficients*

$$\dot{x} = Ax + B\xi \tag{2.50}$$

is completely controllable, then the differential inclusion (2.4)-(2.5) is weakly asymptotically stable at the point $\bar{x}$.

We recall that the system (2.50) is completely controllable if and only if the matrices A, B satisfy the algebraic relation $\text{Rank}\big[B,\ AB,\ \dots\ ,\ A^{n-1}B\big] = n$. This guarantees that the system can be steered from any initial state to any final state, within any given time interval [8, 58].

To prove the theorem, consider the control system

$$\dot{x} = f(x) + \gamma(x,\xi). \tag{2.51}$$

By a classical result in control theory, the above assumptions imply that, for every point $x^\sharp$ sufficiently close to $\bar{x}$, there exists a trajectory starting from $x^\sharp$ reaching $\bar{x}$ in finite time. In particular, in view of (2.49), the system (2.51) is asymptotically stabilizable at the point $\bar{x}$. Since all trajectories of (2.51) are also trajectories of the differential inclusion (2.4), the result follows. ♢

Remark 2.3.2 Toward the construction of smooth selections from the multifunction $\mathcal{F}$ we observe that each closed convex set $\mathcal{F}(x)$ can be equivalently written as

$$\begin{aligned}\mathcal{F}(x) \ &\doteq\ f(x) + \mathcal{F}_1(x) + \mathcal{F}_2(x)\\ &= f(x) + \overline{co}\Big\{\textstyle\sum_{\alpha=1}^m g_\alpha(x)\, w^\alpha + \sum_{\alpha,\beta=1}^m h_{\alpha,\beta}(x)\, w^\alpha w^\beta\,; \quad (w^1,\dots,w^m)\in I\!R^m\Big\}\\ &\qquad + \overline{co}\left\{\sum_{\alpha,\beta=1}^m h_{\alpha,\beta}(x)\, w^\alpha w^\beta\,; \quad (w^1,\dots,w^m)\in I\!R^m\right\}\end{aligned} \tag{2.52}$$

Indeed, by definition we have $\mathcal{F}(x) = f(x) + \mathcal{F}_1(x)$. To establish the identity (2.52) it thus suffices to prove that

$$\mathcal{F}_1 + \mathcal{F}_2 \subseteq \mathcal{F}_1\,. \tag{2.53}$$

Since the set $\mathcal{F}_1(x)$ is convex and contains the origin, for every $(w^1,\dots,w^m)\in I\!R^m$ and $\varepsilon\in[0,1]$ we have

$$y_\varepsilon \ \doteq\ \varepsilon\left(\sum_{\alpha=1}^m g_\alpha(x)\,\frac{w^\alpha}{\sqrt{\varepsilon}} + \sum_{\alpha,\beta=1}^m h_{\alpha,\beta}(x)\frac{w^\alpha w^\beta}{\varepsilon}\right)\ \in\ \mathcal{F}_1\,.$$

Letting $\varepsilon \to 0$ we find

$$\lim_{\varepsilon\to 0+} y_\varepsilon \;=\; \sum_{\alpha,\beta=1}^{m} h_{\alpha,\beta}(x)\, w^\alpha w^\beta\,. \tag{2.54}$$

Since $\mathcal{F}_1(x)$ is closed, it must contain the right hand side of (2.54). This proves the inclusion $\mathcal{F}_2 \subseteq \mathcal{F}_1$. Next, observing that $\mathcal{F}_2$ is a cone, for every $y_2 \in \mathcal{F}_2$ and $\varepsilon > 0$ we have $\varepsilon^{-1} y_2 \in \mathcal{F}_2 \subseteq \mathcal{F}_1$. Therefore, if $y_1 \in \mathcal{F}_1$ we can write

$$y_1 + y_2 = \lim_{\varepsilon\to 0+} (1-\varepsilon)y_1 + \varepsilon(\varepsilon^{-1}y_2) \in \mathcal{F}_1$$

because $\mathcal{F}_1$ is closed and convex. This proves (2.53).

Remark 2.3.3 By Theorem 2.3.1 and the above remark, one may establish a stability result be constructing suitable selections $\gamma(x,\xi) \in \mathcal{F}_2(x)$ from the cone $\mathcal{F}_2$.

Chapter 3

Affine control systems

In this chapter we shall deal with systems affine in the control, namely the case when the quadratic coefficients $h_{\alpha,\beta}$ in the general control equation (2.1) vanish identically and reduces to

$$\dot{x} = f(x) + g_1(x)\dot{u}^1 + \cdots + g_m(x)\dot{u}^m. \tag{3.1}$$

It is clear that classes of controls larger than those of the general case can be considered for equation (3.1): certainly absolutely continuous controls are o.k., but it is also intuitive –even though not trivial– that even discontinuous controls might be allowed. There is a wide literature for systems like (3.1), and we give here just a small and non-exhaustive account of the existing results. Moreover, systems like (3.1) are more interesting in optimization with slow growth than in stabilization, even though they provide an interesting example where *vibrational stabilizability* is *not* achievable.

As we have mentioned, the interest in this extension of the ordinary notion of trajectory for equation (3.1) is motivated e.g. by optimal control problems *with slow growth*, where minimizing sequences possibly converge to discontinuous maps. We stress two main facts. First, the genuine nonlinear nature of the problem implies that a naive distributional approach does not work, even if the control is scalar valued. In other words the dynamical equation governing the motion cannot be interpreted as an equality between distributions. Secondly, as soon as the control is vector-valued, the noncommutativity of Lie brackets of vector fields makes the problem of determining a discontinuity of the trajectory much more involved. For instance, assume that the control [appears linearly in the dynamics and] consists of a [vector-valued] measure concentrated at time $\hat{t}$. It turns out that the mere knowledge of this measure is not sufficient for determining the corresponding jump of the trajectory. In fact, in order to compute this jump one needs a description of the path bridging (istantaneously at $t = \hat{t}$) the gap of the control's primitive. This leads to a notion of *space-time control*, which can be regarded as the limit of the graphs of primitives of ordinary –i.e. absolutely continuous– controls. Let us point out that two distinct space-time controls may have the same spatial projection. And, unless all involved Lie brackets vanish identically, two such space-time controls possibly generate distinct (space-time) trajectories and, hence, distinct costs. Let us remark that the set of [the graphs of] ordinary trajectories is dense in the set of space-time trajectories. Moreover, for a large class of minimum problems an optimal space-time trajectory does exist. Hence, the space-time embedding can be considered as a natural *extension* of the original problem with slow growth.Besides Mechanics, further applications can be found in mathematical modelling of optimal advertising [20].

There is a lot of important references moving around nonlinear impulsive problems. The following is a short, incomplete, list of these works: [20],[40], [53],[54],[55],[62],[63].

3.1 Introduction

While there is no difficulty to give a robust notion of solution to

$$\begin{cases} \dot{x} = f(x) + G\xi = f(x) + g_1\xi_1 + \cdots + g_m\xi_m \\ x(\bar{t}) = \bar{x} \end{cases} \qquad t \in [\bar{t}, T] \tag{3.2}$$

when G is a constant matrix (and $g_1, \ldots, g_m$ are its columns), ξ is a first order distribution –i.e., $\xi = \dot{u}, u \in L^1_{loc}$ – a *distributional approach* turns out to be not adequate as soon as the g_i are x-dependent:

$$\begin{cases} \dot{x} = f(x) + G(x)\xi = f(x) + g_1(x)\xi_1 + \cdots + g_m(x)\xi_m \\ x(\bar{t}) = \bar{x} \end{cases} . \tag{3.3}$$

In fact, we observe that x is a solution of (3.2) corresponding to a control $\xi \in L^1$ if and only if the map

$$z = x - G \cdot u, \tag{3.4}$$

with $u(t) \doteq \int_{\bar{t}}^{t} \xi(s)\,ds$, is a solution of

$$\begin{cases} \dot{z} = \hat{f}(z, u) \doteq f(z + Gu) \\ z(\bar{t}) = \bar{x}. \end{cases} \tag{3.5}$$

Notice that since ξ belongs to L^1, the map u belongs to AC, where AC denotes the set of absolutely continuous maps. Yet both (3.4) and (3.5) are meaningful for a control $u \in L^1$ as well. Moreover in view of the linearity of (3.4) and of wellknown properties of the input-output map of (3.5) we may *pass to the limit* when a sequence of controls $u_n \in AC$ converges (e.g., in the L^1-norm) to a control $u \in L^1$. Hence, given a map $u \in L^1$, one can define *the* solution x of (3.2) corresponding to (the distribution) $\xi = \dot{u}$ as

$$x \doteq z + Gu,$$

where z is the solution to (3.5).

In other words the *input-output map* $\Phi : AC \to L^1$, which to a control u associates the solution $x = \Phi(u)$ of (3.2) corresponding to $\xi = \dot{u}$ [is continuous when AC is endowed with the L^1-topology and] can be *continuously extended* to a map $\tilde{\Phi} : L^1 \to L^1$. That is, the continuous map $\tilde{\Phi}$ renders the diagram commutative, where e denotes the (dense) embedding of AC

in L^1. And the above given notion of solution x corresponding to a control $u \in L^1$ is nothing but $\tilde{\Phi}(u)$. Notice that the solution $\tilde{\Phi}(u)$ is defined up to a set of null measure. (Yet, a notion of solution pointwise defined of a subset $I \subset [\bar{t}, T]$ can be trivially given by considering sequences u_n converging to u in L^1 and pointwise in I (see [6] and [9])).

We remark that as soon as u has bounded variation – which implies that $\xi = \dot{u}$ is a Radon-measure – the solution $x = \tilde{\Phi}(u)$ is a solution "in measure" of (3.2), that is, (3.2) is verified as an identity between the measure $\dot{x}$ and the measure on the right-hand side. Summing up the above considerations we can say that for equations like (3.2) no difficulty arises when one wishes to extend the notion of solution in order to include the case where the control ξ is the (distributional) derivative of a L^1 function (see also [52] for the case when G depend on t). This can be done by continuous extension of the input-output map. And, as soon as u has bounded variation, this is equivalent to give a notion of solution "in measure".

However, there are essentially two crucial drawbacks in the attempt of extending the arguments of the linear case to the nonlinear equation (3.3):

(i) for $m = 1$, i.e., when (3.3) reduces to

$$\begin{cases} \dot{x} = f(x) + g_1(x)\xi \\ x(\bar{t}) = \bar{x} \end{cases} \tag{3.6}$$

the approach based on the continuous extension of the input-output map still works. On the contrary, the approach "in measure" lacks natural features of well-posedness (see below);

(ii) the "continuous extension" approach, valid for the scalar control-case (i.e., $m = 1$), does not hold any longer when $m \geq 2$.

While the problem described in (i) is a genuine analytical question (relying essentially in the fact that a naive distributional approach does not work for nonlinear systems), the difficulty pointed out in (ii) is directly related to a more differential geometric issue, namely the fact that (in general)

$$[g_i, g_j](x) \neq 0,$$

where $[g_i, g_j]$ is the Lie bracket of g_i and g_j:

$$[g_i, g_j] = \nabla g_j g_i - \nabla g_i g_j.$$

Actually, when

$$[g_i, g_j](x) = 0 \quad \forall x \in \mathbb{R}^n, \quad i, j = 1, \dots, m, \tag{3.7}$$

the "extension approach" *does work* [9], just as in the case when $m = 1$ (which trivially verifies the commutativity hypothesis (3.6)).

3.2 The scalar case and the commutative case

As mentioned above, when $m = 1$, we can address the problem of giving a robust notion of solution to (3.3) by exploiting the same extension argument which turned out to be successful in the linear case. At various levels of generality this approach has been pursued in [59], [6] (see also [50],[21],[51] and the references therein). More precisely, under standard regularity assumptions one proves that the input-output map of (3.6)

$$\Phi : AC \to L^1$$

–where $\Phi(u)$ is the classical (i.e., Caratheodory) solution of (3.6) corresponding to $u(t) = \int_{\bar{t}}^{t} \xi(s)\, ds$ – is uniformly continuous when AC is endowed with the L^1 topology: hence Φ can be continuously extended to (3.6).

Therefore, just as in the linear case, there is a continuous mapping $\tilde{\Phi}$ which renders the diagram commutative. The reason why this happens relies

on the fact that (3.6) can be transformed into a linear system. Indeed, let us consider the system

$$\begin{cases} \dot{\alpha} = \xi \\ \dot{x} = f(x) + g_1(x)\xi \\ (\alpha, x)(\bar{t}) = (0, \bar{x}), \end{cases} \tag{3.8}$$

which is formally obtained from (3.6) by adding the variable α and the equation $\dot{\alpha} = \xi$.

It is straightforward to show that the change of coordinates

$$(\alpha, y) = \Gamma(\alpha, y) \doteq (\alpha, exp[-\alpha g_1]x)$$

transforms the vector field $(1, g_1)$ into the constant vector $(1, 0, \dots, 0)$.

Hence, $(\alpha, x)(\cdot) = (u, x)(\cdot)$ is the solution of (3.8) corresponding to ξ if and only if

$$(\alpha, y) \doteq \Gamma \circ (\alpha, x)$$

is the solution to

$$\begin{cases} \dot{\alpha} = \xi \\ \dot{y} = \hat{f}(\alpha, y) \\ (\alpha, y)(\bar{t}) = (0, \bar{x}), \end{cases} \tag{3.9}$$

where, for every $(\alpha, y) = \Gamma(\alpha, x)$,

$$(0, \hat{f})(\alpha, y) \doteq d\Gamma(\alpha, x) \cdot (0, f(x)),$$

$d\Gamma$ denoting the derivative of Γ.

Let us set $u \doteq \int_{\bar{t}}^{t} \xi(s)\, ds$ and let us denote the solution to

$$\begin{cases} \dot{y} = \hat{f}(u, y) \\ y(\bar{t}) = \bar{x} \end{cases}$$

by $\hat{\Phi}(u)$. Hence

$$(u, \Phi(u)) = \Gamma^{-1} \cdot (u, \hat{\Phi}(u)). \tag{3.10}$$

It is easy to show that $\hat{\Phi}(u)$ can be extended from AC to L^1, continuously with respect to the L^1-topology (see [6], [9]). Hence by (3.10), the same can be stated for Φ.

In other words the solution x of (3.6) corresponding to $\xi = \dot{u}$, with $u \in L^1$, coincides with the L^1-limit of (classical) solutions x_n corresponding to a sequence $u_n \in AC$ converging to u in L^1. The above arguments show that x is *independent of the particular sequence* u_n *approximating* u and is defined up to a set of measure zero. (Yet the notion of solution can be refined in order to get a trajectory pointwise defined e.g., at the final time). We conclude this section by remarking that a direct approach "in measure" does not work although it works for the transformed system (3.9). Moreover, it is easy to check that the above given notion of solution x corresponding to a distribution $\xi = \dot{u}$ is nothing but the Γ^{-1}-transform of the (measure) solution (y_0, y) of (3.9), namely $(u, x) = \Gamma^{-1}(y_0, y)$.

A quite critical (and convincing) argument against a direct approach in measure can be found in [22]. An heuristic, non-rigorous explanation of the reason why a "solution in measure" cannot verify elementary properties of well-posedness can be argued by considering, the control $\xi = \dot{u}$ where $u(t) = 0$ for all $t \in [\bar{t}, \hat{t}]$ and $u(t) = 1$ for $t \in]\hat{t}, T]$. Is is clear that a *solution* x to (3.6) corresponding to ξ *has to* jump at $t = \hat{t}$ and that the jump should satisfy a relation like

$$x(\hat{t}_+) - x(\hat{t}_-) = g_1(x(\hat{t}))(u(\hat{t}_+) - u(\hat{t}_-)).$$

Now the question is: which is the "value of $x(\hat{t})$" where g has to be evaluated? It is obvious that an answer to this question is crucial in the attempt of giving to (3.6) a distributional meaning. Yet, any such answer (for example one could decide to take the right-limit of x at $\hat{t}$, or the left-limit, the intermediate point, ect.) turns out to be unsatisfactory in that it does not agree with elementary well-posedness requirements (again, see [22]).

When $g_1, \ldots, g_m$ *verify the commutativity conditions*

$$[g_i, g_j] = 0 \quad \forall i, j = 1, \ldots, m \tag{3.11}$$

the approach pursued in the scalar case can be easily extended to system (3.3) for any $m \geq 1$. The details can be found e.g. in [9]. We only remark that a trasformation of coordinates analogous to the one performed in the linear-case plays a crucial role in the proof of the extendability of the input-output map. More precisely one considers the trasformation $\Gamma : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n$ defined by[1]

$$\Gamma(\alpha, x) \doteq (\alpha, exp[-\alpha_1 g_1] \circ \cdots \circ exp[-\alpha_m g_m] x), \tag{3.12}$$

for every $(\alpha, x) = (\alpha_1, \ldots, \alpha_m, x_1, \ldots, x_n) \in \mathbb{R}^m \times \mathbb{R}^n$.

It is straigthforward to check that Γ changes the system

$$\begin{cases} \dot{\alpha} = \xi \\ \dot{x} = f(x) + \sum_{i=1}^m g_i(x)\xi_i \\ (\alpha, x)(\bar{t}) = (0, \bar{x}) \end{cases} \tag{3.13}$$

into the simpler system

$$\begin{cases} \dot{\alpha} = \xi \\ \dot{y} = \hat{f}(\alpha, y) \\ (\alpha, y)(\bar{t}) = (0, \bar{x}) \end{cases} \tag{3.14}$$

where, for every $(\alpha, y) = \Gamma(\alpha, y)$,

$$(0, \hat{f})(\alpha, y) \doteq d\Gamma(\alpha, x) \cdot (0, f(x)).$$

Therefore one can proceed as in the scalar case and prove the extendability of the input-output map of (3.14), which in turn implies the extendability of the input-output map of (3.13).

[1] If h is a vector field, we use the standard notation $exp[th]$ to mean the function that maps an initial condition y into the value at t of the solution of the Cauchy problem $\dot{x} = f(x),\ x(0) = y$.

3.3 Non commutative systems

For the purpose of continuously extending the input-output map from the set of L^1 controls to the set of controls which are the distributional derivatives of L^1-maps the commutativity condition (3.11) is crucial, as is illustrated by the following example.

Example 3.3.1 Let g_1 and g_2 be smooth vector fields in $\mathbb{R}^n$ such that $[g_1, g_2] \neq 0$ and let us consider the system

$$\begin{cases} \dot{x} = g_1(x)\xi_1 + g_2(x)\xi_2 \\ x(0) = 0, \end{cases} \qquad t \in [0, 2].$$

Let us fix the control $(\xi_1, \xi_2) = \dot{u}, u = (\chi_{[0,1[}, \chi_{[0,1[})$, where χ_E denotes the characteristic function of E.

Both sequences

$$u_n^1(t) = \begin{cases} (1,1) & t \in [0, 1-1/n] \\ (1, 1-n(t-1+1/n)) & t \in [1-1/n, 1] \\ (1-n(t-1), 0) & t \in [1, 1+1/n] \\ (0,0) & t \in [1+1/n, 2] \end{cases}$$

and

$$u_n^2(t) = \begin{cases} (1,1) & t \in [0, 1-1/n] \\ (1-n(t-1+1/n), 1) & t \in [1-1/n, 1] \\ (0, 1-n(t-1)) & t \in [1, 1+1/n] \\ (0,0) & t \in [1+1/n, 2] \end{cases}$$

(are in AC and) converge to u in the L^1-topology.

Yet the solutions x_n^1 corresponding to the controls u_n^1 converge to the map

$$x^1(t) = a^1 \chi_{[1,2]} \quad a^1 \doteq exp(g_1) \circ exp(g_2)(0)$$

while the solutions x_n^2 corresponding to the controls u_n^2 converge to the map

$$x^2(t) = a^2 \chi_{[1,2]} \quad a^2 \doteq exp(g_2) \circ exp(g_1)(0)$$

Since $[g_1, g_2] \neq 0$, in general one has $a^1 \neq a^2$.

This example might suggest that, whereas the input-output of (3.6) fails to be continuously extendable to a set of controls that includes measures, one could try to extend the input-output map corresponding to the graphs of the controls (and of the trajectories). Actually this is the idea underlying the approach pursued in [12], [30]-[33], [34]-[38], [53]-[54]. Let us briefly illustrate the main idea of this approach.

To begin with, let us consider the set

$$\mathcal{U}_K \doteq \{u \in AC([\bar{t}, T], \mathbb{R}^m), V_{\bar{t}}^T(u) \leq K\},$$

where $AC([\bar{t}, T], \mathbb{R}^m)$ denotes the set of absolutely continuous maps from $[\bar{t}, T]$ into $\mathbb{R}^m$ and $V_{\bar{t}}^T(u)$ denotes the variation of u in $[\bar{t}, T]$. Since u is absolutely continuous, one has

$$V_{\bar{t}}^T(u) = \int_{\bar{t}}^T |\dot{u}| \, dt.$$

Now let $\phi_0 : [0, I] \to [\bar{t}, T]$ be a differentiable surjective map, with $\phi_0'(s) \geq 0$ $\forall s \in [0, 1]$, where the sign "$'$" denotes differentiation with respect to the variable s. It is trivial to verify that a map x is the trajectory of (3.6) corresponding to a control $\xi = \dot{u}, u \in \mathcal{U}_K$, if and only if the *space-time trajectory* $(y_0, y) : [0, 1] \to \mathbb{R}^{1+n}$ defined by

$$(y_0, y)(s) \doteq (\phi_0(s), x \circ \phi_0(s)) \quad \forall s \in [0, 1]$$

is the solution of

$$\begin{cases} y_0' = \phi_0' \\ y' = f(y)\phi_0' + \sum g_i(y)\phi_i' \\ (y_0, y)(0) = (0, \bar{x}) \end{cases} \tag{3.15}$$

corresponding to the *space-time control*

$$(\phi_0, \phi)(s) \doteq (\phi_0(s), u \circ \phi_0(s)).$$

In other words (3.15) is the system resulting from (3.6) – more precisely, from (3.6) supplemented with the equation $\dot{t} = 1$– after reparametrizing time by means of $t = \phi_0(s)$. And the *equivalence* between (3.6) and (3.15) relies on the injectivity of ϕ_0. Yet, in principle there is no problem in considering space-time controls (ϕ_0, ϕ) with a ϕ_0 merely non decreasing. Of course, unless ϕ_0 is injective, the space-time control (ϕ_0, ϕ) is not the reparametrization of (the graph of a) control $u \in \mathcal{U}_K$.

In fact a "spatial projection" of (ϕ_0, ϕ), say a selection of $\phi \circ \phi_0^{-1}(t)$, is in general discontinuous. Hence it cannot belong to $\mathcal{U}_K$. In a sense, (ϕ_0, ϕ) contains the information of a "spatial control" $u(t) \in \phi \circ \phi_0^{-1}(t)$ *plus* the extra-information represented by the restriction of ϕ to those intervals where ϕ_0 is constant. Actually, let $[s_1, s_2] \subseteq [0, 1]$ be such that $\phi_0^{'}(s) = 0 \; \forall s \in [s_1, s_2]$. One has $\phi_0(s) = \phi_0(s_1) = \hat{t} \quad \forall s \in [s_1, s_2]$, while in general $\phi(s_1) \neq \phi(s_2)$. Hence $(\phi_0, \phi)(s)$ cannot be the reparametrization of the graph of a continuous control $u(t)$: actually one has $u(\hat{t}_-) = \phi(s_1), u(\hat{t}_+) = \phi(s_2)$. The extra-information is here provided by the restriction $\phi_{|[s_1,s_2]}$. Indeed, in order to compute the "jump" of the trajectory x at t_1, namely the vector

$$x(\hat{t}_+) - x(\hat{t}_-) = y(s_2) - y(s_1),$$

it is not sufficient to know the gap $u(\hat{t}_+) - u(\hat{t}_-) = \phi(s_2) - \phi(s_1)$, because "during" the s-interval $[s_1, s_2]$ the state y evolves according to the controlled dynamics

$$y^{'} = \sum_{i=1}^{m} g_i(y)\phi_i^{'}. \tag{3.16}$$

Therefore in order to calculate the "jump" $y(s_2) - y(s_1)$ one has to know the whole restriction $\phi_{|[s_1,s_2]}$.

Remark 3.3.2 Of course in the commutative case the mere knowledge of the gap $\phi(s_2) - \phi(s_1)$ *is sufficient* for the computation of $y(s_2) - y(s_1)$. This very fact, together with the density theorem below, explain why there are no problems in extending the input-output map when the vector fields g_i commute.

Let us make the above considerations more precise. We shall consider the set Φ_K of *space-time controls with variation bounded by* K, defined by

$$\Phi_K = \Big\{(\phi_0, \phi) \in Lip([0,1], [\hat{t}, T] \times \mathbb{R}^m) : |(\phi_0^{'}, \phi^{'})(s)| \leq K + T, \phi_0^{'}(s) \geq 0 \\ a.e.\, in\, [0,1], \quad \phi_0([0,1]) = [\hat{t}, T], V_0^1(\phi) \leq K\Big\},$$

where $0 \leq \hat{t} < T$ and $Lip(E, F)$ denotes the set of Lipshitz continuous maps from E into F. Let us denote the set of corresponding solution to [3.15] by Λ_K. In the space-time setting, the set $\mathcal{U}_K$ has to be identified with the subset

$$\Phi_K^+ \doteq \big\{(\phi_0, \phi) \in \Phi_K, \phi_0^{'} \geq 0 \quad a.e.\, in\, [0,1]\big\} \subset \Phi_K$$

in a obvious way (see e.g. [34], [36], [47]). Let us denote the corresponding set of solutions to (3.15) by Λ_K^+. The following can be trivially proved:

Lemma 3.3.3 *Let $\bar{\Phi}_K^+$ denote the closure of Φ_K^+ in $C^0([0,1],\mathbb{R}^{1+n})$. Then $\bar{\Phi}_K^+ = \Phi_K$.*

Remark 3.3.4 In Lemma 3.3.3 and in the results below concerning the density of Λ_K^+ and the continuity of the input-output map a topology could be considered on both Φ_K and Λ_K, that is (weaker and) more appropriate than the C^0-topology (see e.g., [6] and [34]). We only point out that this topology keeps track of the free-parameter character of system (3.15), namely of the fact that if $(y_0, y)(s)$ is the solution corresponding to $(\phi_0, \phi)(s)$ and $s(\sigma)$ is a reparametrization of $[0,1]$ than $(y_0, y) \circ s(\sigma)$ is the solution corresponding to $(\phi_0, \phi) \circ s(\sigma)$. Yet we use the C^0 topology, for this choice is not too restrictive and, at the same time, allows one to simplify the presentation of several subjects.

Theorem 3.3.5 *([12]). The input-output map $\mathcal{I} : \Phi_K :\to \Lambda_K$, which to every $(\phi_0, \phi) \in \Phi_K$ associates the corresponding solution to (3.15), is continuous w.r. to the C^0 norm on both Φ_K and Λ_K.*

From the relations $\Lambda_K^+ = \mathcal{I}(\Phi_K^+), \Lambda_K = \mathcal{I}(\Phi_K)$ and the continuity of $\mathcal{I}$ we obtain:

Corollary 3.3.1 *Let $\bar{\Lambda}_K^+$ denote the closure of Λ_K^+ in the C^0 norm. Then $\bar{\Lambda}_K^+ = \Lambda_K$.*

Now let us come back to our original question, that is: *can one give a notion of solution to (3.3) corresponding to a control $\xi = \dot{u}$, with u an L^1-map?* Let us restrict our attention (see Remark 3.3.7 below) to the set of controls

$$\mathcal{U}_K^* \doteq \left\{ u \in BV([\bar{t}, T], \mathbb{R}^m), V_{\bar{t}}^T(u) \leq K \right\}.$$

Observe that the original set $\mathcal{U}_K$ is dense in $\mathcal{U}_K^*$ w.r. to the L^1-topology.

Fix $u \in \mathcal{U}_K^*$. On one hand, the previous example show that the we cannot extend continuously the notion of solution corresponding to a control u from $\mathcal{U}_K$ to $\mathcal{U}_K^*$. In fact, distinct sequences, $(u_n), (\tilde{u}_n)$ in $\mathcal{U}_K$ approximating u give rise to sequences of solutions (x_n) and $(\tilde{x}_n)$ respectively, that in general converge to distinct limits.

On the other hand one can consider a *graph-completion* of u.

Definition 3.3.6 *A* graph-completion *of u is a space-time control $(\phi_0, \phi) \in \Phi_K$ such that $u([\hat{t}, T]) \subseteq \phi([0,1])$.*

Notice that a graph-completion always exists: it is sufficient to bridge the gaps of u with straight lines and to parametrize the so-obtained (rectifiable) $(m+1)$ path with an abscissa that is proportional to the arc-length (in this case the total variation of ϕ equals the total variation of u). Of course, there is no special reason to prefer *this* graph-completion to another one where gaps are filled with curves which are not rectilinear (see the next section).

Now, whereas we are not able to define a solution of (3.3) corresponding to a control $u \in \mathcal{U}_K^*$, we can consider the solution $(y_0, y)(s)$ of (3.15) corresponding to a graph-completion (ϕ_0, ϕ) of u. If one whishes to come back to the parameter t, any selection of $y_0 \circ y_0^{-1}(t)$ can be considered as a solution of (3.3) corresponding to the graph-completion (ϕ_0, ϕ).

Notice that in view of Corollary 3.3.1 the space-time trajectory (y_0, y) can be uniformly approximated by means of trajectories $(y_0^n, y^n) \in \Lambda_K^+$ corresponding to space-time controls (ϕ_0^n, ϕ^n) uniformly converging to (ϕ_0, ϕ). In practice this means that the solution (y_0, y) corresponding to a graph-completion (ϕ_0, ϕ) of u can be approached by the graphs of the solutions x^n corresponding to controls u^n whose (reparametrized) graphs converge to (ϕ_0, ϕ).

We have constructed a class of trajectories, namely the set Λ_K, in which the class of original trajectories, here identified with Λ_K^+, is dense. The set Λ_K concides with the set of outputs of system (3.15) corresponding to the inputs in Φ_K, the set of graph-completions (it is obvious that *each* space-time control (ϕ_0, ϕ) is the graph-completion of a suitable control $u \in \mathcal{U}_K^*$). This gives an answer to the question of defining a notion of output of (3.3) corresponding to a control $\xi = \dot{u}$, with $u \in \mathcal{U}_K^*$. In fact, in order to define such an output we have to *choose* –in principle, arbitrarily – a graph-completion (ϕ_0, ϕ) of u and to consider the corresponding trajectory. Let us remark that for a wide class of optimal control problems having (3.3) as dynamics the optimal control *exists* only in the set Φ_K (see next section). In this case the optimal strategy consists of a control $u \in \mathcal{U}_K^*$ and a *suitable* graph-completion of u.

Remark 3.3.7 When the variation of u is not bounded the situation is much more involved. For example, the input-output map is not continuous *even when it is restricted to continuous controls*, as is shown in [59]. The question is addresses in [10] where a concept of *looping control* is introduced: roughly speaking a looping control is a represantation of the *limit* of a sequence of controls with larger and larger variation.

3.4 Optimal control problems

In order to give an idea of the questions involved in optimal control problems subject to a dynamics of the form (3.3), let us consider the following Mayer problem: let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a continuous map and let us define the final cost functional as

$$C(u) \doteq \psi(x[u](T)),$$

where $x[u](\cdot)$ denotes the trajectory of

$$\begin{cases} \dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)\dot{u}_i \\ x(\bar{t}) = \bar{x} \end{cases} \tag{3.17}$$

corresponding to the control $\xi = \dot{u}$. The problem

$$minimize_{u \in \mathcal{U}_K} C(\dot{u}) \tag{3.18}$$

consists in looking for a control $\hat{u} \in \mathcal{U}_K$ such that $C(\hat{u}) = min_{u \in \mathcal{U}_K} C(u)$.

We begin by observing that an optimal control $\hat{u} \in \mathcal{U}_K$ in general does not exists. This is due essentially to the lack of compactness of $\mathcal{U}_K$. In order to frame this phenomenon of non-existence in a more theoretical picture we begin by observing that it is not related to any lack of convexity: no chattering controls are needed here. Instead, the non-existence of an optimal control $\hat{u} \in \mathcal{U}_K$ is linked to a fact that, in analogy with a similar pathology occurring in the Calculus of Variations, could be called *lack of coerciveness*. As a matter of fact, in the classical problem

$$\begin{array}{l} minimize \\ \end{array} \begin{array}{l} x \in AC \\ x(a) = A \\ x(b) = B \end{array} \int_a^b L(x, \dot{x})\, dt, \tag{3.19}$$

the term *coerciveness* refers to the fact that the Lagrangian L is *superlinear* in $\dot{x}$. Roughly speaking, this avoids the occurrence of minimizing sequences of trajectories that converge to a discontinuous map. In Control Theory, on one hand several problems involve a *bounded* dynamics: this is sufficient to avoid jumps of the optimal trajectories. On the other hand it often happens that, although the dynamics is unbounded, a sort of coerciveness condition penalizes the use of too large velocities. For instance, this is the case of the so-called linear-quadratic problems.

On the contrary, problem (3.18) involves a unbounded dynamics – $\dot{u}$ has no bounds – and no condition makes the use of larger and larger controls unfavorable for the pursuit of the infimum of C. The Example 3.4.1 below provides a typical situation of non-existence of an optimal control. Successively we shall present an extension of the general problem (3.18) and test this extension with the special problem of Example 3.4.1.

Example 3.4.1 Let us consider the system

$$\begin{cases} \dot{x}_1 = \dot{u}_1 \\ \dot{x}_2 = \dot{u}_2 \\ \dot{x}_3 = arctan(x_1^2 + x_2^2) - \frac{x_2 \dot{u}_1}{2} + \frac{x_1 \dot{u}_2}{2} \\ (x_1, x_2, x_3) = (0,0,0) \end{cases} \qquad t \in [0,1] \tag{3.20}$$

and let us address the problem of minimizing the Mayer functional

$$C(u) = x_3[u](1)$$

over the set $\mathcal{U}_\pi$, where $x[u](\cdot)$ denotes the solution to (3.20) and $x_i[u], i = 1,2,3$, is its i-th component. Since the integral

$$\int_0^\alpha \frac{x_1(\sigma)\dot{x}_2(\sigma) - \dot{x}_2(\sigma)x_1(\sigma)}{2} \, d\sigma$$

coincides with the *area spanned by the vector* $(x_1, x_2)(\sigma)$ *in the interval* $[0, \alpha]$, it is not difficult to check that the maps

$$\hat{u}_n(t) \doteq [cos(\pi n(1 - 1/n - t)) - 1, sin(\pi n(1 - 1/n - t))]\chi_{[1-1/n,1]}$$

define a minimizing sequence. Moreover, one has

$$inf_{u \in \mathcal{U}_\pi} C(u) = lim_{n\to\infty} C(\hat{u}_n) = -\pi/2.$$

Now, one can easily verify that every minimizing sequence $\tilde{u}_n$ verifies

$$\begin{aligned} &lim_{n\to\infty} x_3(\tilde{u}_n)(t) = 0 \quad \forall t \in [0,1[, \\ &lim_{n\to\infty} x_3(\tilde{u}_n)(t) = -\pi/2, \end{aligned}$$

which, in particular, implies that no optimal control exists in $\mathcal{U}_\pi$. Also, observe that although the sequence

$$\check{u}_n \doteq [cos(\pi n(t - 1 + 1/n)) - 1, sin(\pi n(t - 1 + 1/n))]\chi_{[1-1/n,1]}$$

it is not minimizing (in that $lim_{n\to\infty} C(\check{u}_n) = \pi/2$) one has

$$lim|\hat{u}_n(t) - \check{u}_n(t)| = 0 \quad \forall t \in [0,1].$$

Extension of problem (3.18). We now introduce an *extension* of the problem (3.18), that is, a new minimum problem (3.22) having the following features:

(i) the infimum value of (3.18) coincides with the infimum value of (3.22);

(ii) the class of elements where one seeks the infimum of (3.18) is dense in the corresponding class of (3.22), with respect to a topology for which the functional of (3.22) is continuous.

We shall consider the extended cost functional

$$C_e(\phi_0, \phi) \doteq \psi(y[\phi_0, \phi](1))$$

where $(y_0, y)[\phi_0, \phi](\cdot)$ denotes the solution of

$$\begin{cases} y_0' = \phi_0' \\ y' = f(y)\phi_0' + \sum_{i=1}^m g_i(y)\phi_i' \\ (y_0, y)(0) = (0, \bar{x}) \end{cases} \tag{3.21}$$

corresponding to the space-time control $(\phi_0, \phi) \in \Phi_K$. We define the extended problem (3.22) as

$$minimize_{(\phi_0,\phi)\in\Phi_K} C_e(\phi_0, \phi). \tag{3.22}$$

Theorem 3.4.2 *Problem (3.22) admits an optimal control* $(\hat{\phi}_0, \hat{\phi}) \in \Phi_K$, *that is*

$$C_e(\hat{\phi}_0, \hat{\phi}) = inf_{(\phi_0,\phi)\in\Phi_K} C(\phi_0, \phi). \tag{3.23}$$

Moreover, for every sequence $(\phi_0^n, \phi^n)(\cdot)$ *in* Φ_K^+ *which converges to* (ϕ_0, ϕ) *one has that the sequence* (u^n) *in* $\mathcal{U}_K$ *defined by*

$$u^n(t) = \phi^n \circ (\phi_0^n)^{-1}(t) \quad \forall t \in [\bar{t}, T]$$

verifies

$$lim_{n\to\infty} C(u^n) = inf_{u\in\mathcal{U}_K} C(u) = C_e(\hat{\phi}_0, \hat{\phi}). \tag{3.24}$$

Remark 3.4.3 In particular (3.23)-(3.24) yield

$$inf_{u\in\mathcal{U}_K} C(u) = inf_{(\phi_0,\phi)\in\Phi_K} C_e(\phi_0, \phi)$$

Proof of Theorem 3.4.2 By Theorem 3.3.5 the map

$$(\phi_0, \phi) \to \psi(y[\phi](1))$$

is continuous with respect to the C^0-norm. Moreover by Ascoli-Arzelà's theorem the set Φ_K is compact, from which the existence of an optimal control follows. The remaining part of the thesis is a straightforward consequence of the density of Φ_K^+ in Φ_K and of the continuity of the input-output map $\mathcal{I}$.

Application to Example 3.4.1. The space-time versions of the system and the cost-fuctional in Example 3.4.1 are given by

$$\begin{cases} y_0' = \phi_0' \\ y_1' = \phi_1' \\ y_2' = \phi_2' \\ y_3' = arctan(y_1^2 + y_2^2) - \frac{y_2 \phi_1'}{2} + \frac{y_1 \phi_2'}{2} \\ (y_0, y_1, y_2, y_3)(0) = (0, 0, 0, 0), \end{cases}$$

and

$$C_e(\phi_0, \phi) = y_3[\phi_0, \phi_1](1),$$

respectively. By Theorem 3.4.2 the extended problem

$$minimize_{(\phi_0,\phi)\in\Phi_\pi} C_e(\phi_0, \phi)$$

has one solution. By the considerations made in Example 3.4.1 it is clear that the space-time control

$$(\hat{\phi}_0, \hat{\phi}(s)) \doteq (2s, 0, 0)\chi_{[0,1/2[}(s) + (0, cos(\pi - 2\pi s) - 1, sin(\pi - 2\pi s))\chi_{[1/2,1]}(s)$$

is optimal for the extended problem.

It is straightforward to check that the sequence $(\hat{\phi}_0^n, \hat{\phi}^n)$ defined by

$$(\hat{\phi}_0^n, \hat{\phi}^n(s)) \doteq \Big((s - s/n)\chi_{[0,1/2[}(s) + (1 + s/n - 1/n)\chi_{[1/2,1]}(s), \hat{\phi}(s) \Big)$$

lies in Φ_π, converges to $(\hat{\phi}_0, \hat{\phi})$, and verifies

$$\hat{u}_n(t) = \hat{\phi}^n \circ (\hat{\phi}_0^n)^{-1}(t)$$

for every n and every $t \in [0, 1]$, where $(\hat{u}_n)$ is the sequence of maps defined in Example 3.4.1. This explains why $(\hat{u}_n)$ is a minimizing sequence, while

$(\check{u}_n)$ *is not* a minimizing sequence. In fact, the space-time controls $(\check{\phi}_0^n, \check{\phi}^n)$ defined by

$$(\check{\phi}_0^n, \check{\phi}^n)(s) \doteq (\hat{\phi}_0^n(s), \check{\phi}(s))$$
$$\check{\phi}(s) \doteq (0, cos(2\pi s - \pi) - 1, sin(2\pi s - \pi))$$

are space-time representations of the $\check{u}_n$, that is

$$\check{u}_n(t) = \check{\phi}^n \circ (\check{\phi}_0^n)^{-1}(t),$$

and they converge uniformly to the space-time control $(\hat{\phi}_0, \check{\phi})$, which is not optimal, for

$$C_e(\hat{\phi}_0, \check{\phi}) = \frac{\pi}{2}.$$

We conclude by remarking that both $(\hat{u}_n)$ and $(\check{u}_n)$ converge pointwise to the control $\hat{u} = (-1, 0)\chi_{\{1\}}$. Yet the information that $(\hat{u}_n)$ [resp. $\check{u}_n$] is [resp. is not] a minimizing sequence cannot be deduced by the mere knowledge of $\hat{u}$ (or equivalently, of $\xi = (-1, 0)u = (-1, 0)\delta_{\{1\}}$). On the contrary, this information is contained in the specification of the graph-completion $(\hat{\phi}_0, \hat{\phi})$ of $\hat{u}$.

Part II

Moving, bilateral, constraints as controls for mechanical systems

Chapter 4

Control of holonomic systems

4.1 Introduction

Let N, M, ν be positive integers such that $\nu \leq N$. Let $\mathcal{Q}$ be an $(N+M)$-dimensional differential manifold, which will be regarded as the space of configurations of a mechanical system [1].

The system will be controlled by means of an active (holonomic, time-dependent) constraint. Let an M-dimensional differential manifold $\mathcal{U}$ be given, together with a surjective submersion

$$\pi : \mathcal{Q} \mapsto \mathcal{U}. \tag{4.1}$$

The fibers $\pi^{-1}(\mathbf{u}) \subset \mathcal{Q}$ will be regarded as the states of the active constraint. The set of all these fibers can be identified with the *control manifold* $\mathcal{U}$.

Let $\mathcal{I}$ be a time interval and let $\mathbf{u} : \mathcal{I} \mapsto \mathcal{U}$ be a continuously differentiable map. We say that *a trajectory* $\mathbf{q} : \mathcal{I} \to \mathcal{Q}$ *agrees with the control* $\mathbf{u}(\cdot)$ *if*

$$\pi \circ \mathbf{q}(t) \;=\; \mathbf{u}(t) \qquad\qquad \forall t \in \mathcal{I}\,. \tag{4.2}$$

For each $\mathbf{q} \in \mathcal{Q}$, consider the subspace of the tangent space at $\mathbf{q}$ given by

$$\Delta_{\mathbf{q}} \;\doteq\; ker\; T_{\mathbf{q}}\pi\,.$$

Here $T_{\mathbf{q}}\pi$ denotes the linear tangent map between the tangent spaces $T_{\mathbf{q}}\mathcal{Q}$ and $T_{\pi(\mathbf{q})}\mathcal{U}$. Clearly, Δ is the integrable distribution whose integral manifolds are precisely the fibers $\pi^{-1}(\mathbf{u})$.

The general setting will be as follows:

[1] As customary, we assume that such mechanical system is obtained by a finite collection of material points and rigid bodies connected by time-independent holonomic constraints, physically obeying to D'Alembert Principle.

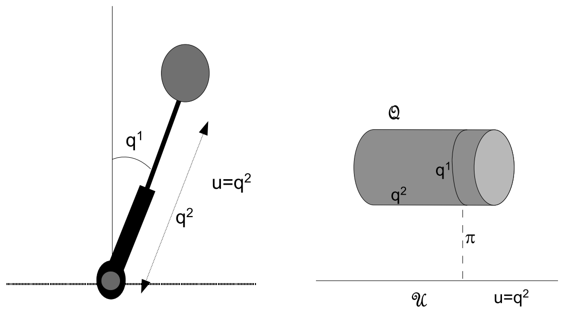

Figure 4.1: Length as control

1) The manifold $\mathcal{Q}$ is endowed with a Riemannian metric $\mathbf{g} = \mathbf{g_q}[\cdot,\cdot]$, the so-called *kinetic metric*, which defines the kinetic energy $\mathcal{T}$. More precisely, for each $\mathbf{q} \in \mathcal{Q}$ and $\mathbf{v} \in T_{\mathbf{q}}\mathcal{Q}$ one has

$$\mathcal{T}(\mathbf{q},\mathbf{v}) \doteq \frac{1}{2}\mathbf{g_q}[\mathbf{v},\mathbf{v}]\,. \tag{4.3}$$

We shall use the notation $\mathbf{v} \mapsto \mathbf{g_q}(\mathbf{v})$ to denote the isomorphism from $T_{\mathbf{q}}\mathcal{Q}$ to $T^*_{\mathbf{q}}\mathcal{Q}$ induced by the scalar product $\mathbf{g_q}[\cdot,\cdot]$. Namely, for every $\mathbf{v} \in T_{\mathbf{q}}\mathcal{Q}$, the 1-form $\mathbf{g_q}(\mathbf{v})$ is defined by

$$\langle \mathbf{g_q}(\mathbf{v}),\mathbf{w}\rangle \doteq \mathbf{g_q}[\mathbf{v},\mathbf{w}] \qquad \forall \mathbf{w} \in T_{\mathbf{q}}\mathcal{Q}\,, \tag{4.4}$$

where $\langle\cdot,\cdot\rangle$ is the duality between the tangent space $T_{\mathbf{q}}\mathcal{Q}$ and the cotangent space $T^*_{\mathbf{q}}\mathcal{Q}$.

If $\mathbf{q} \in \mathcal{Q}$ and $W \subset T_{\mathbf{q}}$, $W^\perp$ denotes the subspace of $T_{\mathbf{q}}$ consisting of all vectors that are orthogonal to every vector in W:

$$W^\perp \doteq \{\mathbf{v} \in T_{\mathbf{q}} \quad | \quad \mathbf{g_q}[\mathbf{v},\mathbf{w}] = 0 \quad \forall \mathbf{w} \in W\}\,. \tag{4.5}$$

For a given distribution $E \subset T\mathcal{Q}$, the *orthogonal distribution* $E^\perp \subset T\mathcal{Q}$ is defined by setting $E^\perp_{\mathbf{q}} \doteq (E_{\mathbf{q}})^\perp$, for every $\mathbf{q} \in \mathcal{Q}$.

We use $\mathcal{P}^E : T_{\mathbf{q}}\mathcal{Q} \to E_{\mathbf{q}}$ to denote the *orthogonal projection* on $E_{\mathbf{q}}$. Namely, for every $\mathbf{v} \in T_{\mathbf{q}}\mathcal{Q}$, $\mathcal{P}^E(\mathbf{v})$ is the unique vector in $E_{\mathbf{q}}$ such that $\mathbf{v} - \mathcal{P}^E(\mathbf{v}) \in E^\perp_{\mathbf{q}}$.

2) The mechanical system is subject to *forces*. In the Hamiltonian formalism, these are represented by vertical vector fields on the cotangent bundle $T^*\mathcal{Q}$. We recall that, in a natural local system of coordinates (q,p), the fact that $\mathbf{F}$ is *vertical* means that its q-component is zero, namely $\mathbf{F} = \sum_{i=1}^{N+M} F_i \frac{\partial}{\partial p_i}$.

3) The constraints (5.1) and (5.3) are dynamically implemented by reaction forces obeying

D'ALEMBERT PRINCIPLE: *If* $t \mapsto \mathbf{q}(t)$ *is a trajectory which satisfies both (5.1) and (5.3), and* $\mathbf{R}(t)$ *is the constraint reaction at a time* t*, then*

$$\mathbf{R}(t) \;\in\; \ker\Big(\Delta_{\mathbf{q}(t)} \cap \Gamma_{\mathbf{q}(t)}\Big). \tag{4.6}$$

In other words, regarding the reaction force $\mathbf{R}(t)$ as an element of the cotangent space $T^*_{\mathbf{q}(t)}\mathcal{Q}$, one has

$$\langle \mathbf{R}(t)\,,\ \mathbf{v}\rangle \;=\; 0 \qquad\qquad \forall \mathbf{v} \;\in\; \Delta_{\mathbf{q}(t)} \cap \Gamma_{\mathbf{q}(t)} \;\subseteq\; T_{\mathbf{q}(t)}\mathcal{Q}\,. \quad {}^{2} \tag{4.7}$$

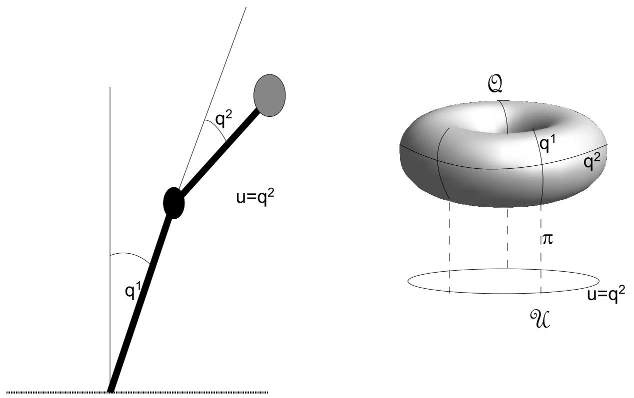

Figure 4.2: Second pendulum as control

We deal here with the case without non-holonomic constraints (formally $\Gamma = T\mathcal{Q}$). We begin by presenting the control equation of a curve agree-

[2]By considering the "vector $V_{\mathbf{R}}(\cdot)$ corresponding to $\mathbf{R}(\cdot)$ ", namely $V_{\mathbf{R}} \doteq \mathbf{g}^{-1}(\mathbf{R})$, one has $\mathbf{g}_{\mathbf{q}(t)}[V_{\mathbf{R}}, \mathbf{v}] = 0$ for all $\mathbf{v} \in \Delta_{\mathbf{q}(t)}$, that is $\mathbf{R}$ *is "orthogonal" to* $\Delta_{\mathbf{q}(t)}$, which is the way D'Alembert Principle is sometimes formulated.

ing with a control $\mathbf{u}(\cdot)$(under the assumptions stated in the Introduction). These equations can be formulated in an intrinsic way (see [29], [11]). We prefer to state them in coordinates, and we refer e.g. to [?], [?], [8] for their deductions (see also [19] and [29]). Successively we will see how the functional dependence of the dynamics on the derivative $\dot{u}$ is related to the geometric properties of the foliation $\Lambda \doteq \{\pi^{-1}(\mathbf{u}) \quad \mathbf{u} \in \mathcal{U}\}$. In particular

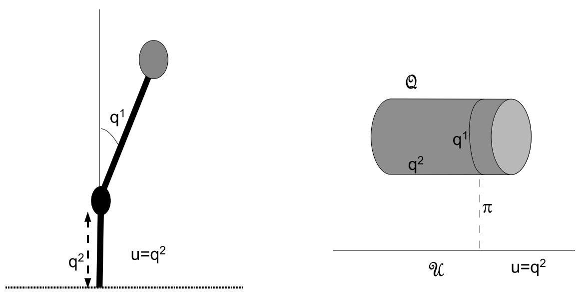

Figure 4.3: A pendulum with vertically moving pivot.

the quadratic dependence is equivalent to the existence geodesics orthogonal to Λ. This, in turn, is necessary and sufficient for a certain *orthogonal curvature* is equal to zero. We remark (see Part I) that the non-vanishing of the term quadratic in $\dot{u}$ is crucial for *vibrational* stabilization.

4.2 Control equations

Consider a Δ-adapted coordinate chart $(q^i, q^{N+\alpha})$, $i = 1, \ldots, N, \alpha = 1, \ldots, M$, on an open subset $\mathcal{O}$ of $\mathcal{U}$, and a coordinate chart (u^α) on a subset of the projection $\pi(\mathcal{O}) \subset \mathcal{U}$. We can obviously set $x^{N+\alpha} = u^\alpha$, $\alpha = 1, \ldots, M$. So, without danger of confusion, we shall sometimes refer to (q^i, u^α) (instead of $(q^i, q^{N+\alpha})$) as Δ-adapted coordinates. We also use the compact notation $(q^\sharp, q^\flat)$, where $q^\sharp \doteq (q^1, \ldots, q^N)$, $q^\flat \doteq (q^{N+1}, \ldots, q^{N+M})$. Let $\Big((q^i, u^\alpha), (p_i, p_{N+\alpha})\Big) = (q^\sharp, q^\flat, p_\sharp, p_\flat)$ be the corresponding local coordinates on the fiber bundle $T^*\mathcal{Q}$. Let $G = (g_{r,s})_{r,s=1,\ldots,N+M}$ be the matrix representing the kinetic metric $\mathbf{g}$, and let $G^{-1} = (g^{r,s})_{r,s=1,\ldots,N+M}$ denote its inverse.

In the following, we consider the sub-matrices

$$G_1 \doteq (g_{i,j})\,, \qquad G_2 \doteq (g_{N+\alpha,N+\beta})\,, \qquad (G^{-1})_2 \doteq \left(g^{N+\alpha,N+\beta}\right)\,,$$

$$G_{12} \doteq (g_{i,N+\alpha})\,, \qquad (G^{-1})_{12} \doteq \left(g^{i,N+\alpha}\right)\,,$$

with the convention that the Latin indices i,j run from 1 to N, while the Greek indices α,β run from 1 to M. Let us also set

$$\begin{aligned} A = (a^{i,j}) \doteq (G_1)^{-1}, \qquad E = (e_{\alpha,\beta}) \doteq ((G^{-1})_2)^{-1}, \\ K = (k^i_{N+\alpha}) \doteq (G^{-1})_{12}E\,. \end{aligned} \tag{4.8}$$

Proposition 4.2.1 *Let $\mathbf{u}(\cdot) : I \mapsto \mathcal{U}$ be twice continuously differentiable, and let $(\mathbf{q},\mathbf{p}) : I \mapsto T^*\mathcal{Q}$ be a curve verifying the Hamiltonian equations of motion and such that $\mathbf{q}$ agrees with $\mathbf{u}$.* [3]

Then the coordinate maps $t \mapsto (q(t), p_\sharp(t)) = \left(q^\sharp(t), q^\flat(t), p_\sharp(t)\right)$ satisfy the control equation differential equation

$$\begin{pmatrix} \dot{q^\sharp} \\ \dot{q^\flat} \\ \dot{p_\sharp} \end{pmatrix} = \begin{pmatrix} A{p_\sharp}^t \\ 0 \\ -\frac{1}{2}{p_\sharp}^t\frac{\partial A}{\partial q}p_\sharp \end{pmatrix} + \begin{pmatrix} K\dot{u} \\ \dot{u} \\ -{p_\sharp}^t\frac{\partial K}{\partial q^\sharp}\dot{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2}\dot{u}^t\frac{\partial E}{\partial q}\dot{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ F^{u(\cdot),\dot{u}(\cdot)} \end{pmatrix}, \tag{4.9}$$

where

$$F^{u(\cdot),\dot{u}(\cdot)} \doteq \left(F_1^{u(\cdot),\dot{u}(\cdot)}, \ldots, F_N^{u(\cdot),\dot{u}(\cdot)}\right). \tag{4.10}$$

For convenience, in (4.9) we have written all vectors as column vectors, while the superscript ${}^{\mathtt{t}}$ denotes transposition. Componentwise, (4.9) reads:

$$\begin{cases} \dot{q}^i = a^{i,j}p_j \;+\; k^i_{N+\alpha}\dot{u}^\alpha \\ \dot{q}^\alpha = \;\dot{u}^\alpha \\ \dot{p}_i = -\frac{1}{2}\frac{\partial a^{\ell,j}}{\partial q^i}p_\ell p_j \;-\; \frac{\partial k^j_\alpha}{\partial q^i}p_j\dot{u}^\alpha + \frac{1}{2}\frac{\partial e_{\alpha,\beta}}{\partial q^i}\dot{u}^\alpha\dot{u}^\beta + F_i^{u(\cdot),\dot{u}(\cdot)}. \end{cases} \tag{4.11}$$

As announced above, we do not give here a proof of Proposition 4.2.1, for which we refer to the quoted references. However, we shall hint how to deduce the control equations (4.11) from the more general case where non holonomic constraints act on the system as well (see Subsection 5.3.3).

[3] By possibly restricting the size of the interval I, we can assume that $\mathbf{q}(t)$ remains inside the domain of the single chart $(q, q^\flat)$ for every $t \in I$.

4.3 The input-output map

The presence of the derivative $\dot{u}$ in the dynamic equations (4.9) depends on the Riemannian metric $\mathbf{g}$ defining the kinetic energy and the integral foliation Γ of Δ,

$$\Lambda \doteq \{\pi^{-1}(\mathbf{u}) \quad \mathbf{u} \in \mathcal{U}\}. \tag{4.12}$$

In this section we review the main results in this direction. To simplify the discussion, throughout this section we shall assume that the additional forces $\mathbf{F}$ vanish identically, so that in (4.10) one has

$$F_{\mathcal{Q}}^{u(\cdot),\dot{u}(\cdot)} \equiv 0\,.$$

The following definitions were introduced in [**?**].

Definition 4.3.1 *A local, Δ-adapted, system of coordinates $(q^\sharp, q^\flat)$ on $\mathcal{Q}$ is called* N-fit for hyperimpulses *if, for every differentiable control function* $\mathbf{u}(\cdot)$*, the right-hand side of the corresponding equation of motion (4.11) does not contains any quadratic term in the variable $\dot{u}$.*

A local, Δ-adapted, system of coordinates $(q^\sharp, q^\flat)$ on $\mathcal{Q}$ is called strongly N-fit for hyperimpulses *if, for every differentiable control function* $\mathbf{u}(\cdot)$*, the right-hand side of (4.11) is independent of the variable $\dot{u}$.*

Moreover, we shall call generic *any local , Δ-adapted, system of coordinates $(q^\sharp, q^\flat)$ which is not N-fit for hyperimpulses.*

Remark 4.3.2 The denomination "*N-fit for hyperimpulses*" for a system of coordinates $(q^\sharp, q^\flat)$ refers to the fact that, if the dependence on $\dot{u}$ is only linear, one can then construct solutions $\left(q^\sharp(\cdot), p_\sharp(\cdot)\right)$ also for discontinuous controls $u(\cdot)$. In general, a jump in $u(\cdot)$ will produce a discontinuity of both $q(\cdot)$ and $p(\cdot)$. For this reason we call it a *hyperimpulse*, as opposite to *impulse*, which can cause a discontinuity in the component $p(\cdot)$ only.

A first characterization of N-fit coordinates was derived in [**?**]. It is important to observe that the property of being N-fit depends only on the metric g and on the foliation Λ, while it is independent of the particular the system of Δ-adapted coordinates. This allows one to give the following definitions.

Definition 4.3.3 [**?**] *The foliation Λ is called* N-fit for hyperimpulses *if there exists an atlas of Δ-adapted charts that are also N-fit for hyperimpulses. In this case,* all *Δ-adapted charts are N-fit for hyperimpulses.*

The foliation Λ is called strongly N-fit for hyperimpulses *if there exists an atlas of Δ-adapted charts which are strongly N-fit for hyperimpulses.*

Moreover, the foliation Λ will be called generic *if it is not N-fit for hyperimpulses.*

The paper [**?**] established the connection between the N-fitness of the foliation Λ and the bundle-like property of the metric, introduced in [49, 48]. We recall here the main definitions and results.

Definition 4.3.4 *The metric* $\mathbf{g}$ *is* bundle-like *with respect to the foliation Λ if, for one (hence for every) Λ- adapted chart, it has a local representation of the form*

$$\sum_{i,j=1}^{N} g_{i,j}(q^\sharp, q^\flat)\omega^i \otimes \omega^j + \sum_{\alpha,\beta=1}^{M} g_{N+\alpha,N+\beta}(q^\flat)dq^{N+\alpha} \otimes dq^{N+\beta}\,,$$

where $\omega^1, \dots, \omega^N$ are linearly independent 1-forms such that, for each $\mathbf{q} \in \mathcal{Q}$ in the domain of the chart, one has

(i) $\left(\omega^1, \dots, \omega^N, dq^{N+1}, \dots, dq^{N+M}\right)$ is a basis of the cotangent space $T^*_{\mathbf{q}}\mathcal{Q} \times T^*_{\mathbf{u}}\mathcal{U}$;

(ii) $\left\langle \omega^i(\mathbf{q}), Y \right\rangle = 0$, for every $Y \in \Delta^\perp_{\mathbf{q}}$ and all $i = 1, \dots, N$.

We recall that $\Delta^\perp_{\mathbf{q}}$ is the orthogonal bundle (see defined at (4.5)). If g is bundle-like with respect to the foliation Λ, the latter is also called a *Riemannian foliation*, because in this case a Riemannian structure can be well defined also on the quotient space.

Theorem 4.3.5 *Consider the foliation Λ as in (4.12). The following statements are equivalent:*

i) The foliation Λ is N-fit for hyperimpulses.

ii) The metric $\mathbf{g}$ *is bundle-like with respect to the foliation Λ, i.e., the foliation Λ is Riemannian.*

iii) For any $\mathbf{u}, \bar{\mathbf{u}} \in \mathcal{U}$ the map $d_{\mathbf{u},\bar{\mathbf{u}}}(\cdot) : \mathcal{Q} \mapsto I\!R$ defined by

$$d_{\mathbf{u}}(\mathbf{q}) \doteq dist\, \left(\mathbf{q}, \pi^{-1}(\bar{\mathbf{u}})\right)$$

is constant. In other words, given two leaves, the points of one of the two are all at the same distance from the other leaf. (This allows one to define a metric on the set of leaves.)

iv) If $t \mapsto \mathbf{q}(t)$ *is any geodesic curve with respect to the metric* $\mathbf{g}$*, and if* $\dot{\mathbf{q}}(\tau) \in \Delta^{\perp}_{(\mathbf{q}(\tau),\mathbf{u}(\tau))}$ *at some time* τ*, then* $\dot{\mathbf{q}}(t) \in \Delta^{\perp}_{\mathbf{q}(t)}$ *for all* t*. In other words, if a geodesic crosses perpendicularly one of the leaves, then it crosses perpendicularly also every other leaf it meets.*

v) If $(q^\sharp, q^\flat)$ *is a* Δ*-adapted system of coordinates, then*

$$\frac{\partial g^{N+\alpha,N+\beta}}{\partial q_i} = 0 \qquad i = 1,\ldots,N, \quad \alpha,\beta = 1,\ldots,M\,, \tag{4.13}$$

where $G^{-1} = (g^{r,s})$ *denotes the inverse of the matrix* $G = (g_{r,s})$ *representing the metric* $\mathbf{g}$ *in the coordinates* $(q^\sharp, q^\flat)$*.*

Proof. The equivalence of $i)$ and $ii)$ is a trivial consequence of the definitions of bundle-like metric and of N-fit system of coordinates. The equivalence of $ii)$, $iii)$, and $iv)$, is a classical result on bundle-like metrics [49]. Moreover, by (4.11), the foliation is fit for jumps if and only if $\partial e_{\alpha,\beta}/\partial q^i \equiv 0$. Recalling that the matrix $(e_{\alpha,\beta})$ is the inverse of $(G^{-1})_2 = (g^{N+\alpha,N+\beta})$, we conclude that $i)$ is equivalent to $v)$.

$\diamondsuit$

Theorem 4.3.6 *The following statements are equivalent:*

i) The foliation Λ *is strongly* N*-fit for hyperimpulses .*

ii) The foliation Λ *is* N*-fit for hyperimpulses and the orthogonal bundle* $\Delta^{\perp}_{\mathbf{q}}$ *in (4.5) is integrable.*

iii) There is an atlas such that, for every (Δ*-adapted) chart* $(q^\sharp, q^\flat)$*, one has*

$$\frac{\partial g^{N+\alpha,N+\beta}}{\partial q_i} = 0, \qquad g^{i,N+\alpha} = 0 \qquad \forall i = 1,\ldots,N, \quad \alpha,\beta = 1,\ldots,M.$$

Indeed the equivalence of $i)$ and $ii)$, formulated in terms of Riemannian foliations, was proved in [49]. The equivalence between $i)$ and $iii)$ follows from (4.11). See again [49] and [**?**] for details.

4.4 The orthogonal curvature

For any $(q^\sharp, q^\flat)$ in the range of a Δ-adapted chart, let us consider

$$\frac{\partial e_{\alpha,\beta}}{\partial q^i}\, dc^\alpha \otimes dc^\beta \otimes dq^i \tag{4.14}$$

Definition 4.4.1 *We shall refer to the function (4.14) as the* orthogonal curvature tensor of the foliation.

Lemma 4.4.2 *The function in* (4.14) *is intrinsically defined with respect to the foliation* Λ*. This means that if* $(\tilde{q^\sharp}, \tilde{q^\flat})$ *is a* Δ*-adapted chart then*

$$\frac{\partial \tilde{e}_{\alpha,\beta}}{\partial \tilde{q}^i} = \frac{\partial u^\gamma}{\partial \tilde{q}^{N+\alpha}} \frac{\partial q^{N+\delta}}{\partial \tilde{q}^{N+\beta}} \frac{\partial q^j}{\partial \tilde{q}^i} \frac{\partial e_{\gamma,\delta}}{\partial q^j}\,. \tag{4.15}$$

Proof. Since $(q^\sharp, q^\flat)$ and $(\tilde{q}, \tilde{q^\flat})$ are Δ-adapted, the coordinate transformation $(q^\sharp, q^\flat) \mapsto (\tilde{q}, \tilde{q^\flat})$ satisfies $\frac{\partial \tilde{q^\flat}}{\partial q} = 0$. Therefore,

$$\tilde{g}^{N+\alpha,N+\beta} = \frac{\partial \tilde{q}^{N+\alpha}}{\partial q^{N+\gamma}} \frac{\partial \tilde{q}^{N+\beta}}{\partial q^{N+\delta}} g^{N+\gamma,N+\delta}.$$

By inverting the matrices on both sides of the above identity one obtains

$$\tilde{e}_{\alpha,\beta} = \frac{\partial q^{N+\gamma}}{\partial \tilde{Q}^{N+\alpha}} \frac{\partial q^{N+\delta}}{\partial \tilde{q}^{N+\beta}} e_{\gamma,\delta},$$

which implies (4.15), because $q^\flat = q^\flat(\tilde{q^\flat})$ is independent of $\tilde{q^\sharp}$.
◊

Remark 4.4.3 *Although the quantity in (4.14) is not a tensor in the strict sense of the word, by (4.15) it still transforms like a tensor w.r.t. to changes of* Δ*-adapted coordinates. Hence, it is intrinsically defined in terms of the foliation. By a slight abuse of language, we thus define (4.14) as the* orthogonal curvature tensor of the foliation Λ*.*

According to Theorem 4.3.5, the foliation Λ is N-fit for hyperimpulses if and only if the the corresponding orthogonal curvature is identically equal to zero. We now give a geometric construction which clarifies the meaning of the coefficients $\partial e_{\alpha,\beta}/\partial q^i$ in (4.14), in the general case (see Figure 4.4).

In the following, given a tangent vector $\mathbf{V} \in T_{(\bar{\mathbf{q}},\bar{\mathbf{u}})}(\mathcal{Q} \times \mathcal{U})$, we denote by $\tau \mapsto \gamma_{\mathbf{V}}(\tau)$ the geodesic curve starting from $(\bar{\mathbf{q}}, \bar{\mathbf{u}})$ with velocity $\mathbf{V}$. In other words,

$$\gamma_{\mathbf{V}}(0) = (\bar{\mathbf{q}}, \bar{\mathbf{u}})\,, \qquad\qquad \frac{d\gamma_{\mathbf{V}}}{d\tau}(0) = \mathbf{V}\,.$$

The exponential map is then defined by setting

$$Exp_{(\bar{\mathbf{q}},\bar{\mathbf{u}})}(\mathbf{V}) \doteq \gamma_{\mathbf{V}}(1).$$

This is well defined for all vectors $\mathbf{V}$ in a neighborhood of the origin.

Fix any point $\mathbf{q} \in \mathcal{Q}$ and consider any non-zero vector $\mathbf{V} \in \Delta^{\perp}_{(\mathbf{q})}$. Construct the geodesic curve that originates at $\mathbf{q}$ with speed $\mathbf{V}$, namely

$$s \mapsto \gamma_{\mathbf{V}}(s) \doteq Exp_{\mathbf{q}}(s\mathbf{V})\,. \tag{4.16}$$

Next, for each $s \neq 0$, consider the orthogonal space $\Delta^{\perp}_{\mathbf{q}_s}$ at the point $\mathbf{q}_s = \gamma_{\mathbf{V}}(s)$. Assuming that s is sufficiently small, a transversality argument yields the existence of a unique vector $\mathbf{W} \in \Delta^{\perp}_{\mathbf{q}_s}$ such that

$$Exp_{\mathbf{q}_s}\mathbf{W} \;=\; \bar{\mathbf{q}}_s \;\in\; \pi^{-1}(\pi(\mathbf{q}))\,. \tag{4.17}$$

In other words, we are moving back to a point $\bar{\mathbf{q}}_s$ on the original leaf $\pi^{-1}(\pi(\mathbf{q}))$, following a second geodesic curve that is orthogonal to $\pi^{-1}(\pi(\mathbf{q}_s))$. In general, $\bar{\mathbf{q}}_s \neq \mathbf{q}$. We claim that, setting $\sigma \doteq s^2$, the map

$$\sigma \mapsto \bar{\mathbf{q}}_{\sqrt{\sigma}}$$

defines a unique tangent vector $\mathbf{Z}(\mathbf{V}) \in T_{\mathbf{q}}$. Moreover, the map $\mathbf{V} \mapsto \mathbf{Z}(\mathbf{V})$ is a homogeneous quadratic map from $\Delta^{\perp}_{\mathbf{q}}$ into the tangent space $T_{\mathbf{q}}\mathcal{Q} \subset T_{\mathbf{q}}(\mathcal{Q})$. In turn, this determines a unique symmetric bilinear mapping $B : \Delta^{\perp}_{\mathbf{q}} \otimes \Delta^{\perp}_{\mathbf{q}} \mapsto T_{\mathbf{q}}\mathcal{Q}$ such that $B(\mathbf{V},\mathbf{V}) = \mathbf{Z}(\mathbf{V})$, namely

$$B(\mathbf{V}_1,\mathbf{V}_2) \doteq \frac{1}{4}\mathbf{Z}(\mathbf{V}_1+\mathbf{V}_2) - \frac{1}{4}\mathbf{Z}(\mathbf{V}_1-\mathbf{V}_2)\,. \tag{4.18}$$

The relation between the bilinear mapping (4.18) and the curvature tensor (4.14) can be best analyzed by using coordinates. Consider an orthonormal basis $(J_1,\ldots,J_M)$ of $\Delta^{\perp}_{\mathbf{q}}$, together with local "$\mathcal{U}$-orthonormal coordinates" $(q^{\sharp}, q^{\flat})$, constructed as [11]. If $\mathbf{V} = w_1 J_1 + \cdots + w_M J_M$, then by construction the point $(\mathbf{q}_s, \mathbf{u}_s)$ has coordinates $(0, sw) = (0,\ldots,0,sw_1,\ldots,sw_M)$. Let $(q^1_w(s),\ldots,q^N_w(s),0,\ldots,0)$ be the coordinates of the point $\bar{\mathbf{q}}_s$, (constructed as in [11]). We have:

Theorem 4.4.4 *The curve $s \mapsto q_w(s) \in I\!R^N$ is continuous and satisfies*

$$\lim_{s\to 0}\frac{\bar{q}^i_w(s)}{s^2} = \frac{1}{2}\sum_{\alpha,\beta=1}^{M}\frac{\partial e_{\alpha,\beta}}{\partial q^i}w^{\alpha}w^{\beta} \qquad\qquad i = 1,\ldots,N\,. \tag{4.19}$$

♢

We refer to [11] for the proof of this result, which is based on the construction of *$\mathcal{U}$-orthonormal coordinates.*

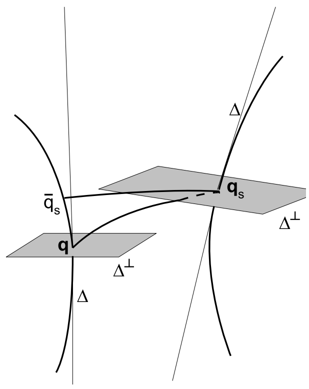

Figure 4.4: The geodesics involved in the computation of the orthogonal curvature.

Chapter 5

Control of non-holonomic systems

We consider a mechanical system as in Chapter 4 with, in addition, a *non holonomic constraint*. This means that, at each point $\mathbf{q} \in \mathcal{Q}$, velocities must belong to a given subset of the tangent space. We shall consider "linear" non holonomic constraints, which means that at each $\mathbf{q} \in \mathcal{Q}$ the velocities have to lie in a subspace $\Gamma_{\mathbf{q}} \subseteq T_{\mathbf{q}}\mathcal{Q}$. The fact that the constraint is *non holonomic* means that the *distribution* $\mathbf{q} \to \Gamma_{\mathbf{q}}$ is not integrable: there is no foliation whose leaves have the Γ_q as tangent spaces. (see Appendix A.) In other words, the constraints on the velocities cannot be deduced by differentiations from configuration constraints.

5.1 Non-holonomic systems with active constraints as controls

Let N, M, ν be positive integers such that $\nu \leq N$. Let $\mathcal{Q}$ be an $(N+M)$-dimensional differential manifold, which will be regarded as the space of configurations of a mechanical system.

Let Γ be a distribution on $\mathcal{Q}$, i.e. a vector sub-bundle of the tangent space $T\mathcal{Q}$. Throughout the following, we consider trajectories $t \mapsto \mathbf{q}(t) \in \mathcal{Q}$ of the mechanical system which are continuously differentiable and satisfy the geometric constraint

$$\dot{\mathbf{q}}(t) \ \in \ \Gamma_{\mathbf{q}(t)} \, . \tag{5.1}$$

We do not assume Γ to be integrable, so that (5.1) is a *non holonomic constraint*.

In our model, the system will be controlled by means of an active (holonomic, time-dependent) constraint. To describe this constraint, let an M-dimensional differential manifold $\mathcal{U}$ be given, together with a submersion

$$\pi : \mathcal{Q} \mapsto \mathcal{U}. \tag{5.2}$$

The fibers $\pi^{-1}(\mathbf{u}) \subset \mathcal{Q}$ will be regarded as the states of the active constraint. The set of all these fibers can be identified with the *control manifold* $\mathcal{U}$.

Let $\mathcal{I}$ be a time interval and let $\mathbf{u} : \mathcal{I} \mapsto \mathcal{U}$ be a continuously differentiable map. We say that *a trajectory* $\mathbf{q} : \mathcal{I} \to \mathcal{Q}$ *agrees with the control* $\mathbf{u}(\cdot)$ *if*

$$\pi \circ \mathbf{q}(t) \;=\; \mathbf{u}(t) \qquad\qquad \forall t \in \mathcal{I}\,. \tag{5.3}$$

For each $\mathbf{q} \in \mathcal{Q}$, consider the subspace of the tangent space at $\mathbf{q}$ given by

$$\Delta_{\mathbf{q}} \;\doteq\; ker\ T_{\mathbf{q}}\pi\,.$$

Here $T_{\mathbf{q}}\pi$ denotes the linear tangent map between the tangent spaces $T_{\mathbf{q}}\mathcal{Q}$ and $T_{\pi(\mathbf{q})}\mathcal{U}$. Clearly, Δ is the holonomic distribution whose integral manifolds are precisely the fibers $\pi^{-1}(\mathbf{u})$.

5.1.1 General setting

1) The manifold $\mathcal{Q}$ is endowed with a Riemannian metric $\mathbf{g} = \mathbf{g}_{\mathbf{q}}[\cdot,\cdot]$, the so-called *kinetic metric*, which defines the kinetic energy $\mathcal{T}$. More precisely, for each $\mathbf{q} \in \mathcal{Q}$ and $\mathbf{v} \in T_{\mathbf{q}}\mathcal{Q}$ one has

$$\mathcal{T}(\mathbf{q},\mathbf{v}) \;\doteq\; \frac{1}{2}\mathbf{g}_{\mathbf{q}}[\mathbf{v},\mathbf{v}]\,. \tag{5.4}$$

We shall use the notation $\mathbf{v} \mapsto \mathbf{g}_{\mathbf{q}}(\mathbf{v})$ to denote the isomorphism from $T_{\mathbf{q}}\mathcal{Q}$ to $T^*_{\mathbf{q}}\mathcal{Q}$ induced by the scalar product $\mathbf{g}_{\mathbf{q}}[\cdot,\cdot]$. Namely, for every $\mathbf{v} \in T_{\mathbf{q}}\mathcal{Q}$, the 1-form $\mathbf{g}_{\mathbf{q}}(\mathbf{v})$ is defined by

$$\langle \mathbf{g}_{\mathbf{q}}(\mathbf{v}), \mathbf{w}\rangle \;\doteq\; \mathbf{g}_{\mathbf{q}}[\mathbf{v},\mathbf{w}] \qquad \forall \mathbf{w} \in T_{\mathbf{q}}\mathcal{Q}\,, \tag{5.5}$$

where $\langle\cdot,\cdot\rangle$ is the duality between the tangent space $T_{\mathbf{q}}\mathcal{Q}$ and the cotangent space $T^*_{\mathbf{q}}\mathcal{Q}$.

If $\mathbf{q} \in \mathcal{Q}$ and $W \subset T_{\mathbf{q}}$, $W^{\perp}$ denotes the subspace of $T_{\mathbf{q}}$ consisting of all vectors that are perpendicular to every vector in W:

$$W^{\perp} \;\doteq\; \{\mathbf{v} \in T_{\mathbf{q}} \quad | \quad \mathbf{g}_{\mathbf{q}}[\mathbf{v},\mathbf{w}] = 0 \quad \forall \mathbf{w} \in W\}\,.$$

For a given distribution $E \subset T\mathcal{Q}$, the *orthogonal distribution* $E^\perp \subset T\mathcal{Q}$ is defined by setting $E^\perp_{\mathbf{q}} \doteq (E_{\mathbf{q}})^\perp$, for every $\mathbf{q} \in \mathcal{Q}$.

We use $\mathcal{P}^E : T_{\mathbf{q}}\mathcal{Q} \to E_{\mathbf{q}}$ to denote the *orthogonal projection* on $E_{\mathbf{q}}$. Namely, for every $\mathbf{v} \in T_{\mathbf{q}}\mathcal{Q}$, $\mathcal{P}^E(\mathbf{v})$ is the unique vector in $E_{\mathbf{q}}$ such that $\mathbf{v} - \mathcal{P}^E(\mathbf{v}) \in E^\perp_{\mathbf{q}}$.

2) Throughout the following, we shall assume that the holonomic distribution Δ and the non-holonomic distribution Γ satisfy the *tranversality condition*

$$\Delta_{\mathbf{q}} + \Gamma_{\mathbf{q}} \;=\; T_{\mathbf{q}}\mathcal{Q} \qquad\qquad \forall \mathbf{q} \in \mathcal{Q}\,. \tag{5.6}$$

Notice that this is equivalent to

$$\Delta^\perp \cap \Gamma^\perp \;=\; \{0\}\,. \tag{5.7}$$

3) The mechanical system is subject to *forces*. In the Hamiltonian formalism, these are represented by vertical vector fields on the cotangent bundle $T^*\mathcal{Q}$. We recall that, in a natural local system of coordinates (q,p), the fact that $\mathbf{F}$ is *vertical* means that its q-component is zero, namely $\mathbf{F} = \sum_{i=1}^{N+M} F_i \frac{\partial}{\partial p_i}$.

4) The constraints (5.1) and (5.3) are dynamically implemented by reaction forces obeying

D'Alembert Principle: *If $t \mapsto \mathbf{q}(t)$ is a trajectory which satisfies both (5.1) and (5.3), and $\mathbf{R}(t)$ is the constraint reaction at a time t, then*

$$\mathbf{R}(t) \;\in\; \ker\Big(\Delta_{\mathbf{q}(t)} \cap \Gamma_{\mathbf{q}(t)}\Big). \tag{5.8}$$

In other words, regarding the reaction force $\mathbf{R}(t)$ as an element of the cotangent space $T^*_{\mathbf{q}(t)}\mathcal{Q}$, one has

$$\langle \mathbf{R}(t)\,,\; \mathbf{v}\rangle \;=\; 0 \qquad\qquad \forall \mathbf{v} \;\in\; \Delta_{\mathbf{q}(t)} \cap \Gamma_{\mathbf{q}(t)} \;\subseteq\; T_{\mathbf{q}(t)}\mathcal{Q}\,. \tag{5.9}$$

5.1.2 Equations of motion

For each $\mathbf{q} \in \mathcal{Q}$, we shall use $\mathbf{g}_{\mathbf{q}}^{-1}$ to denote the inverse of the isomorphism $\mathbf{g}_{\mathbf{q}}$ at (5.5). Moreover we define the scalar product on the cotangent space $\mathbf{g}_{\mathbf{q}}^{-1}[\cdot,\cdot] :\; T^*_{\mathbf{q}}\mathcal{Q} \times T^*_{\mathbf{q}}\mathcal{Q} \;\mapsto\; I\!R$ by setting

$$\mathbf{g}_{\mathbf{q}}^{-1}[\mathbf{p}, \tilde{\mathbf{p}}] \;\doteq\; \mathbf{g}_{\mathbf{q}}\,[\mathbf{g}_{\mathbf{q}}^{-1}(\mathbf{p}),\, \mathbf{g}_{\mathbf{q}}^{-1}(\tilde{\mathbf{p}})] \qquad\qquad \forall\;\; (\mathbf{p}, \tilde{\mathbf{p}}) \;\in\; T^*_{\mathbf{q}}\mathcal{Q} \times T^*_{\mathbf{q}}\mathcal{Q}\,.$$

For every $\mathbf{q} \in \mathcal{Q}$, we shall use $\mathcal{H}(\mathbf{q}, \cdot)$ to denote the Legendre transform of the map $\mathbf{v} \to \mathcal{T}(\mathbf{q}, \mathbf{v})$, so that

$$\mathcal{H}(\mathbf{q}, \mathbf{p}) \; = \; \frac{1}{2}\mathbf{g}_{\mathbf{q}}^{-1}[\mathbf{p}, \mathbf{p}] \qquad \big(= \mathcal{T}(\mathbf{q}, \mathbf{g}^{-1}(\mathbf{p}))\big) \qquad \forall\, \mathbf{p} \in T^*_{\mathbf{q}}\mathcal{Q}\,. \tag{5.10}$$

The map $\mathcal{H} : T^*\mathcal{Q} \to I\!\!R$ will be called the *Hamiltonian corresponding to the kinetic energy* $\mathcal{T}$.

Let $(q) = (q^\sharp, q^\flat)$ be local, Δ-adapted, coordinates on $\mathcal{Q}$, and let (u) coordinates on $\mathcal{U}$, such that the domain of the chart (q) is mapped by π into the domain of the chart (u). We can assume, as usual, $q^\flat (= (q^{N+1}, \ldots, q^{N+M})) = u(= (u^1, \ldots, q^M))$. From Classical Mechanics it follows that, given a smooth control $t \mapsto u(t)$ (here regarded as a time-dependent holonomic constraint verifyng d'*Alembert Hypothethis*), the corresponding motion $t \mapsto (q, p)(t)$ on $T^*\mathcal{Q}$ verifies the relations

$$\begin{cases} \dot{q}(t) \; = \; \dfrac{\partial H}{\partial p}(q(t), p(t)) \\[2ex] \dot{p}(t) \; = \; -\dfrac{\partial H}{\partial q}(q(t), p(t)) + F(t, q(t), p(t)) + R(t) \\[2ex] q^\flat(t) \; = \; u(t) \\[2ex] p(t) \; \in \; g_{q(t)}(\Gamma_{q(t)}) \\[2ex] R(t) \; \in \; \ker \Delta_{q(t)} + \ker \Gamma_{q(t)}\,. \end{cases} \tag{5.11}$$

We have used H, F, R, and g to denote the local expressions of $\mathcal{H}$, $\mathbf{F}$, $\mathbf{R}$, and $\mathbf{g}$, respectively. For simplicity, we use the same notation for the distributions Γ, Δ on $\mathcal{Q}$ and their local expressions in coordinates. Moreover, we use the notation $v \mapsto g_q(v)$ to denote the local expression of the isomorphism $\mathbf{v} \mapsto \mathbf{g}_{\mathbf{q}}(\mathbf{v})$.

The first two equations in (5.11) are the dynamical equations written in Hamiltonian form. The first one yields the inverse of the Legendre transform. The third equation represents the (holonomic) control-constraint. Relying on the fact that π is a submersion, we shall always choose local coordinates (q^r) and (u^α) such that $q^{N+\alpha} = u^\alpha$, for all $\alpha = 1, \ldots, M$. We thus regard this equation as prescribing a priori the evolution of the last M coordinates: $q^{N+1}, \ldots, q^{N+M}$. The fourth relation in (5.11) is the Hamiltonian version of the non holonomic constraint (5.1). The fifth relation is clearly equivalent to (5.8), i.e. it represents d'Alembert's Principle.

Remark 5.1.1 Although a global, intrinsic formulation of these equations can be given (see [29]) we are here mainly interested in their local coordinate-wise expression. Indeed, a major goal of our analysis is to understand the functional dependence on the time derivative $\dot{u}$ of the control.

The special case where no active constraints are present can be obtained by taking $\Delta \equiv T^*\mathcal{Q}$, i.e. $\ker(\Delta) = \{0\}$. In this case, (5.11) reduces to the standard Hamiltonian version of the dynamical equations with non-holonomic constraints, namely

$$\begin{cases} \dot{q}(t) &= \dfrac{\partial H}{\partial p}(q(t), p(t)) \\[2ex] \dot{p}(t) &= -\dfrac{\partial H}{\partial q}(q(t), p(t)) + F(t, q(t), p(t)) + R(t) \\[2ex] p(t) &\in g_{q(t)}(\Gamma_{q(t)}) \\[2ex] R(t) &\in \ker \Gamma_{q(t)} \,, \end{cases} \tag{5.12}$$

5.2 Orthogonal decompositions of the tangent and the cotangent bundles

To derive a set of equations describing the constrained motion, it will be convenient to decompose both the tangent bundle $T\mathcal{Q}$ and the cotangent bundle $T^*\mathcal{Q}$ as direct sums of three suitable vector sub-bundles. We recall that $\mathcal{Q}$ is a manifold of dimension $N+M$, while Δ and Γ are distributions on $\mathcal{Q}$, having dimensions N and $N+M-\nu$, respectively.

5.2.1 Tangent bundle

Definition 5.2.1 *For every $\mathbf{q} \in \mathcal{Q}$, we define the following three subspaces of $T_\mathbf{q}\mathcal{Q}$.*

$$(T_\mathbf{q}\mathcal{Q})_I \doteq \Delta_\mathbf{q} \cap \Gamma_\mathbf{q} \,, \qquad (T_\mathbf{q}\mathcal{Q})_{II} \doteq \Gamma_\mathbf{q}^\perp, \qquad (T_\mathbf{q}\mathcal{Q})_{III} \doteq (\Delta_\mathbf{q} \cap \Gamma_\mathbf{q})^\perp \cap \Gamma_\mathbf{q} \,. \tag{5.13}$$

Proposition 5.2.1 *For each $\mathbf{q} \in \mathcal{Q}$, the three subspaces in (5.13) are mutually orthogonal and span the entire tangent space, namely*

$$T_\mathbf{q}\mathcal{Q} = (T_\mathbf{q}\mathcal{Q})_I \oplus (T_\mathbf{q}\mathcal{Q})_{II} \oplus (T_\mathbf{q}\mathcal{Q})_{III} \,. \tag{5.14}$$

Moreover,

$$(T_{\mathbf{q}}\mathcal{Q})_I \oplus (T_{\mathbf{q}}\mathcal{Q})_{III} \;=\; \Gamma\,. \tag{5.15}$$

If the transversality condition (5.6) holds, then the above subspaces have dimensions

$$\dim(T_{\mathbf{q}}\mathcal{Q})_I \;=\; N-\nu\,, \qquad \dim(T_{\mathbf{q}}\mathcal{Q})_{II} \;=\; \nu\,, \qquad \dim(T_{\mathbf{q}}\mathcal{Q})_{III} \;=\; M\,. \tag{5.16}$$

Proof. The orthogonality of the subspaces $(T_{\mathbf{q}}\mathcal{Q})_I$ and $(T_{\mathbf{q}}\mathcal{Q})_{II}$ is immediately clear from the definitions. Observing that

$$\Big((T_{\mathbf{q}}\mathcal{Q})_I \oplus (T_{\mathbf{q}}\mathcal{Q})_{II}\Big)^\perp \;=\; \Big((T_{\mathbf{q}}\mathcal{Q})_I\Big)^\perp \cap \Big((T_{\mathbf{q}}\mathcal{Q})_{II}\Big)^\perp \;=\; (\Delta_{\mathbf{q}} \cap \Gamma_{\mathbf{q}})^\perp \cap \Gamma_{\mathbf{q}} \;=\; (T_{\mathbf{q}}\mathcal{Q})_{III}\,,$$

we obtain the orthogonal decomposition (5.14). In particular, this implies

$$(T_{\mathbf{q}}\mathcal{Q})_I \oplus (T_{\mathbf{q}}\mathcal{Q})_{III} \;=\; \Big((T_{\mathbf{q}}\mathcal{Q})_{II}\Big)^\perp \;=\; \Gamma_{\mathbf{q}}\,.$$

Finally, if (5.6) holds, then

$$\dim(\Delta_{\mathbf{q}} \cap \Gamma_{\mathbf{q}}) \;=\; \dim(\Delta_{\mathbf{q}}) + \dim(\Gamma_{\mathbf{q}}) - \dim(T_{\mathbf{q}}\mathcal{Q}) \;=\; N + (N+M-\nu) - (N+M) \;=\; N-\nu\,.$$

Moreover, $\dim\Big((T_{\mathbf{q}}\mathcal{Q})_{II}\Big) = \dim(T_{\mathbf{q}}\mathcal{Q})$ - $\dim(\Gamma_{\mathbf{q}}) = \nu$. The last equality in (5.16) now follows from (5.14).
$\diamondsuit$

For $J \in \{I, II, III\}$, the perpendicular projection onto the subspace $(T\mathcal{Q})_J$ will be denoted by

$$\mathcal{P}_J : T_{\mathbf{q}}\mathcal{Q} \;\mapsto\; (T_{\mathbf{q}}\mathcal{Q})_J\,. \tag{5.17}$$

Remark 5.2.2 Under the assumption (5.6), the third subspace in the decomposition (5.14) can be equivalently written as

$$(T_{\mathbf{q}}\mathcal{Q})_{III} \;=\; \mathcal{P}^\Gamma(\Delta_{\mathbf{q}}^\perp).$$

Indeed, for any vector $\mathbf{v} \in T_{\mathbf{q}}\mathcal{Q}$ one has $\mathbf{v} \;\in\; \mathcal{P}^\Gamma(\Delta_{\mathbf{q}}^\perp)$ iff

$$\mathbf{v} \;\in\; \Gamma_{\mathbf{q}} \cap (\Delta_{\mathbf{q}}^\perp + \Gamma_{\mathbf{q}}^\perp) \;=\; \Gamma_{\mathbf{q}} \cap (\Delta_{\mathbf{q}} \cap \Gamma_{\mathbf{q}})^\perp.$$

5.2.2 Cotangent bundle

Thanks to the isomorphism $\mathbf{g} : T\mathcal{Q} \mapsto T^*\mathcal{Q}$ defined at (5.5), one can use (5.14) to obtain a similar decomposition of the cotangent bundle as the direct sum of three vector bundles:

$$T^*\mathcal{Q} \;=\; (T^*\mathcal{Q})_I \oplus (T^*\mathcal{Q})_{II} \oplus (T^*\mathcal{Q})_{III}, \tag{5.18}$$

where, for $J \in \{I, II, III\}$,

$$(T^*_{\mathbf{q}}\mathcal{Q})_J \;\doteq\; \mathbf{g}_{\mathbf{q}}\Big((T_{\mathbf{q}}\mathcal{Q})_J\Big). \tag{5.19}$$

We denote by $\mathcal{P}^*_J : T^*_{\mathbf{q}}\mathcal{Q} \to (T^*_{\mathbf{q}}\mathcal{Q})_J$ the perpendicular projection w.r.t. the metric $\mathbf{g}^{-1}$. The above construction yields

$$\mathcal{P}^*_J \;=\; \mathbf{g} \circ \mathcal{P}_J \circ \mathbf{g}^{-1}. \tag{5.20}$$

In view of Proposition 5.2.1 one obtains:

Proposition 5.2.2 *If the transversality condition (5.6) holds, then*

$$\dim\Big((T^*_{\mathbf{q}}\mathcal{Q})_I\Big) \;=\; N-\nu\,, \qquad \dim\Big((T^*_{\mathbf{q}}\mathcal{Q})_{II}\Big) \;=\; \nu\,, \qquad \dim\Big((T^*_{\mathbf{q}}\mathcal{Q})_{III}\Big) \;=\; M\,. \tag{5.21}$$

Moreover, for every $\mathbf{q} \in \mathcal{Q}$ the three subspaces $(T^*_{\mathbf{q}}\mathcal{Q})_I, (T^*_{\mathbf{q}}\mathcal{Q})_{II}, (T^*_{\mathbf{q}}\mathcal{Q})_{III}$ are pairwise orthogonal (w.r.t. the metric $\mathbf{g}^{-1}$). In particular, $(T^*_{\mathbf{q}}\mathcal{Q})_{II} = \ker(\Gamma_{\mathbf{q}})$.

5.3 A closed system of control equations

The following result, providing different ways to express the constraint (5.1), is straightforward.

Lemma 5.3.1 *Let $t \mapsto \mathbf{q}(t) \in \mathcal{Q}$ be a $\mathcal{C}^1$ map. Let $\mathbf{p}(t) \doteq \mathbf{g}(\dot{\mathbf{q}}(t))$. Moreover, set*

$$\dot{\mathbf{q}}_{II}(t) \doteq \mathcal{P}_{II}(\dot{\mathbf{q}}(t)), \qquad \mathbf{p}_{II} \doteq \mathcal{P}^*_{II}(\mathbf{p}(t)).$$

Then the non-holonomic constraint $\dot{\mathbf{q}}(t) \in \Gamma_{\mathbf{q}(t)}$ can be expressed in any of the following equivalent forms:

$$\mathbf{p}(t) \in \mathbf{g}(\Gamma_{\mathbf{q}(t)}) \quad\Longleftrightarrow\quad \dot{\mathbf{q}}_{II}(t) = 0 \quad\Longleftrightarrow\quad \mathbf{p}_{II}(t) = 0. \tag{5.22}$$

On a given, natural chart (q,p) on $T^*\mathcal{Q}$, and for $J \in \{I,\, II,\, III\}$, we denote by P_J^* the (q-dependent) matrix representing the projection $\mathcal{P}_J^*$. In this coordinate system, the components of the force and of the constraint reaction will be denoted as $F_J \doteq P_J^*\, F$ and $R_J \doteq P_J^*\, R$. For every $J = I, II, III$, consider the bilinear (q-dependent, possibly not symmetric), $I\!R^{N+M}$-valued map

$$(p,\tilde{p}) \;\mapsto\; \theta_J[p,\tilde{p}] \;\doteq\; \Big([\theta_J]_\ell^{r,s} p_r \tilde{p}_s\Big)_{\ell=1,\dots,N+M}, \tag{5.23}$$

where

$$[\theta_J]_\ell^{r,s} \;\doteq\; \sum_{i=1}^{N+M} \left(\frac{\partial (P_J^*)^r{}_\ell}{\partial q^i}\, g^{i,s} - \frac{1}{2}\frac{\partial g^{r,s}}{\partial q^i} (P_J^*)^i{}_\ell \right). \tag{5.24}$$

The following result will be essential in deriving the equations of motions, relative to the bundle decomposition introduced above.

Lemma 5.3.2 *Let $t \mapsto (q,p)(t)$ be a $\mathcal{C}^1$ map. For $J \in \{I, II, III\}$, define*

$$p_J(t) \doteq P_J^*\, p(t)\,, \tag{5.25}$$

so that $p(t) = p_I(t) + p_{II}(t) + p_{III}(t)$ for all t. Let $t \mapsto R(t)$ be a continuous map. Then the following statements are equivalent.

A) *The map $p(\cdot)$ satisfies the equation*

$$\dot{p}(t) \;=\; -\frac{\partial H}{\partial q}(q(t),p(t)) + F(t,q(t),p(t)) + R(t)\,. \tag{5.26}$$

B) *The maps $(p_I)(\cdot)$, $(p_{II})(\cdot)$, $(p_{III})(\cdot)$ satisfy the system*

$$\dot{p}_J(t) \;=\; \theta_J[p(t),p(t)] + F_J(t,q(t),p(t)) + R_J(t), \qquad J \in \{I,\, II,\, III\}\,. \tag{5.27}$$

Remark 5.3.1 In general $\theta_J[p,p] \notin P_J^*(I\!R^{N+M})$. Indeed, the derivative $\dot{p}_J$ may not be contained in the sub-space $P_J^*(I\!R^{N+M})$.

Proof of Lemma 5.3.2. Assume that **B)** holds. By summing the three equations in (5.27) one obtains

$$\begin{aligned}
\dot{p}(t) &= \dot{p}_I(t) + \dot{p}_{II}(t) + \dot{p}_{III}(t) \\
&= \theta_I[p(t),p(t)] + \theta_{II}[p(t),p(t)] + \theta_{III}[p(t),p(t)] + F(t,q(t),p(t)) + R(t) \\
&= (P_I^* + P_{II}^* + P_{III}^*)\left(-\frac{\partial H}{\partial q}(q(t),p(t))\right) + F(t,q(t),p(t)) + R(t) \\
&= -\frac{\partial H}{\partial q}(q(t),p(t)) + F(t,q(t),p(t)) + R(t)
\end{aligned}$$

where we have exploited the identities

$$P_I^* + P_{II}^* + P_{III}^* \; = \; Id, \qquad \frac{\partial P_I^*}{\partial q} + \frac{\partial P_{II}^*}{\partial q} + \frac{\partial P_{III}^*}{\partial q} \; = \; \frac{\partial (Id)}{\partial q} \; = \; 0.$$

This proves the validity of the differential equation (5.26).

Conversely, assume that **A)** holds. Then, for $J \in \{I, II, III\}$ one has

$$\begin{aligned} \dot{p}_J \;&=\; \frac{d}{dt}(P_J^* p) \;=\; \Big(\frac{\partial P_J^*}{\partial q} \cdot \dot{q}(t)\Big)\, p(t) + P_J^*\, \dot{p}(t) \\ &=\; \Big(\frac{\partial P_J^*}{\partial q} \cdot g^{-1}(p)\Big)\, p(t) + P_J^* \left(-\frac{1}{2}\frac{\partial g^{-1}}{\partial q}[p(t), p(t)] + F(t, q(t), p(t)) + R(t)\right) \\ &=\; \theta_J[p(t), p(t)] + F_J(t, q(t), p(t)) + R_J(t)\,. \end{aligned}$$

This proves the validity of the differential equations for p_I, p_{II}, and p_{III} in (5.27).

♢

5.3.1 The system without control-constraints

The equations of motion for a non-holonomic system with external forces can be found in several textbooks, see for example [1, 5, 4]. We review here a form of these equations, relative to decompositions (5.14) and (5.18)

Theorem 5.3.1 *Let $t \mapsto q(t)$ be a continuously differentiable path and set $p(t) \doteq g(\dot{q}(t))$. Then the path $(q, p)(\cdot)$ satisfies the non-holonomic equations* (5.12) *(with no active constraints) if and only if it is a solution of*

$$\left\{ \begin{array}{rcl} \dot{q} &=& g^{-1}(p) \\ \dot{p}_I &=& \theta_I[p, p] + F_I \\ p_{II} &=& 0 \\ \dot{p}_{III} &=& \theta_{III}[p, p] + F_{III}\,. \end{array} \right. \tag{5.28}$$

Proof. Assume that the relations (5.12) hold. For every $J \in \{I, II, III\}$ consider the projection $\dot{q}_J(t) \doteq P_J\, \dot{q}(t)$, where P_J is the coordinate representation of the projection $\mathcal{P}_J$. By Lemma 5.3.1 one has $\dot{\mathbf{q}}_{II}(t) = 0$ and hence

$\mathbf{p}_{II}(t) = 0$, for every t. Therefore, the third and the first equation in (5.28) are straightforward:

$$p_{II} \equiv 0\,, \qquad\qquad \dot q = g^{-1}(p_I + p_{III}) = g^{-1}(p).$$

Next, since $P_I^*(I\!R^{N+M}) \oplus P_{III}^*(I\!R^{N+M}) = g(\Gamma)$, the last relation in (5.12) (i.e. the d'Alembert condition) implies $P_I^*(R(t)) = P_{III}^*(R(t)) = 0$ for all t. Using the second equation in (5.12 and Lemma 5.3.2, for $J \in \{I,\, III\}$ we thus obtain

$$\dot p_J = \theta_J[p,p] + F_J(t)\,.$$

This proves the validity of the differential equations for p_I and p_{III} in (5.28).

Conversely, let the equations in (5.28) hold, with θ_I, θ_{III} defined as in (5.23)-(5.24).

The third equation in (5.12) follows from $p_{II} = 0$. In particular $p = p_I + p_{III}$. The first equation in (5.12) follows by

$$\dot q = g^{-1}(p_I + p_{III}) = g^{-1}(p) = \frac{\partial H}{\partial p}\,.$$

If the *constraint reaction* $R(t)$ is defined by

$$R(t) \doteq \dot p(t) + \frac{\partial H}{\partial q}(q(t), p(t)) - F(t, q(t), p(t)), \tag{5.29}$$

then the second equation in (5.12) is trivially satisfied.

We claim that the fourth relation in (5.12) also holds, namely $R(t) \in \ker \Gamma_{q(t)}$ for every time t. Indeed, by Lemma 5.3.2, (5.29) is equivalent to

$$\begin{cases} \dot p_I(t) = \theta_I(p(t),p(t)) + F_I(t,q(t),p(t)) + R_I(t) \\ \dot p_{II}(t) = \theta_{II}(p(t),p(t)) + F_{II}(t,q(t),p(t)) + R_{II}(t) \\ \dot p_{III}(t) = \theta_{III}(p(t),p(t)) + F_{III}(t,q(t),p(t)) + R_{III}(t)\,. \end{cases}$$

Hence, by the second and third equations in (5.28) one has $R_I(t) = R_{III}(t) = 0$. This implies $R(t) \in P_{II}^*(I\!R^{N+M}) = \ker \Gamma_{q(t)}$.

$\diamondsuit$

5.3.2 The non-holonomic system with control-constraints

In this section we write the equations of motion controlled by means of active constraints (5.11) in an equivalent form, as a closed system of controlled differential equations, supplemented by algebraic relations.

Since $\pi : \mathcal{Q} \mapsto \mathcal{U}$ is a submersion, the transversality assumption (5.6) implies that, for every $\mathbf{q} \in \mathcal{Q}$, the restriction of the tangent map $T_{\mathbf{q}}\pi$ to the subspace $(T_{\mathbf{q}}\mathcal{Q})_{III}$ is an isomorphism from $(T_{\mathbf{q}}\mathcal{Q})_{III}$ onto $T\mathcal{U}_{\pi(\mathbf{q})}$. We shall use $\mathbf{h}$ to denote the inverse of this isomorphism, and define $\mathbf{k} \doteq \mathbf{g} \circ \mathbf{h}$.

We now choose local coordinates (q, p) on $T^*\mathcal{Q}$, together with coordinates u on $\mathcal{U}$, such that

$$q^{N+1} = u^1, \quad \dots \quad , q^{N+M} = u^M. \tag{5.30}$$

The expressions of $\mathbf{h}$ and $\mathbf{k}$ in local coordinates will be denoted by h and k, respectively. We recall that the bilinear map θ_I was defined at (5.23)-(5.24). In the following, it is understood that the maps u, q should take values within the domains of the respective charts.

Theorem 5.3.2 *Let $\mathcal{I}$ be a time interval and let $u : \mathcal{I} \to I\!R^M$ be a $\mathcal{C}^1$ control function. Let $q : \mathcal{I} \mapsto I\!R^N$ be a $\mathcal{C}^1$ path and set $p(t) \doteq g(\dot{q}(t))$. Then the path $(q, p) : \mathcal{I} \to T^*\mathcal{Q}$ satisfies the non holonomic equations (5.11) (with constraints as controls) if and only if the following two conditions are satisfied.*

(i) Setting $p_J(t) \doteq P_J^*(p(t))$, for every $t \in \mathcal{I}$ one has

$$\begin{cases} \dot{q} = g^{-1}(p_I) + h(\dot{u}) \\ \dot{p}_I = \theta_I[p_I + k(\dot{u}) \,,\ p_I + k(\dot{u})] + F_I \\ p_{II} = 0 \\ p_{III} = k(\dot{u})\,. \end{cases} \tag{5.31}$$

(ii) At some time $t_0 \in \mathcal{I}$ one has $q^{N+\alpha}(t_0) = u^\alpha(t_0)$ for all $\alpha = 1, \dots, M$.

Proof. 1. Assume that all relations in (5.11) are satisfied. By (5.22) the fourth relation in (5.11) implies $p_{II} = 0$. Using the third equation in

(5.11) we obtain

$$\begin{cases} p_{III} = g(\dot q_{III}) = g(h(\dot u)) = k(\dot u) \\ p = p_I + p_{III} = p_I + k(\dot u) \\ \dot q(t) = g^{-1}(p) = g^{-1}(p_I + k(\dot u)) = g^{-1}(p_I) + h(\dot u)\,. \end{cases} \tag{5.32}$$

Finally, the fifth relation in (5.11), i.e. the d'Alembert condition, implies $P_I^*\, R(t) = 0$. Hence, by the second equation on (5.11) and Lemma 5.3.2, we obtain

$$\dot p_I = \frac{d}{dt}\Big(P_I^*\, p\Big) = \theta_I[p,p] + F_I = \theta_I[p_I + k(\dot u),\, p_I + k(\dot u)] + F_I\,.$$

This yields the second equation in (5.31).

2. Viceversa, assume that all the relations in (5.31) hold, and let us set $p \doteq p_I + k(\dot u)$. By (5.22), the equation $p_{II} = 0$ implies the fourth relation in (5.11). Moreover, from the first equation in (5.31) and the identities in (5.32) we obtain

$$\dot q = g^{-1}(p_I) + h(\dot u) = g^{-1}(p) = \frac{\partial H}{\partial p}(q,p)\,.$$

By defining

$$R(t) \doteq \dot p + \frac{\partial H}{\partial q}(q,p) - F(t, q(t), p(t)), \tag{5.33}$$

the second equation in (5.11) is clearly satisfied. It remains to check the fifth equation in (5.11), namely $R(t) \in \ker \Delta_q(t) + \ker \Gamma_q(t)$. Since by construction $\ker \Delta + \ker \Gamma = (T^*\mathcal{Q})_{II} + (T^*\mathcal{Q})_{III}$, this is equivalent to proving that $R_I = 0$. By Lemma 5.3.2, (5.33) is equivalent to the three equalities

$$\begin{cases} \dot p_I = \theta_I[p,p] + F_I + R_I \\ \dot p_{II} = \theta_{II}[p,p] + F_{II} + R_{II} \\ \dot p_{III} = \theta_{III}[p,p] + F_{II} + R_{III}\,. \end{cases} \tag{5.34}$$

By subtracting the second equation in (5.31) from the first identity in (5.34) we obtain $R_I(t) = 0$.

◇

Remark 5.3.2 In order to express more clearly the functional dependence on $\dot{u}$, let us write the equations (5.31) in the form

$$\begin{cases} \dot{q} \;=\; g^{-1}(p_I) + h(\dot{u}) \\ \dot{p}_I \;=\; \theta_I[p_I\,,\; p_I] + \Upsilon[p_I\,, \dot{u}] + \Psi[\dot{u}, \dot{u}] + F_I \\ p_{II} \;=\; 0 \\ p_{III} \;=\; k(\dot{u})\,, \end{cases} \tag{5.35}$$

where

$$\Upsilon[p_I\,,\, k(\dot{u})] \;\doteq\; \theta_I[p_I\,, k(\dot{u})] + \theta_I[k(\dot{u})\,,\, p_I], \qquad \Psi[\dot{u},\, \dot{u}] \;\doteq\; \theta_I[k(\dot{u})\,, k(\dot{u})]. \tag{5.36}$$

5.3.3 The case without non-holonomic constraints

When there is no presence of non-holonomic constraints, that is, when $\Gamma_{\mathbf{q}} = T\mathcal{Q}$ for all $\mathbf{q} \in \mathcal{Q}$, the equations in (5.35) should reduce to (4.9), the controls system we have introduced in Chapter 4. In fact, we had not proved the validity of those equations and we have referred to the fact that they could be obtained as a degenerate case ($\nu = 0$) of the case with non holonomic constraints. We leave to the reader the explicit deduction of (4.9) from (5.35). While this is a bit laborious, we point out that it reduces to elementary steps based on basic matrix computations and on the following crucial issues:

1. when $\nu = 0$, the decomposition $T^*\mathcal{Q} = (T^*\mathcal{Q})_I \oplus (T^*\mathcal{Q})_{II} \oplus (T^*\mathcal{Q})_{III}$ reduces to $(T^*\mathcal{Q})_I = \mathbf{g}(\Delta)$, $(T^*\mathcal{Q})_{II} = \{0\}$, $(T^*\mathcal{Q})_{III} = \mathbf{g}(\Delta^\perp)$, $\mathbf{g}(\Delta) \oplus \mathbf{g}(\Delta^\perp)$.

2. In any natural system of coordinates on $T^*\mathcal{Q}$ connected with Δ-adapted coordinates $(q^\sharp, q^\flat)$, the projection $\mathcal{P}_I^*$ reduces to the projection on the first n-dimensional component, namely it is given by the $(N+M) \times (N+M)$ *constant matrix* where upper left $N \times N$ minor coincides with the unit matrix and the other digits are zero.

3. The space $\Delta_{\mathbf{q}}^\perp$ is generated my the last M columns of the matrix g representing the kinetic metric $\mathbf{g}$.

Remark 5.3.3 Notice that the simplifying issue in passing from the case with non holonomic constraints to that without such constraints relies just on the fact that (in all Δ- adapted coordinates) *the projection* $\mathcal{P}_I^*$ *is constant*. Hence, the only possible occurrence of the $\dot{u}$-quadratic term in the control equations is due to the non vanishing of the orthogonal curvature introduced in Chapter 4. On the contrary, in the general case, *an extra occurrence of the quadratic term* is due to the fact that in general the projection $\mathcal{P}_I^*$ *is not constant*, which is indeed a manifestation of the non-holonomy of the additional constraint $\dot{q} \in \Gamma_{\mathbf{q}}$.

5.4 A control system with a reduced number of equations

Let $q = (q^1, \ldots, q^{N+M})$ be Δ-adapted coordinates on a open set $\emptyset \subseteq \mathcal{Q}$, so that

$$\Delta \;=\; span\Big\{\frac{\partial}{\partial q^1}\,,\,\ldots\,,\,\frac{\partial}{\partial q^N}\Big\}.$$

The main goal of this section is to write the equations of motion in terms of these state coordinates, together with additional coordinates $\xi_1, \ldots, \xi_{N+M}$ corresponding to suitable bases of the cotangent bundle $T^*\mathcal{Q}$, decomposed as in (5.18). It will turn out that the number of control equations can be reduced to $2N - \nu$. In fact, the relevant equations will involve only the variables $q^1, \ldots, q^N, \xi_1, \ldots, \xi_{N-\nu}$.

Consider a family $\{\mathcal{V}_1, \ldots, \mathcal{V}_{N+M}\}$ of smooth, linearly independent vector fields on $\mathcal{Q}$, such that

$$\begin{aligned}
(T_{\mathbf{q}}\mathcal{Q})_I &= span\Big\{\mathcal{V}_1(\mathbf{q}), \ldots, \mathcal{V}_{N-\nu}(\mathbf{q})\Big\}\,,\\
(T_{\mathbf{q}}\mathcal{Q})_{II} &= span\Big\{\mathcal{V}_{N-\nu+1}(\mathbf{q}), \ldots, \mathcal{V}_{N}(\mathbf{q})\Big\}\,,\\
(T_{\mathbf{q}}\mathcal{Q})_{III} &= span\Big\{\mathcal{V}_{N+1}(\mathbf{q}), \ldots, \mathcal{V}_{N+M}(\mathbf{q})\Big\}.
\end{aligned} \tag{5.37}$$

Throughout the following, we assume that the vectors $\{\mathcal{V}_1, \ldots, \mathcal{V}_{N-\nu}\}$, which generate $(T_{\mathbf{q}}\mathcal{Q})_I$, are orthogonal, i.e.

$$\mathbf{g}[\mathcal{V}_r, \mathcal{V}_s] \;=\; 0 \qquad \forall r, s \in \{1, \ldots, N-\nu\}, \quad r \neq s\,. \tag{5.38}$$

In addition, for $i = 1, \ldots, N+M$, we define the basis of cotangent vectors

$$\mathbf{\Omega}_i \;\doteq\; \frac{\mathbf{g}(\mathcal{V}_i)}{\mathbf{g}[\mathcal{V}_i,\, \mathcal{V}_i]}, \tag{5.39}$$

By (5.19), this yields

$$(T^*_{\mathbf{q}}\mathcal{Q})_I \;=\; span\Big\{\mathbf{\Omega}_1(\mathbf{q}),\ldots,\mathbf{\Omega}_{N-\nu}(\mathbf{q})\Big\}\,,$$

$$(T^*_{\mathbf{q}}\mathcal{Q})_{II} \;=\; span\Big\{\mathbf{\Omega}_{N-\nu+1}(\mathbf{q}),\ldots,\mathbf{\Omega}_{N}(\mathbf{q})\Big\}\,,$$

$$(T^*_{\mathbf{q}}\mathcal{Q})_{III} \;=\; span\Big\{\mathbf{\Omega}_{N+1}(\mathbf{q}),\ldots,\mathbf{\Omega}_{N+M}(\mathbf{q})\Big\}\,.$$

By (5.40) and (5.39), the differential forms $\{\mathbf{\Omega}_1,\,\ldots\,,\mathbf{\Omega}_{N-\nu}\}$ are mutually orthogonal with respect to the metric $\mathbf{g}^{-1}$, namely

$$\mathbf{g}^{-1}[\mathbf{\Omega}_r,\,\mathbf{\Omega}_s] \;=\; 0 \qquad \forall r,s\in\{1,\ldots,N-\nu\},\quad r\neq s\,. \tag{5.40}$$

Moreover, the basis $\Big\{\mathbf{\Omega}_1(\mathbf{q}),\ldots,\mathbf{\Omega}_{N-\nu}(\mathbf{q})\Big\}$ is dual to the basis $\Big\{\mathcal{V}_1(\mathbf{q}),\ldots,\mathcal{V}_{N-\nu}(\mathbf{q})\Big\}$, i.e.

$$\langle\mathbf{\Omega}_r,\,\mathcal{V}_s\rangle \;=\; \delta_{r,s} \qquad \forall r,s=1,\ldots,N-\nu.$$

This choice of orthogonal bases makes it easy to compute the projections $\mathcal{P}_I$ and $\mathcal{P}^*_I$. Indeed, for any tangent vector $\mathbf{w}$ and any cotangent vector $\mathbf{p}$ one has

$$\mathcal{P}_I(\mathbf{w}) \;=\; \sum_{\ell=1}^{N-\nu}\langle\mathbf{\Omega}_\ell,\,\mathbf{w}\rangle\,\mathcal{V}_\ell \;=\; \sum_{\ell=1}^{N-\nu}\frac{\mathbf{g}[\mathbf{w},\mathcal{V}_\ell]}{\mathbf{g}[\mathcal{V}_\ell,\mathcal{V}_\ell]}\mathcal{V}_\ell, \tag{5.41}$$

$$\mathcal{P}^*_I(\mathbf{p}) \;=\; \sum_{\ell=1}^{N-\nu}\langle\mathbf{p},\,\mathcal{V}_\ell\rangle\,\mathbf{\Omega}_\ell \;=\; \sum_{\ell=1}^{N-\nu}\frac{\mathbf{g}^{-1}[\mathbf{p},\mathbf{\Omega}_\ell]}{\mathbf{g}^{-1}[\mathbf{\Omega}_\ell,\,\mathbf{\Omega}_\ell]}\,\mathbf{\Omega}_\ell\,. \tag{5.42}$$

In the following, to simplify notation, whenever repeated indices taking values from 1 to $N+M$ are summed, the summation symbol will be omitted. On the other hand, summations ranging over a smaller set of indices will be explicitly written. Let $g=(g_{r,s})$ be the matrix representing the Riemannian metric $\mathbf{g}$ in the q-coordinates. In turn, the inverse matrix $g^{-1}=(g^{r,s})$ represents the metric $\mathbf{g}^{-1}$ on the cotangent space.

For $\ell=1,\ldots,N+M$, let ${V^1}_\ell,\ldots,{V^{N+M}}_\ell$ be the q-components of $\mathcal{V}_\ell$, so that $\mathcal{V}_\ell={V^i}_\ell\frac{\partial}{\partial q_i}$. If the (column) vector $v\in I\!R^{N+M}$ yields the coordinate representation of $\mathbf{v}$, then by (5.41) the projected vector $\mathcal{P}_I(\mathbf{v})$ has coordinates $P_I\,v$, where the $(N+M)\times(N+M)$ matrix P_I is defined by

$$(P_I)^s{}_r \;\doteq\; \sum_{\ell=1}^{N-\nu}\frac{g_{r,k}{V^k}_\ell{V^s}_\ell}{g_{i,j}{V^i}_\ell{V^j}_\ell} \qquad\qquad r,s=1,\ldots,N+M\,.$$

Similarly, let $\Omega_{r,1}, \ldots, \Omega_{r,N+M}$ be the components of Ω_r, so that $\mathbf{\Omega}_r = \Omega_{r,s} dq^s$. If the (row) vector $p \in I\!R^{M+N}$ yields the coordinate representation of the covector $\mathbf{p}$, then by (5.42) the projected the vector $\mathcal{P}_I^*(p)$ has coordinates given by $p\, P_I^*$, where the $(N+M) \times (N+M)$ matrix P_I^* is defined by

$$(P_I^*)^r{}_s \;\doteq\; \sum_{\ell=1}^{N-\nu} \frac{g^{r,k}\,\Omega_{\ell,k}\,\Omega_{\ell,s}}{g_{i,j}\,V^i{}_\ell\,V^j{}_\ell}\,. \tag{5.43}$$

In order to derive a reduced form of the system (5.31), we need to write an explicit expression of the (q-dependent) matrices h and k. Let us define the $(N+M) \times M$ matrix V_{III} and the $M \times M$ matrix $V^{III}{}_{III}$ by setting

$$V_{III} \doteq \Big(V^r{}_{N+\alpha}\Big), \qquad V^{III}{}_{III} \doteq \Big(V^{N+\beta}{}_{N+\alpha}\Big)$$

Here and in the sequel, Greek indices such as α, β range from 1 to M, while Latin indices such as r, s range from 1 to $N+M$. Recalling the identities (5.30), it is easy to check that the injective linear map $\mathbf{h} : T\mathcal{U}_{\pi(\mathbf{q})} \mapsto (T_{\mathbf{q}}\mathcal{Q})_{III} \subset T_{\mathbf{q}}\mathcal{Q}$ introduced in Section 5.3.2 is represented by the $(N+M) \times M$ matrix

$$h \;=\; V_{III} \cdot \Big(V^{III}{}_{III}\Big)^{-1}.$$

In turn, the linear map $\mathbf{k}(\cdot) \;=\; \mathbf{g}(\mathbf{h}(\cdot))$ is represented by the $(N+M) \times M$ matrix

$$k \doteq g \cdot h \;=\; g \cdot V_{III} \cdot \Big(V^{III}{}_{III}\Big)^{-1}.$$

5.4.1 Reduction of the number of equations

It clearly suffices to derive equations describing the first N components $(q^1, , \ldots, q^N)$ of q, because by (5.30) the last M components coincide with those of the control u. This already reduces the number of differential equations in (5.31) from $2N + 2M$ to $2N + M$.

Relying on the fact that the projection $\mathcal{P}_I^*$ takes values in the subbundle $(T^*\mathcal{Q})_I$, which has dimension $N - \nu$, one can reduce the number of equations to $2N - \nu$, as shown in the following theorem.

Theorem 5.4.1 *Let $(q, p, u)(\cdot)$ be as in Theorem 5.3.2. Moreover, let $(\xi,\, \eta,\, \lambda)$ be the components of $\mathbf{p} = p_i dq^i$ w.r.t. the frame $\{\mathbf{\Omega}_1, \ldots, \mathbf{\Omega}_{N+M}\}$, so that*

$$\mathbf{p}(t) \;=\; \sum_{\ell=1}^{N-\nu} \xi_\ell(t)\mathbf{\Omega}_\ell + \sum_{\ell=N-\nu+1}^{N} \eta_\ell(t)\mathbf{\Omega}_\ell + \sum_{\ell=N+1}^{N+M} \lambda_\ell(t)\mathbf{\Omega}_\ell\,.$$

Then the curve $t \mapsto (q(t), p_I(t), p_{II}(t), p_{III}(t), u(t))$ *satisfies the control system (5.35) if and only if the curve* $t \mapsto (q(t), \xi(t), \eta(t), \lambda(t), u(t))$ *is a solution of*

$$\left\{\begin{array}{rcll} \dot{q}^r &=& \sum_{\ell=1}^{N-\nu} W^r{}_\ell \xi_\ell + h^r{}_{N+\alpha}\dot{u}^\alpha & r = 1, \ldots, N \\ \dot{\xi}_m &=& \tilde{\theta}_m[\xi,\, \xi] + \tilde{\Upsilon}_m\,[\xi\,, \dot{u}] + \tilde{\Psi}_m[\dot{u}, \dot{u}] + \tilde{F}_m & m = 1, \ldots, N-\nu \\ q^{N+\alpha} &=& u^\alpha & \alpha = 1, \ldots, M \\ \eta &=& 0 & \\ \lambda &=& (V_{III})^t \cdot k \cdot \dot{u} & \end{array}\right. \tag{5.44}$$

where: i) for every $\ell = 1, \ldots, N-\nu$, *and* $r = 1, \cdots, N$ *we have set* $W^r{}_\ell \doteq V^r{}_\ell/(g_{h,k}V^h{}_\ell V^k{}_\ell)$; *ii) the superscript* ^{t} *denotes transposition; and iii) for every* $m = 1, \ldots, N-\nu$,

$$\begin{array}{c} \tilde{\theta}_m[\xi, \hat{\xi}] \doteq \sum_{a,b=1}^{N-\nu} \tilde{\theta}_m^{a,b}\, \xi_a \hat{\xi}_b, \quad \textit{with} \\ \tilde{\theta}_m^{a,b} \;\doteq\; \left([\theta_I]_i^{r,s}\Omega_{r,a}\Omega_{s,b} - \sum_{j=1}^{N} \frac{\partial \Omega_{i,a}}{\partial q^j} g^{j,s}\Omega_{s,b}\right) V^i{}_m, \\ \tilde{\Upsilon}_m\,[\xi\,, \dot{u}] \;\doteq\; \sum_{a=1}^{N-\nu} \tilde{\Upsilon}^a_{\alpha,m}\, \xi_a\, \dot{u}^\alpha, \quad \textit{with} \\ \tilde{\Upsilon}^a_{\alpha,m} \;\doteq\; \left(\Upsilon^r_{\alpha,i}\,\Omega_{r,a} - \frac{\partial \Omega_{i,a}}{\partial q^{N+\alpha}} - \sum_{j=1}^{N} \frac{\partial \Omega_{i,a}}{\partial q^j} V^i{}_m h^j{}_{N+\alpha}\right) V^i{}_m, \\ \tilde{\Psi}_m\,[w, \tilde{w}] \;\doteq\; \tilde{\Psi}_{\alpha,\beta,m}\, w^\alpha \tilde{w}^\beta, \quad \textit{with} \\ \tilde{\Psi}_{\alpha,\beta,m} \;\doteq\; \Psi_{\alpha,\beta,i} V^i{}_m\,, \qquad \tilde{F}_m \;\doteq\; F_i V^i{}_m\,. \end{array} \tag{5.45}$$

Proof. Recalling that $p_{I_i} = \sum_{\ell=1}^{N-\nu} \xi_\ell \Omega_{i,\ell}$, one obtains

$$\begin{aligned}
\dot{p}_i &= \sum_{\ell=1}^{N-\nu} \dot{\xi}_\ell \Omega_{i,\ell} + \sum_{\ell=1}^{N-\nu}\sum_{j=1}^{N} \xi_\ell \frac{\partial \Omega_{i,\ell}}{\partial q^j}\dot{q}^j \\
&= \sum_{\ell=1}^{N-\nu} \dot{\xi}_\ell \Omega_{i,\ell} + \sum_{\ell=1}^{N-\nu}\sum_{j=1}^{N} \left(\xi_\ell \frac{\partial \Omega_{i,\ell}}{\partial q^j} g^{jr} p_{I_r} + h^j_{N+\alpha}\dot{u}^\alpha\right) + \sum_{\ell=1}^{N-\nu}\sum_{\alpha=1}^{M} \xi_\ell \frac{\partial \Omega_{i,\ell}}{\partial q^{N+\alpha}}\dot{u}^\alpha \\
&= \sum_{\ell=1}^{N-\nu} \dot{\xi}_\ell \Omega_{i,\ell} + \sum_{\ell=1}^{N-\nu}\sum_{j=1}^{N} \left(\xi_\ell \frac{\partial \Omega_{i,\ell}}{\partial q^j} g^{jr} (\sum_{\ell=1}^{N-\nu} \xi_\ell \Omega_{i,\ell}) + h^j_{N+\alpha}\dot{u}^\alpha\right) + \sum_{\ell=1}^{N-\nu}\sum_{\alpha=1}^{M} \xi_\ell \frac{\partial \Omega_{i,\ell}}{\partial q^{N+\alpha}}\dot{u}^\alpha .
\end{aligned} \tag{5.46}$$

Since $\Omega_{i,r} V^i{}_m = \delta_{r,m}$, for all $r, m = 1, \ldots, N-\nu$, by (5.46) and the second equation in (5.31) one obtains the equations for the adjoint variable ξ. The other equations in (5.44) are trivially derived by substitution.

5.5 Continuity properties of the input-output map

We now examine the connections between the following properties:

i) The continuity of the input-output functional $u(\cdot) \mapsto (q(\cdot), p(\cdot))$.

ii) The vanishing of the "centrifugal" term $\Psi[\dot{u}, \dot{u}]$ in (5.36).

iii) The invariance of the distribution $\Gamma \cap (\Gamma \cap \Delta)^\perp$ in Γ-constrained inertial motions.

As in Section 2, $\mathcal{T}$ will denote the kinetic energy associated with the metric $\mathbf{g}$, defined at (5.4), while Γ denotes a non-holonomic distribution on the manifold $\mathcal{Q}$.

Definition 5.5.1 *Let $\mathcal{I}$ be a time interval. A $\mathcal{C}^2$ map $\mathbf{q} : \mathcal{I} \to \mathcal{Q}$ will be called a* free inertial motion *if, in any set of coordinates, its local expression $q(\cdot)$ provides a solution to the Euler-Lagrange equations*

$$\frac{d}{dt}\frac{\partial \mathcal{T}}{\partial \dot{q}} - \frac{\partial \mathcal{T}}{\partial q} = 0. \tag{5.47}$$

Definition 5.5.2 *A $\mathcal{C}^2$ map $q : \mathcal{I} \to \mathcal{Q}$ will be called a* Γ-constrained inertial motion *if, in any set of coordinates, its local expression $q(\cdot)$ is a solution of*

$$\frac{d}{dt}\frac{\partial \mathcal{T}}{\partial \dot{q}} - \frac{\partial \mathcal{T}}{\partial q} \in \ker \Gamma \tag{5.48}$$

and satisfies non holonomic constraints

$$\dot{q} \ \in \ \Gamma. \tag{5.49}$$

Definition 5.5.3 *Let $S \subseteq T\mathcal{Q}$ be any set. We say that S is* invariant for the free inertial flow, *(or equivalently:* inertially invariant*), if, for every free inertial motion $\mathbf{q} : \mathcal{I} \to \mathcal{Q}$ such that $(\mathbf{q}, \dot{\mathbf{q}})(t_0) \in S$ for some $t_0 \in \mathcal{I}$, one has*

$$(\mathbf{q}, \dot{\mathbf{q}})(t) \in S \qquad \forall t \in \mathcal{I}. \tag{5.50}$$

Definition 5.5.4 *Let $S \subseteq \mathcal{T}\mathcal{Q}$ be any set. We say that S is* invariant for the Γ-constrained inertial flow, *(or equivalently:* Γ-inertially invariant*), if, for every Γ-constrained inertial motion $\mathbf{q} : \mathcal{I} \to \mathcal{Q}$ such that $(\mathbf{q}, \dot{\mathbf{q}})(t_0) \in S$ for some $t_0 \in \mathcal{I}$, one has*

$$(\mathbf{q}, \dot{\mathbf{q}})(t) \in S \qquad \forall t \in \mathcal{I}. \tag{5.51}$$

Theorem 5.5.1 *The following conditions are equivalent:*

1) *If $F \equiv 0$, for every $\mathcal{C}^1$ control $u(\cdot)$ and every solution $q(\cdot)$ of the control differential equation (on $\mathcal{Q}$)*

$$\dot{q} \ = \ h(\dot{u}),$$

the map

$$t \ \mapsto \ (q(t), p(t)) \ \doteq \ \Big(q(t)\, , \, k(\dot{u}(t))\Big)$$

is a solution of (5.31)*. Equivalently, $p_I(t) = 0$ for all $t \in \mathcal{I}$.*

2) *For every local chart (q) the (vector-valued) quadratic form Ψ defined at (5.36) vanishes identically.*

3) *For every local chart (q), the (vector-valued) quadratic form $v \mapsto \theta_I[v, v]$ in* (5.23) *vanishes on the subspace $P^*_{III}(I\!R^{N+M})$.*

4) *The sub-bundle $(T\mathcal{Q})_{III} = \Gamma \cap (\Gamma \cap \Delta)^{\perp}$ is Γ-inertially invariant.*

Proof. Let a continuously differentiable control $u(\cdot)$ be given. If condition 2) holds true then, by the linearity of the equation for p_I, there is a solution $(q, p)(\cdot)$ of (5.31) such that $p_I \equiv 0$. This solution verifies the equation for q, so that $\dot{q} = h(\dot{u})$ and $p = p_{III} = g(\dot{q}) = k(\dot{u})$. Hence condition 1) is verified. Conversely, if Ψ is not identically zero, then it is not possible that for every initial condition and every control there exists a solution (q, p) such that $p_I \equiv 0$. Hence conditions 1) and 2) are equivalent.

The equivalence of 1) and 3) follows directly by the definition of θ_I and Ψ.

Finally, 1) holds true if and only if 4) is verified: indeed, on one hand it is well known that the equations (5.31) can be (locally) written in the Lagrangian form (5.48)-(5.49). On the other hand the condition $p_I = 0$ is equivalent to $\dot{q} \in \Gamma \cap (\Gamma \cap \Delta)^{\perp}$.

5.6 A geometric interpretation of the quadratic term

Let us fix $\bar{\mathbf{u}} \in \mathcal{U}$, $\bar{\mathbf{w}} \in T_{\mathbf{u}}\mathcal{U}$ and let $u : V \to I\!R^M$ be a chart defined on a neighborhood of $\bar{\mathbf{u}}$.

Let $\bar{u} = (\bar{u}^1, \ldots, \bar{u}^M)$ be the coordinates of $\bar{\mathbf{u}}$, and let the M-tuple $\bar{w} = (\bar{w}^1, \ldots, \bar{w}^M)$ be defined by $\bar{\mathbf{w}} = \sum_{\alpha=1}^{M} \bar{w}^{\alpha} \frac{\partial}{\partial u^{\alpha}}$. Fix a map $\eta : [0,1] \to I\!R$ of class C^1 such that

$$\|\dot{\eta}\|_2 = 1, \qquad \eta(0) = \eta(1) = 0, \qquad \dot{\eta}(0+) = \dot{\eta}(1-) \tag{5.52}$$

Finally, let the family of curves

$$\{\mathbf{u}_\epsilon : [0,\epsilon] \to \mathcal{U} \quad \epsilon \in [0,1]\}$$

be such that for every $\epsilon \in]0,1]$ (the image $\mathbf{u}_\epsilon([0,1])$ is entirely included in the coordinates' domain V and) the local expression $u_\epsilon(\cdot)$ of $\mathbf{u}_\epsilon$ is given by :

$$u_\epsilon(t) \doteq \bar{u} + \epsilon\eta(t/\epsilon)w \qquad \forall t \in [0,1] \tag{5.53}$$

Notice that

$$\dot{\mathbf{u}}_\epsilon(0) = \mathbf{w} \ \forall \epsilon \in]0,1], \qquad \int_0^\epsilon u_\epsilon(s)ds = 0, \qquad \int_0^\epsilon \dot{\eta}^2(t/\epsilon)dt = \epsilon \tag{5.54}$$

Now choose $\bar{\mathbf{q}} \in \mathcal{Q}$ such that $\bar{\mathbf{u}} = \pi(\bar{\mathbf{q}})$. For every $\epsilon \in]0,1]$, let us consider the solution $(\mathbf{q}_\epsilon, \mathbf{p}_\epsilon) : [0,\epsilon] \to T^*Q$ of the control system locally expressed by (5.35), with initial condition $(\mathbf{q}_\epsilon(0), \mathbf{p}_\epsilon(0)) = (\bar{\mathbf{q}}, \mathbf{g}(\mathbf{h} \cdot \bar{\mathbf{w}}))$. [1]

Let us consider the continuous curve $\mathbf{b} : [0,1] \to \mathcal{Q}$

$$\epsilon \mapsto \mathbf{b}(\epsilon) \doteq \mathbf{q}_\epsilon(\epsilon)$$

Remark 5.6.1 The image $\mathbf{b}([\mathbf{0},\mathbf{1}])$ lies on the leaf $\pi^{-1}(\bar{\mathbf{u}})$, i.e. $\pi(\mathbf{b}(\epsilon)) = \mathbf{u}$, for all $\epsilon \in [0,1]$.

[1] See Section 5.4 for the definition of the isomorphism $\mathbf{h}$.

Theorem 5.6.1 *The map* $\mathbf{b}$ *is twice differentiable at* $\epsilon = 0$ *and, for any natural system of coordinates* (q, v) *one has*

$$\frac{d\,\mathbf{b}}{d\epsilon}(0) = 0 \qquad \frac{d^2\,\mathbf{b}}{d\epsilon^2}(0) = g^{ji}\Psi_i[w, w]\frac{\partial}{\partial v_j}$$

where the Ψ_i*'s are the components of the* ($I\!R^{N+M}$*-valued) quadratic form* Ψ *defined in (5.36).*

Definition 5.6.2 *In view of Theorem 5.6.1, for every* $\mathbf{q} \in \mathcal{Q}$ *let us call the bilinear map*

$$\mathbf{w} \to \Psi_i[w, w]g^{ij}\frac{\partial}{\partial v_j} \qquad \left(= \frac{d^2\mathbf{b}}{d\epsilon^2}(0)\right)$$

the curvature of the Γ-constrained orthogonal bundle $(T_{\mathbf{q}}\mathcal{Q})_{III}$.

Proof of Theorem 5.6.1. Let local coordinates (q) and (u) be chosen near $\mathbf{q}$ and $\mathbf{u} = \pi(\mathbf{q})$, respectively. It is not restrictive to assume that $q(\mathbf{q}) = 0$, $u(\mathbf{u}) = 0$. For every $r = 1, \ldots, N + M$ let $\bar{v}^r \doteq h^r_\alpha \bar{w}^\alpha$, so that, $\mathbf{h}\bar{\mathbf{w}} = \bar{\mathbf{v}} = \bar{v}^r \frac{\partial}{\partial q^r} \in (T_{\mathbf{q}}\mathcal{Q})_{III})$. Let q_ϵ be the coordinate representation of the curve $\mathbf{q}_\epsilon$. For every $s \in [0, \epsilon]$ let us set $p_\epsilon(s) \doteq g(\dot{q}_\epsilon)$ (where the matrix g is evaluated at $(q_\epsilon(s))$), and let us notice that $(q_\epsilon, p_\epsilon)(\cdot)$ verify

$$\begin{cases} \dot{q} \;=\; p_I \cdot g^{-1}(q) + h \cdot \dot{u} \\ \dot{p}_I \;=\; \theta_I[p_I + k(\dot{u})\,,\; p_I + k(\dot{u})] \\ p_{II} \;=\; 0 \\ p_{III} \;=\; k(\dot{u}) \\ q(0) = 0 \\ (p_I, p_{II}, p_{III})(0) = (0, 0, kw). \end{cases} \tag{5.55}$$

where $p_J \doteq \mathcal{P}_J(p)$, for $J = I, II, III$. In particular, since the derivatives $\dot{u}_\epsilon$ are equi-bounded, by standard estimates it follows that the trajectories

(q_ϵ, p_ϵ) are equi-bounded as well. Therefore, since $p_I(0) = 0$,

$$q_\epsilon(\epsilon) = \int_0^\epsilon \big((p_\epsilon)_I(s) \cdot g^{-1}(q_\epsilon(s)) + h(q_\epsilon(s)) \cdot \dot{u}_\epsilon(s)\big)\, ds =$$

$$\int_0^\epsilon \Big((p_\epsilon)_I(0) \cdot g^{-1}(0) + h(0) \cdot \dot{u}_\epsilon(s) + o(\epsilon)\Big) ds = \tag{5.56}$$

$$o(\epsilon^2) + h(0) \cdot \int_0^\epsilon \dot{u}_\epsilon(s)\ ds = o(\epsilon^2)$$

and

$$\dot{q}_\epsilon(\epsilon) - v = g^{-1}\Big(q_\epsilon(\epsilon)\Big) \cdot p_\epsilon(\epsilon) - g^{-1}\Big((0)\Big) \cdot (k(0) \cdot w) =$$

$$g^{-1}\Big(q_\epsilon(\epsilon)\Big) \cdot \int_0^\epsilon \theta_I(q_\epsilon(s))\Big[p_{I\epsilon}(s) + k(q_\epsilon(s)) \cdot \dot{u}_\epsilon(s)\,,\ p_{I\epsilon}(s) + k(q_\epsilon(s)) \cdot \dot{u}_\epsilon(s)\Big] ds =$$

$$g^{-1}\,(0) \cdot \int_0^\epsilon \theta_I(0)\Big[k(0) \cdot \dot{u}_\epsilon(s)\,, k(0) \cdot \dot{u}_\epsilon(s)\Big] ds + o(\epsilon^2)\ =$$

$$g^{-1}\,(0) \cdot \theta_I(0)[k(0) \cdot w\,,\ k(0) \cdot w] \int_0^\epsilon \dot{\eta}^2(t/\epsilon)\, dt + o(\epsilon^2)$$

$$= \epsilon g^{-1}\,(0) \cdot \theta_I(0)[k(0) \cdot w\,,\ k(0) \cdot w] + o(\epsilon^2) \tag{5.57}$$

By (5.56) and (5.57), if $b(\cdot)$ is the local expression of $\mathbf{b}(\cdot)$, we get

$$\frac{db}{d\epsilon} = \lim_{\epsilon \to 0} \frac{\Big(q_\epsilon(\epsilon), \dot{q}_\epsilon(\epsilon)\Big)\ -\ (0, v)}{\epsilon} = \Big(0\ ,\ g^{-1}\,(0) \cdot \Big(\theta_I(0)\Big[k(0) \cdot w\,, k(0) \cdot w\Big]\Big)$$

$$= \Big(0\ ,\ g^{-1}\,(0) \cdot \Psi(0)\Big[w\,, w\Big]\Big) = (0, g^{1,i}\Psi_i[w, w], \ldots, g^{N+M,i}\Psi_i[w, w])$$

Chapter 6

Stabilization by control-constraints

6.1 Holonomic systems

Let us recall the control equations from Chapter 4. Let $G = (g_{r,s})_{r,s=1,\dots,N+M}$ be the matrix that represents the covariant inertial tensor in a given coordinate chart $(q^\sharp, q^\flat)$. In particular, the kinetic energy of the whole system at a state $(q^\sharp, q^\flat)$ with velocity $(v, w) \in I\!R^{N+M}$ is given by

$$\mathcal{T} = \frac{1}{2} g_{i,j}(q^\sharp, q^\flat) v^i v^j + g_{i,N+\alpha}(q^\sharp, q^\flat) v^i w^\alpha + \frac{1}{2} g_{N+\alpha,N+\beta}(q^\sharp, q^\flat) w^\alpha w^\beta .$$

Here and in the sequel, $i, j = 1, \dots, N$ while $\alpha, \beta = 1, \dots, M$. By $G^{-1} = (g^{r,s})_{r,s=1,\dots,N+M}$ we denote the inverse of G. Moreover, we consider the sub-matrices $G_1 \doteq (g_{i,j})$, $(G^{-1})_2 \doteq \left(g^{N+\alpha,N+\beta}\right)$, and $(G^{-1})_{12} \doteq \left(g^{i,N+\alpha}\right)$. Finally, we introduce the matrices

$$A = \left(a^{i,j}\right) \doteq (G_1)^{-1}, \qquad E = (e_{\alpha,\beta}) \doteq ((G^{-1})_2)^{-1}, \qquad K = \left(k^i_\alpha\right) \doteq (G^{-1})_{12} E . \tag{6.1}$$

We recall that all the above matrices depend on the variables $q = (q^\sharp, q^\flat)$. Concerning the external force, our main assumption will be

Hypothesis (A). *The force $F^{u,w}$ acting on the whole system does not explicitly depend on time,and, moreover, is affine w.r.t. the time derivative of the control, so that*

$$F^{u,\dot{u}} = F^{u,\dot{u}}(q^\sharp, p_\sharp) = F^u_0(q^\sharp, p_\sharp) + F^u_1(q, p_\sharp) \cdot \dot{u} . \tag{6.2}$$

where we have set $p_\sharp = (p_1, \dots, p_N)$.

(E.g., positional forces $F^{u,\dot u}(q) = F_0^u(q, p_\sharp)$ not necessarily conservative, verify this hypothesis. Forces which are affine functions of the velocity, which implies $F^{u,\dot u}(q^\sharp, p_\sharp) = F_0^u(q^\sharp, p_\sharp) + F_1^u(q) \cdot \dot u$ are o.k. as well.)

We can thus write the control equations in the form

$$\begin{pmatrix} \dot{q^\sharp} \\ \\ \dot{q^\flat} \\ \\ \dot{p_\sharp} \end{pmatrix} = \begin{pmatrix} A{p_\sharp}^{\mathtt{t}} \\ \\ 0 \\ \\ -\frac{1}{2} p_\sharp \frac{\partial A}{\partial q^\sharp} {p_\sharp}^{\mathtt{t}} + F_0^u \end{pmatrix} + \begin{pmatrix} K \\ \\ 1_M \\ \\ -p_\sharp \frac{\partial K}{\partial q^\sharp} + F_1^u \end{pmatrix} \dot u + \dot u^{\mathtt{t}} \begin{pmatrix} 0 \\ \\ 0 \\ \\ \frac{1}{2} \frac{\partial E}{\partial q^\sharp} \end{pmatrix} \dot u \,. \tag{6.3}$$

Our main goal is to find conditions which imply that the system (6.3) is stabilizable at a point $(\bar{q^\sharp}, \bar{q^\flat}, 0)$.

Two results concerning stabilizability will be described here. The first one relies on suitable smooth selections from the corresponding set-valued maps, as in Theorem 2.3.1. The second one is based on the use of Lyapunov functions.

For each $(q^\sharp, q^\flat)$, consider the cone

$$\Gamma(q^\sharp, q^\flat) \doteq \overline{co} \left\{ w^{\mathtt{t}} \frac{\partial E(q^\sharp, q^\flat)}{\partial q^\sharp} w \; ; \qquad w \in I\!R^M \right\}. \tag{6.4}$$

Let $\xi \in I\!R^d$ be an auxiliary control variable, ranging on a neighborhood of a point $\bar\xi \in I\!R^d$. Aiming to apply Theorem 2.3.1, let us consider a control system of the form

$$\begin{cases} \dot{q^\sharp} & = A{p_\sharp}^{\mathtt{t}} \\ \\ \dot{p_\sharp} & = F_0^{\bar u}(q^\sharp, p_\sharp) + \gamma(q^\sharp, p_\sharp, \bar u, \xi)\,, \end{cases} \tag{6.5}$$

where γ is a suitable selection from the cone Γ. It will be convenient to write (6.5) in the more compact form

$$\begin{pmatrix} \dot{q^\sharp} \\ \dot{p_\sharp}^{t} \end{pmatrix} = \Phi(q^\sharp, p_\sharp, \bar u, \xi)\,, \tag{6.6}$$

regarding $(q^\sharp, p_\sharp) \in I\!R^{N+N}$ as state variables and $\xi \in I\!R^d$ as control variable. Assume that

$$F_0^{\bar u}(\bar{q^\sharp}, 0) + \gamma(\bar{q^\sharp}, 0, \bar u, \bar\xi) = 0\,. \tag{6.7}$$

By (6.5) this implies $\Phi(\bar{q}^\sharp, 0, \bar{u}, \bar{\xi}) = 0 \in I\!R^{2N}$. To test the local controllability of (6.5) at the equilibrium point $(\bar{q}^\sharp, 0, \bar{u}, \bar{\xi})$ we look at the linearized system with constant coefficients

$$\begin{pmatrix} \dot{q}^\sharp \\ \dot{p}_\sharp{}^t \end{pmatrix} = \Lambda \begin{pmatrix} q^\sharp \\ p_\sharp{}^t \end{pmatrix} + \mathcal{B}\xi\,, \tag{6.8}$$

where

$$\Lambda = \frac{\partial \Phi}{\partial(q^\sharp, p_\sharp)} \qquad\qquad \mathcal{B} = \frac{\partial \Phi}{\partial \xi}$$

with all partial derivatives being computed at the point $(\bar{q}^\sharp, 0, \bar{u}, \bar{\xi})$.[1] We can now state

Theorem 6.1.1 *Assume that a smooth map*

$$(q^\sharp, p_\sharp, u, \xi) \mapsto \gamma(q^\sharp, p_\sharp, u, \xi) \in \Gamma(q^\sharp, q^\flat) \tag{6.9}$$

can be chosen in such a way that (6.7) holds and so that the linear system (6.8) is completely controllable. Then the system (6.3) is asymptotically stabilizable at the point $(\bar{q}^\sharp, 0, \bar{u})$.

Proof. According to Theorem 2.3.1 and Remark 2.3.2, it suffices to show that the control system

$$\begin{pmatrix} \dot{q}^\sharp \\ \\ \dot{p}_\sharp \\ \\ \dot{u} \end{pmatrix} = \begin{pmatrix} A p_\sharp{}^{\mathtt{t}} \\ \\ -\frac{1}{2} p^{\mathtt{t}} \frac{\partial A}{\partial q^\sharp} p_\sharp + F_0^u \\ \\ 0 \end{pmatrix} + \begin{pmatrix} K \\ \\ -p_\sharp{}^{\mathtt{t}} \frac{\partial K}{\partial q^\sharp} + F_1^u \\ \\ 1_M \end{pmatrix} w + \\ w^{\mathtt{t}} \begin{pmatrix} 0 \\ \\ \frac{1}{2}\frac{\partial E}{\partial q^\sharp} \\ \\ 0 \end{pmatrix} w + \begin{pmatrix} 0 \\ \\ \gamma(q^\sharp, p_\sharp, u, \xi) \\ \\ 0 \end{pmatrix} \tag{6.10}$$

is locally controllable at $(\bar{q}^\sharp, 0, \bar{u})$. Notice that in (6.10) the state variables are $q^\sharp, p_\sharp, u$, while w, ξ are the controls. Computing the Jacobian matrices

[1] Exploiting the equality $q^\flat = u$, hereon we are sometimes using the notation u instead $q^\flat$.

of partial derivatives at the point $(q^\sharp, p_\sharp, u, w, \xi) = (\bar{q^\sharp}, 0, \bar{u}, 0, \bar{\xi})$, we obtain a linear system with constant coefficients, of the form

$$\begin{pmatrix} \dot{q^\sharp} \\ \dot{p}_\sharp \\ \dot{u} \end{pmatrix} = \begin{pmatrix} \Lambda_{11} & 0 & 0 \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} q^\sharp \\ p_\sharp \\ u \end{pmatrix} + \begin{pmatrix} 0 & B_{12} \\ B_{21} & B_{22} \\ 0 & 1_M \end{pmatrix} \begin{pmatrix} \xi \\ w \end{pmatrix} \doteq$$

$$\doteq \widetilde{\Lambda} \begin{pmatrix} q^\sharp \\ p_\sharp \\ u \end{pmatrix} + \begin{pmatrix} \widetilde{\mathcal{B}_1} & \widetilde{\mathcal{B}_2} \end{pmatrix} \begin{pmatrix} \xi \\ w \end{pmatrix} \tag{6.11}$$

By assumption, the linear system (6.8) is completely controllable. Therefore

$$\text{Rank}\left[\mathcal{B},\ \Lambda\mathcal{B},\ \ldots\ ,\ \Lambda^{2N-1}\mathcal{B}\right] \ = \ 2N\,. \tag{6.12}$$

We now observe that the matrices $\Lambda, \mathcal{B}$ at (6.8) correspond to the submatrices

$$\Lambda = \begin{pmatrix} \Lambda_{11} & 0 \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix}, \qquad \mathcal{B} = \begin{pmatrix} 0 \\ B_{21} \end{pmatrix}. \tag{6.13}$$

Hence from (6.12) it follows

$$\text{span}\left[\widetilde{\mathcal{B}_1},\ \widetilde{\Lambda}\widetilde{\mathcal{B}_1},\ \ldots\ ,\ \widetilde{\Lambda}^{2N-1}\widetilde{\mathcal{B}_1}\right] \ = \ \left\{ \begin{pmatrix} q^\sharp \\ p_\sharp \\ 0 \end{pmatrix} \ ; \quad q^\sharp \in I\!R^N,\ p_\sharp \in I\!R^N \right\}. \tag{6.14}$$

Adding to this subspace the subspace generated by the columns of the matrix $\widetilde{\mathcal{B}_2}$, we obtain the entire space $I\!R^{2N+M}$. We thus conclude that the linear system (6.11) is completely controllable. In turn, this implies that the nonlinear system (6.10) is asymptotically stabilizable at $(\bar{q^\sharp}, 0, \bar{u})$, completing the proof.

◇

By choosing a special kind of selection and relying of the particular structure of (6.5), we can deduce Corollary 6.1.1 below. To state it, if k is a positive integer such that $kM \geq N$ and $W = (w_1, \ldots, w_k) \in I\!R^{M\times k}$, let us consider the $N \times kM$ matrix

$$M(u,q^\sharp,W) \doteq \begin{pmatrix} \frac{\partial e_{1,\beta}}{\partial q^1} w_1^\beta, \ldots, \frac{\partial e_{M,\beta}}{\partial q^1} w_1^\beta, & \ldots\ldots\ldots, & \frac{\partial e_{1,\beta}}{\partial q^1} w_k^\beta, \ldots, \frac{\partial e_{M,\beta}}{\partial q^1} w_k^\beta \\ & \ldots\ldots & \\ \frac{\partial e_{1,\beta}}{\partial q^N} w_1^\beta, \ldots, \frac{\partial e_{M,\beta}}{\partial q^N} w_1^\beta, & \ldots\ldots\ldots, & \frac{\partial e_{1,\beta}}{\partial q^1} w_k^\beta, \ldots, \frac{\partial e_{M,\beta}}{\partial q^N} w_k^\beta \end{pmatrix}. \tag{6.15}$$

Corollary 6.1.1 *Let k be a positive integer and assume that for a given state $(\bar{q}^\sharp, \bar{u})$ there exists a k-tuple $\bar{W} = (\bar{w}_1, \ldots, \bar{w}_k) \in (R^M)^k$ such that*

$$\text{Rank}\Big(M(\bar{u}, \bar{q}^\sharp, \bar{W})\Big) = N \tag{6.16}$$

and

$$\begin{cases} (F_0^u)^1 + \sum_{\alpha,\beta=1}^M \frac{\partial e_{\alpha,\beta}}{\partial q^1} \sum_{r=1}^k \bar{w}_r^\alpha \bar{w}_r^\beta = 0 \\ \qquad\cdots \\ \qquad\cdots \\ (F_0^u)^N + \sum_{\alpha,\beta=1}^M \frac{\partial e_{\alpha,\beta}}{\partial q^N} \sum_{r=1}^k \bar{w}_r^\alpha \bar{w}_r^\beta = 0, \end{cases} \tag{6.17}$$

where the involved functions are computed at $(q^\sharp, p_\sharp, u) = (\bar{q}^\sharp, 0, \bar{u})$. Then the system (6.3) is asymptotically stabilizable at the point $(\bar{q}^\sharp, 0, \bar{u})$.

Proof. Let us observe that the matrices Λ and $\mathcal{B}$ in (6.13) have the following form:

$$\mathcal{B} = \begin{pmatrix} 0_{N\times d} \\ \frac{\partial \gamma}{\partial \xi} \end{pmatrix} \qquad \Lambda = \begin{pmatrix} 0_{N\times N} & A \\ \frac{\partial(F+\gamma)}{\partial q^\sharp} & \frac{\partial(F+\gamma)}{\partial p_\sharp} \end{pmatrix} \tag{6.18}$$

so that, in particular,

$$\Lambda\mathcal{B} = \begin{pmatrix} A \cdot \frac{\partial\gamma}{\partial\xi} \\ \frac{\partial(F+\gamma)}{\partial p_\sharp} \cdot \frac{\partial\gamma}{\partial\xi} \end{pmatrix}. \tag{6.19}$$

Let us set $d = kM$, $\xi = W = (w_1, \ldots, w_k)$, and

$$\gamma_i(q^\sharp, u, W) \doteq \frac{1}{2} \sum_{\ell=1}^k \frac{\partial e_{\alpha,\beta}}{\partial q^i} w_\ell^\alpha w_\ell^\beta \qquad i = 1, \ldots, N.$$

Notice that, by 2-homogeneity $\gamma = (\gamma^1, \ldots, \gamma^N)$, is in fact a selection of the set-valued map Γ defined in (6.4). In view of Theorem 6.1.1, to prove the asymptotic stability it is sufficient find

$$\bar{\xi} = \bar{W}$$

such (6.17) holds and, moreover,

$$\text{Rank } [\mathcal{B}, \Lambda\mathcal{B}] (\bar{q^\sharp}, 0, \bar{u}, \bar{W}) = 2N.$$

Since A is a non-singular matrix, by (6.19) the latter condition is equivalent to

$$\text{Rank} \left(\frac{\partial \gamma}{\partial W} \right) (\bar{q^\sharp}, 0, \bar{u}, \bar{W}) \ = \ N. \tag{6.20}$$

In turn, this coincides with (6.16), so the proof is concluded.
$\diamondsuit$

We now describe a second approach, based on Corollary 2.2.12 and on the construction of a suitable, energy-like, Lyapunov function. Throughout the following we assume that the external force F in (6.2) admits the representation

$$F = F(q^\sharp, p_\sharp, u, w) = -\frac{\partial U}{\partial (q^\sharp, u)} + F_1(q^\sharp, p_\sharp, u) \cdot w \tag{6.21}$$

in terms of a potential function $U = U(q^\sharp, q^\flat)$.

Definition 6.1.1 *Given a k-tuple of vectors $W \doteq \{w_1, \ldots, w_k\} \subset I\!R^M$, the corresponding* asymptotic effective potential $(q^\sharp, u) \mapsto U_W(q^\sharp, u)$ *is defined as*

$$U_W(q^\sharp, u) \doteq U(q^\sharp, u) - \frac{1}{2} \sum_{\ell=1}^{k} w_\ell^t E(q^\sharp, u) w_\ell$$

$$\left(= U(q^\sharp, u) - \frac{1}{2} \sum_{\ell=1}^{k} \sum_{\alpha,\beta=1}^{M} e_{\alpha,\beta}(q^\sharp, u) w_\ell^\alpha \ w_\ell^\beta \right).$$

Theorem 6.1.2 . *Let the external force F have the form (6.21). For a given state $(\bar{q^\sharp}, \bar{u})$, assume that there exist a neighborhood $\mathcal{N}$ of $(\bar{q^\sharp}, \bar{u})$ and a k-tuple $W \doteq \{w_1, \ldots, w_k\} \subset I\!R^M$, as in Definition 6.1.1 which, in addition, satisfy the following property:*

There exists a continuously differentiable map $u \mapsto \beta(u)$ defined on a neighborhood of $\bar{u}$ such that the function

$$(q^\sharp, u) \mapsto U_W(q^\sharp, u) + \beta(u)$$

has a *strict local minimum* at $(q^\sharp, u) = (\bar{q^\sharp}, \bar{u})$.

Then the system (6.3) is stabilizable at $(\bar{q^\sharp}, 0, \bar{u})$.

Proof. As in Section 2.3, consider the symmetrized differential inclusion corresponding to (6.3), namely

$$\begin{pmatrix} \dot{q^\sharp} \\ \dot{p_\sharp} \\ \dot{z} \end{pmatrix} \in \overline{co} \left\{ \begin{pmatrix} A{p_\sharp}^{\mathtt{t}} \\ -\frac{1}{2}{p_\sharp}^{\mathtt{t}} \frac{\partial A}{\partial q^\sharp} p_\sharp - \frac{\partial U}{\partial q^\sharp} \\ 0 \end{pmatrix} + \ w^{\mathtt{t}} \begin{pmatrix} 0 \\ \frac{1}{2} \frac{\partial E}{\partial q^\sharp} \\ 0 \end{pmatrix} w \,, \qquad w \in I\!R^M \right\}. \tag{6.22}$$

To prove the theorem, it suffices to show that the point $(\bar{q^\sharp}, 0, \bar{u})$ is a stable equilibrium for the differential equation

$$\begin{pmatrix} \dot{q^\sharp} \\ \dot{p_\sharp} \\ \dot{u} \end{pmatrix} = \begin{pmatrix} Ap_\sharp \\ -\frac{1}{2}{p_\sharp}^{\mathtt{t}} \frac{\partial A}{\partial q^\sharp} p_\sharp - \frac{\partial U_W}{\partial q^\sharp} \\ 0 \end{pmatrix}. \tag{6.23}$$

Indeed, by the definition of U_W, the right hand side of (6.23) is a selection of the right hand-side of (6.22). Introducing the Hamiltonian function

$$H_W \doteq \frac{1}{2} pAp^{\mathtt{t}} + U_W,$$

the equation (6.23) can be written in the following Hamiltonian form:

$$\left(\dot{q^\sharp}\,,\ \dot{p_\sharp}\,,\ \dot{u}\right)^{\mathtt{t}} = \left(\frac{\partial H_W}{\partial p_\sharp}\ ,\quad -\frac{\partial H_W}{\partial q^\sharp}\ ,\ \ 0 \right). \tag{6.24}$$

Therefore the map

$$V(q^\sharp, p_\sharp, u) \doteq H_W(q^\sharp, p_\sharp, u) + \beta(u) \tag{6.25}$$

is a Lyapunov function for (6.23), from which it follows that $(\bar{q^\sharp}, 0, \bar{z})$ is a stable equilibrium for (6.23).

$\diamondsuit$

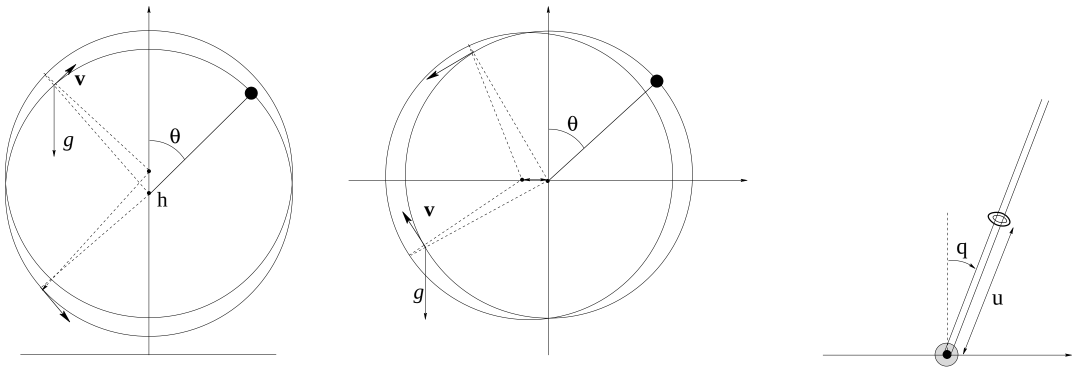

Figure 6.1: A pendulum whose pivot oscillates vertically (on the left) and horizontally (center). On the right: a bead sliding without friction along a rotating axis.

6.2 Examples

Example 1 (pendulum with oscillating pivot).[2] Let us consider a pendulum with fixed length $r = 1$, whose pivot is moving on the vertical y-axis, as shown in Figure 6.2, left. Its position is described by two variables: the clockwise angle θ formed by the pendulum with the y-axis, and the height h of the pivot. We now consider $h = u(t)$ to be our control variable, while the evolution of the other variable $\theta = q(t)$ will be determined by the equations of motion. We assume that the control function $t \mapsto u(t)$ can be assigned as a function of time, ranging over a neighborhood of the origin.

We assume that the both the pendulum and its pivot have unit mass, so that the kinetic matrix G and the matrices in (6.1) take the form

$$G = \begin{pmatrix} 1 & -\sin q \\ -\sin q & 2 \end{pmatrix} \qquad A = (1), \qquad E = (1+\cos^2 q), \qquad K = (\sin q).$$

Remark 6.2.1 To be consistent with the general theory we need to put a mass on the pivot as well. This is needed in order that the matrix G be invertible. On the other hand it is easy to show that the resulting control equations are independent of the mass of the pivot. Actually this should expected, since the motion of the pivot is here considered as a control. Of course, what is not independent of the mass of the pivot is the constraint reaction necessary to produce a given motion of u.

[2]Without danger of confusion, in these examples we shall use the notation q instead of $q^\sharp$.

Notice that orthogonal curvature of the constraint foliation Λ –i.e. the coefficient of $(\dot{u})^2$, see Section 4.4— is different from zero, for $\frac{dE}{dq} = -2\sin q \cos q$.

In the presence of downward gravitational acceleration g, the control equations for q and the corresponding momentum p is given by

$$\begin{cases} \dot{q} = p + (\sin q)\dot{u} \\ \\ \dot{p} = -\frac{\partial U}{\partial q} - p(\cos q)\dot{u} - (\sin q)(\cos q)\dot{u}^2 \,, \end{cases} \tag{6.26}$$

where $U(q,u) \doteq g \cos q$ is the gravitational potential.

Using Theorem 6.1.2, it is easy to check that this system is stabilizable at the upward equilibrium point $(\bar{q}, \bar{p}, \bar{u}) = (0,0,0)$. Indeed, choosing $W = \{w\}$ with $w > g$, the corresponding effective potential

$$U_W = g \, \cos q - \frac{1}{2}(1 + \cos^2 q)w^2 .$$

has a strict local minimum at $q = 0$.

To illustrate an application of Theorem 6.1.1, we now show that the above system is asymptotically stabilizable at every position $(\bar{q}, 0, 0)$ with $0 < |\bar{q}| < \pi/2$. To fix the ideas, assume $\bar{q} > 0$, the other case being entirely similar. For $\xi > 0$, the map $\gamma(q,p,\xi) = -\xi$ provides a smooth selection from the cone

$$\Gamma(q,u) \;\doteq\; \overline{co}\left\{ \frac{\partial E(q,u)}{\partial q} w^2 ; \quad w \in I\!R \right\} \;=\; \{-\xi ; \quad \xi \geq 0\}.$$

The corresponding system (6.5), with ξ as control variable, now takes the form

$$\begin{cases} \dot{q} = p \\ \dot{p} = g \, \sin q - \xi \,. \end{cases} \tag{6.27}$$

It is easy to check that $(\bar{q}, \bar{p}, \bar{\xi}) = (\bar{q}, 0, \; g \sin \bar{q})$ is an equilibrium position and the system is locally controllable at this point. Indeed, the linearized control system with constant coefficients is

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -g\cos\bar{q} & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \xi \,.$$

By Theorem 6.1.1, the system (6.26) is asymptotically stabilizable at $(\bar{q}, 0, 0)$.

By similar arguments one can show that, by means of horizontal oscillations of the pivot, one can stabilize the system at any position of the form $(\bar{q}, 0, 0)$, with $\frac{\pi}{2} \leq |\bar{q}| \leq \pi$.

Example 2 (sliding bead). Consider the mechanical system represented in Figure 6.2(right), consisting of a bead sliding without friction along a bar, and subject to gravity. The bar can be rotated around the origin, in a vertical plane. Calling q the distance of the bead from the origin, while u is the angle formed by the bar with the vertical line. Regarding u as the controlled variable, in this case the kinetic matrix G and the matrices in (6.1) take the form

$$G = \begin{pmatrix} 1 & 0 \\ 0 & q^2 \end{pmatrix}, \qquad A = (1), \qquad E = (q^2), \qquad K = (0).$$

The orthogonal curvature of the constraint foliation Λ is not vanishing identically: indeed, one has $\frac{dE}{dq} = 2q$. The control equations for q and the corresponding momentum p are

$$\begin{cases} \dot{q} = p \\ \dot{p} = -g \, \cos u + q\dot{u}^2 \,. \end{cases} \tag{6.28}$$

This case is more intuitive than the previous ones. Indeed, it is clear that a rapid oscillation of the angle u generates a centrifugal force that can contrast the gravitational force. More precisely, the system can be asymptotically stabilized at each $(\bar{q}, \bar{p}, \bar{u}) \in]0, +\infty[\times\{0\}\times] -\pi/2, \pi/2[$. A simple proof of this fact follows from Theorem 6.1.1. Indeed, for $q > 0$ we trivially have $\Gamma(q, u) = \{qw^2 \,; \;\; w \in I\!R\} = \{\xi \in I\!R \,; \;\; \xi \geq 0\}$. It is now clear that, if $\cos \bar{u} > 0$, then the control system

$$\begin{cases} \dot{q} = p \\ \dot{p} = -g \, \cos \bar{u} + \xi \,, \end{cases} \tag{6.29}$$

admits the equilibrium point $(\bar{q}, 0, \bar{\xi})$, with $\bar{\xi} = g \, \cos \bar{u} > 0$. Moreover, this system is completely controllable around this equilibrium point, using with controls $\xi \geq 0$. An application of Theorem 6.1.1 yields the asymptotic stability property.

We remark that the stabilizing controls cannot be independent of the position q and the velocity p. In particular, the approach in Theorem 6.1.2, based on effective potential, cannot be pursued in this case, because a constant control w cannot stabilize the system

$$\begin{cases} \dot{q} = p \\ \dot{p} = -g \cos u + qw^2. \end{cases}$$

Example 3 (double pendulum with moving pivot). So far we have considered examples with scalar controls. We wish now to study a case where the control u is two-dimensional, hence the cone (6.4) is also two-dimensional. Consider a double pendulum consisting of three point masses P_0, P_1, P_2, such that the distances $|P_0P_1|, |P_1P_2|$ are fixed, say both equal to 1. Let these points be subject to the gravitational force and constrained without friction on a vertical plane. Let (u^1, u^2) be the cartesian coordinates of the pivot P_0 , and let q^1, q^2 the clockwise angles formed by P_0P_1 and P_1P_2 with the upper vertical half lines centered in P_0 and P_1, respectively, see Figure 6.2. Because of the constraints, the state of the system $\{P_0, P_1, P_2\}$ is thus entirely described by the four coordinates (q^1, q^2, u^1, u^2). The reduced system, obtained by regarding the parameters (u^1, u^2) as controls and the coordinates (q^1, q^2) as state-coordinates, is two-dimensional. We assume that the all three points have unit mass, so that the matrix $G = (g_{rs})$ representing the kinetic energy is given by

$$G = \begin{pmatrix} 2 & \cos(q^1 - q^2) & 2\cos q^1 & -2\sin q^1 \\ \cos(q^1 - q^2) & 1 & \cos q^2 & -\sin q^2 \\ 2\cos q^1 & \cos q^2 & 3 & 0 \\ -2\sin q^1 & -\sin q^2 & 0 & 3 \end{pmatrix},$$

Moreover, recalling (6.1), we have

$$E = \begin{pmatrix} 1 - \frac{4(\sin q^1)^2}{-3+\cos 3(q^1-q^2)} & -\frac{2\sin 2q^1}{-3+\cos 3(q^1-q^2)} \\ -\frac{2\sin 2q^1}{-3+\cos 3(q^1-q^2)} & 1 - \frac{4(\sin q^1)^2}{-3+\cos 3(q^1-q^2)} \end{pmatrix},$$

$$(F_0^u)^1 = 2g\sin q^1, \qquad (F_0^u)^2 = g\sin q^2.$$

Let us observe, as in Remark 6.2.1, that the matrix E and the corresponding control equations are independent of the pivot's mass.

Proposition 6.2.1 *For every $\bar{q}^1 \in]0, \pi/4[$ (resp. $\bar{q}^1 \in]-\pi/4, 0[$) there exists $\delta > 0$ such that for all $\bar{q}^2 \in]-\delta, 0[$ (resp. $\bar{q}^2 \in]-\delta, 0[$) the system is stabilizable at $(q^1, q^2, p^1, p^2, u^1, u^2) = (\bar{q}^1, \bar{q}^2, 0, 0, 0, 0)$.*
Moreover, the system is stabilizable at $(q^1, q^2, p^1, p^2, u^1, u^2) = (0, 0, 0, 0, 0, 0)$.

Remark 6.2.2 For obvious reasons of translational invariance, if we replace $(u_1, u_2) = (0, 0)$ with any other value $(\bar{u}^1, \bar{u}^2) \in I\!R^2$ *the result holds true as well.*

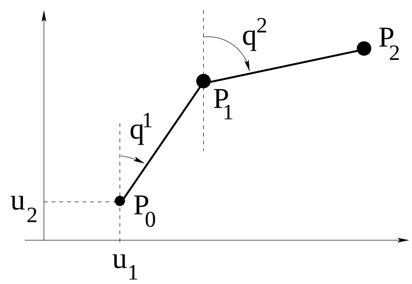

Figure 6.2: Controlling the double pendulum by moving the pivot at P_0.

Proof of Proposition 6.2.1. Using Corollary 6.1.1 with $N = M = 2$ and $k = 1$, we have that the system can be stabilized to $(\bar{q}^1, \bar{q}^2, \bar{u}^1, \bar{u}^2)$ provided there exist $\bar{w} \in I\!R^2$ such that

$$\begin{cases} 2g\sin\bar{q}^1 + \sum_{\alpha,\beta=1}^{2} \frac{\partial e_{\alpha,\beta}}{\partial \bar{q}^1}\bar{w}^\alpha \bar{w}^\beta = 0 \\ g\sin\bar{q}^2 + \sum_{\alpha,\beta=1}^{2} \frac{\partial e_{\alpha,\beta}}{\partial \bar{q}^2}\bar{w}^\alpha \bar{w}^\beta = 0 \end{cases} \tag{6.30}$$

and

$$\det \begin{pmatrix} \frac{\partial e_{1,1}}{\partial q^1}\bar{w}^1 + \frac{\partial e_{1,2}}{\partial q^1}\bar{w}^2 & \frac{\partial e_{2,1}}{\partial q^1}\bar{w}^1 + \frac{\partial e_{2,2}}{\partial q^1}\bar{w}^2 \\ \frac{\partial e_{1,1}}{\partial q^2}\bar{w}^1 + \frac{\partial e_{1,2}}{\partial q^2}\bar{w}^2 & \frac{\partial e_{2,1}}{\partial q^2}\bar{w}^1 + \frac{\partial e_{2,2}}{\partial q^2}\bar{w}^2 \end{pmatrix} \neq 0 \tag{6.31}$$

Notice that the latter relation can be written as

$$Q_{\alpha,\beta}\bar{w}^\alpha \bar{w}^\beta \neq 0 \tag{6.32}$$

where the matrix $Q = \left(Q_{\alpha,\beta}\right)$ is defined by

$$Q \doteq \frac{\partial E}{\partial q^1} \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \frac{\partial E}{\partial q^2}. \tag{6.33}$$

We recall that E denotes the matrix $(e_{\alpha,\beta})$. Moreover, it is meant that the functions in (6.30)-(6.33) are computed at $(\bar{q}^1, \bar{q}^2)$.

Let us fix $\bar{q}^1 \in]0, \pi/4[$. In order to establish the existence of a $\delta > 0$ such that for every $\bar{q}^2 \in] - \delta, 0[$ there is a $\bar{w}$ verifying the relations (6.30),(6.31), we need to study the intersections of the level sets of the quadratic forms $Q, \frac{\partial E}{\partial q^1}, \frac{\partial E}{\partial q^2}$.

Let us write the matrix $\frac{\partial E}{\partial q^1}$ and $\frac{\partial E}{\partial q^2}$ explicitly:

$$\frac{\partial E}{\partial q^1} = \begin{pmatrix} \frac{8\sin q^1\left(-3\cos q^1+\cos(q^1-2q^2)\right)}{\left(-3+\cos(2(q^1-q^2))\right)^2} & -\frac{4\left(-3\cos 2q^1+\cos 2q^2\right)}{\left(-3+\cos(2(q^1-q^2))\right)^2} \\ -\frac{4\left(-3\cos 2q^1+\cos 2q^2\right)}{\left(-3+\cos(2(q^1-q^2))\right)^2} & -\frac{8\cos q^1\left(3\sin q^1+\sin(q^1-2q^2)\right)}{\left(-3+\cos(2(q^1-q^2))\right)^2} \end{pmatrix}$$

$$\frac{\partial E}{\partial q^2} = \begin{pmatrix} \frac{8\sin^2 q^1 \sin(2(q^1-q^2))}{\left(-3+\cos(2(q^1-q^2))\right)^2} & \frac{4\sin 2q^1 \sin(2(q^1-q^2))}{\left(-3+\cos(2(q^1-q^2))\right)^2} \\ \frac{4\sin 2q^1 \sin(2(q^1-q^2))}{\left(-3+\cos(2(q^1-q^2))\right)^2} & \frac{8\cos^2 q^1 \sin(2(q^1-q^2))}{\left(-3+\cos(2(q^1-q^2))\right)^2} \end{pmatrix}$$

In particular, one has

$$\det\left(\frac{\partial E}{\partial q^1}(q^1,q^2)\right) = -\frac{16}{\left(-3+\cos(2(q^1-q^2))\right)^2} < 0\,, \qquad \det\left(\frac{\partial E}{\partial q^2}(q^1,q^2)\right) = 0$$

for all q^1, q^2.

Hence, one has:

(i) The quadratic form $w \mapsto w^{\mathsf{t}}\frac{\partial E}{\partial q^1}w$ is indefinite, so it can be factorized by two linear, independent, forms. Let us assume that, $\bar{q}^2 \in]-\bar{q}^1, 0[$, so that, in particular, $\frac{\partial e_{2,2}}{\partial q^1} < 0$. Hence, for suitable functions $a = a(q^1,q^2)$, $b = b(q^1,q^2)$ such that $a(q^1,q^2) \neq b(q^1,q^2)$ for all q^1, q^2, one has

$$\frac{\partial e_{\alpha,\beta}}{\partial q^1}w^\alpha w\beta = \frac{\partial e_{2,2}}{\partial q^1}(w^2 - aw^1)(w^2 - bw^1).$$

(ii) If $\bar{q}^2 \in]-\bar{q}^1, 0[$, the quadratic form $w \mapsto w^{\mathsf{t}}\frac{\partial E}{\partial q^2}w$ is positive semi-definite. Hence it can be factorized by the positive scalar function $\frac{\partial e_{2,2}}{\partial q^1}$ and the square of a linear function. Moreover *this linear function coincides with one of the two linear factors of the quadratic form* $w \mapsto w^{\mathsf{t}}\frac{\partial E}{\partial q^1}w$. This is a trivial consequence of the identity

$$\left(\frac{\partial e_{1,2}}{\partial q^2}\frac{\partial e_{2,2}}{\partial q^1}\right)^2 - 2\frac{\partial e_{2,1}}{\partial q^1}\frac{\partial e_{2,2}}{\partial q^1}\frac{\partial e_{1,2}}{\partial q^2}\frac{\partial e_{2,2}}{\partial q^1}\frac{\partial e_{2,2}}{\partial q^2} + \frac{\partial e_{1,1}}{\partial q^1}\frac{\partial e_{2,2}}{\partial q^1}\left(\frac{\partial e_{2,2}}{\partial q^2}\right)^2 = 0,$$

which can be verified by direct computation. Let $(w^2 - aw^1)$ be the common factor of the two quadratic forms. Hence, we obtain

$$\frac{\partial e_{\alpha,\beta}}{\partial q^2} w^\alpha w^\beta = \frac{\partial e_{2,2}}{\partial q^2}(w^2 - aw^1)^2.$$

(iii) The quadratic form $w \mapsto w^{\mathrm{t}} Q w$ is semi-definite and, at each (q^1, q^2), *it is proportional to the form* $w^{\mathrm{t}} \frac{\partial E}{\partial q^2} w$. More precisely, one has

$$Q_{\alpha,\beta} w^\alpha w^\beta = \left(\frac{\partial e_{2,2}}{\partial q^1} \cdot \frac{a-b}{2} \right) \frac{\partial e_{\alpha,\beta}}{\partial q^2} w^\alpha w \beta = \left(\frac{\partial e_{2,2}}{\partial q^1} \cdot \frac{\partial e_{2,2}}{\partial q^2} \cdot \frac{a-b}{2} \right) (w^2 - aw^1)^2.$$

This is easily deduced by (6.33). Notice, in particular, the form $Q_{\alpha,\beta} w^\alpha w^\beta$ is never equal to the null form, since $a(q^1, q^2) \neq b(q^1, q^2)$ for all q^1, q^2.

If S is a 2×2 matrix and $\rho \in I\!R$ let us set

$$\{w^{\mathrm{t}} S w = \rho\} \doteq \{w \in I\!R^2 \mid w^{\mathrm{t}} S w = \rho\}.$$

Since $w^{\mathrm{t}} \frac{\partial E}{\partial q^2} w$ is positive definite and $\sin q^2 < 0$, there exists a real number $\eta > 0$ such that

$$\left\{ w^{\mathrm{t}} \frac{\partial E}{\partial q^2} w = -\sin \bar{q}^2 v \right\} = \{w \in I\!R^2 \,:\, (w^2 - aw^1) = \eta\} \cup \{w \in I\!R^2 \,:\, (w^2 - aw^1) = -\eta\},$$

so that, in particular,

$$\left\{ w^{\mathrm{t}} \frac{\partial E}{\partial q^2} w = -g \sin \bar{q}^2 \right\} \cap \{w \in I\!R^2 \,:\, (w^2 - aw^1) = 0\} = \emptyset.$$

By (iii) this implies

$$\left\{ w^{\mathrm{t}} \frac{\partial E}{\partial q^2} w = -g \sin \bar{q}^2 \right\} \cap \{w^{\mathrm{t}} Q w = 0\} = \emptyset. \qquad (6.34)$$

Moreover, by (i) the line $\{w \in I\!R^2 \,:\, (w^2 - aw^1) = 0\}$ is asymptotic to the hyperbolic arc

$$\left\{ w^{\mathrm{t}} \frac{\partial E}{\partial q^1} w = -2g \sin \bar{q}^1 \right\},$$

which implies

$$\left\{ w^{\mathrm{t}} \frac{\partial E}{\partial q^1} w = -2g \sin \bar{q}^1 \right\} \cap \left\{ w^{\mathrm{t}} \frac{\partial E}{\partial q^2} w = -g \sin \bar{q}^2 \right\} \neq \emptyset\,. \qquad (6.35)$$

Putting (6.34) and (6.35) together, we obtain the first statement of the theorem.

On the other hand, the second statement will be proved by an application of Theorem 6.1.2. Since $U(q) = g(2\cos q^1 + \cos q^2)$ is a potential, by letting $W = \{(0, \eta)\}$ and $\beta(u) \doteq (u^1)^2 + (u^2)^2$, we have that the effective potential

$$U_W(q, u) \doteq U(q) + \eta^2 e_{2,2}(q) + \beta(u)$$

has a strict minimum at $(q, u) = (0, 0, 0, 0)$ as soon as $|\eta|$ is large enough. In view of Theorem 6.1.2, this implies the that the system is stabilizable at $(q^1, q^2, p_1, p_2, u^1, u^2) = (0, 0, 0, 0, 0, 0)$.

$\diamondsuit$

6.3 Non-holonomic systems

Let us remark a crucial difference with the case without non holonomic constraints. In that case the presence of the quadratic term is completely determined by the relation between the kinetic metrics and the distribution Δ. In particular this quadratic term accounts for the curvature of the orthogonal distribution $\Delta^\perp$ (see Chapter 4 and [11]). Instead, when the non holonomic constraint $\dot{q} \in \Gamma$ is acting on the system we may well have that the orthogonal distribution $\Delta^\perp$ has zero-curvature while the dynamical equations still contain a term which is quadratic in $\dot{\mathbf{u}}$. This is the case, for instance, of the *Roller Racer*.

6.4 An Example: the Roller Racer

The Roller Racer is a classical example of a non-holonomic system, widely investigated within the theory of the *momentum map*, see e.g. [5, 18]. It consists of two rigid planar bodies, connected at a point by a rotating joint, as shown in fig.6.4. One of the two bodies has a much larger mass than the other. Let ρ be the distance between the joint and the center of mass of the heavier body. To simplify computations we also assume that the center of mass of the lighter body coincides with the joint.

The coordinates (q^1, q^2, q^3, u) are as follows . We let $(q^3, q^1) = (x, y)$ be the Euclidean coordinates of the center of mass of the large body. Moreover, $q^2 = \theta$ is the counter-clockwise angle between the horizontal axis and the major axis of the large body, and $u = \phi$ is the counter-clockwise angle between the major axes of the two bodies.

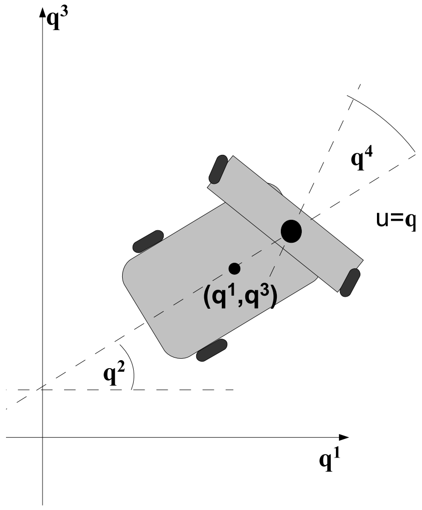

Figure 6.3: The roller racer.

The non-holonomic constraint consists in assuming that each pair of wheels rotate (without slipping) parallel to the corresponding body. This corresponds to the condition

$$(\dot{q}, \dot{u}) \;\in\; \Gamma \;\doteq\; ker(\omega^1) \cap ker(\omega^2),$$

where

$$\begin{cases} \omega^1 \doteq \cos q^2 \, dq_1 - \sin q^2 \, dq^3 \\ \omega^2 \doteq \cos(q^2 + u) \, dq_1 + \rho \, \cos u \, dq^2 - \sin(q^2 + u) \, dq^3 \, . \end{cases} \tag{6.36}$$

Γ is the (non-integrable) distribution determining the non-holonomic constraint. In other words, admissible motions $t \mapsto (q, u)(t)$ are subject to

$$\begin{aligned} \cos q^2(t) \, \dot{q}^1(t) - \sin q^2(t) \, \dot{q}^3(t) &= 0 \, , \\ \cos(q^2(t) + u(t)) \, \dot{q}^1(t) + \rho \, \cos u(t) \, \dot{q}^2(t) - \sin(q^2(t) + u(t)) \, \dot{q}^3(t) &= 0 \, . \end{aligned} \tag{6.37}$$

The coordinate $\phi = u$ is regarded as a control.

The transversality condition is trivially satisfied, because

$$\Delta^{ker} \;=\; span\{du\}, \qquad \Gamma^{ker} \;=\; span\{\omega^1, \omega^2\},$$

and $span\{\omega^1, \omega^2\} \cap span\{du\} = \{0\}$.

For simplicity, we normalize the mass of the large body, setting it $= 1$. The moment of inertia of the large body with respect to the vertical axis passing through its center of mass is denoted by I. The mass of the small body is regarded as negligible, but its moment of inertia J with respect to the vertical axis passing through the center of mass (of the small body) is assumed to be different from zero.[3] The kinetic matrix $(g_{i,j})$ and its inverse $(g^{i,j})$ are computed as

$$(g_{i,j}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & I+J & 0 & J \\ 0 & 0 & 1 & 0 \\ 0 & J & 0 & J \end{pmatrix} \qquad (g^{i,j}) = \mathbf{G}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{I} & 0 & -\frac{1}{I} \\ 0 & 0 & 1 & 0 \\ 0 & -\frac{1}{I} & 0 & \frac{I+J}{JI} \end{pmatrix}.$$

Let us start by finding a basis for the decomposition $T\mathcal{Q} = (T\mathcal{Q})_I \oplus (T\mathcal{Q})_{II} \oplus (T\mathcal{Q})_{III}$. Notice that the vectors

$$\mathbf{w}_1 = 2\rho\cos u \sin q^2 \, \frac{\partial}{\partial q^1} + \; 2\sin u \frac{\partial}{\partial q^2} + \; 2\rho \cos u \cos q^2 \frac{\partial}{\partial q^3} \text{ and } \; \mathbf{w}_2 = \frac{\partial}{\partial u}$$

form a basis for Γ. Since $\Gamma = span\{\frac{\partial}{\partial q^1}, \frac{\partial}{\partial q^2}, \frac{\partial}{\partial q^3}\}$, $\mathbf{w}_1$ is a basis for $(T_\mathcal{Q})_I = \Gamma \cap \Delta$.

It is also straightforward to verify that

$$(T\mathcal{Q})_{II} = \Gamma^\perp \;=\; \ker\{\mathbf{g}(\mathbf{w}_1), \mathbf{g}(\mathbf{w}_2)\} \;=\; span\,\{\mathbf{v}_2, \mathbf{v_3}\}\;, \qquad (6.38)$$

where

$$\mathbf{v}_2 \;=\; \frac{I\csc q^2 \tan u}{\rho}\,\frac{\partial}{\partial q^1} - \frac{\partial}{\partial q^2} + \frac{\partial}{\partial u}, \qquad \mathbf{v}_3 \;=\; -\cot q^2\,\frac{\partial}{\partial q^1} + \frac{\partial}{\partial q^3}.$$

Since $(T\mathcal{Q})_{III}$ is perpendicular to both $(T\mathcal{Q})_I \oplus (T\mathcal{Q})_{II}$ and we get

$$(T\mathcal{Q})_{III} \;=\; \ker\Big\{\mathbf{g}(\mathbf{v}_1), \mathbf{g}(\mathbf{v}_2 O), \mathbf{g}(\mathbf{w}_3)\Big\} = \; span\{\mathbf{v}_4\} \qquad (6.39)$$

where we have set

$$\mathbf{v}_4 \;\doteq\; \frac{1}{\Delta_0}\left(-\frac{1}{2}J\rho\sin q^2 \sin 2u\,\frac{\partial}{\partial q^1} \;-J\sin^2 u\,\frac{\partial}{\partial q^2} \;-\frac{1}{2}J\rho\cos q^2 \sin 2u\,\frac{\partial}{\partial q^3}\right),$$

[3] This is a standard approximation adopted in the existing literature.

$\Delta_0 \doteq \rho^2 \cos^2 u + (I+J) \sin^2 u$.

Then setting $\mathcal{V}_1 \doteq \mathbf{v}_1, \mathcal{V}_2 \doteq \mathbf{v}_2, \mathcal{V}_3 \doteq \mathbf{v}_3, \mathcal{V}^4 \doteq \mathbf{v}_4$ we have obtained a family of vector fields (on an open subset of the configuration manifold) as in (5.37).

To get the equation we need primarily the computation of the coefficients (5.45). For this purpose, let us begin with observing that the form $\mathbf{\Omega}_1$ generating $(T^*\mathcal{Q})_I$ is given by

$$\mathbf{\Omega}_1 = \Omega_{1,1} dq^1 + \Omega_{2,1} dq^2 + \Omega_{3,1} dq^3 + \Omega_{4,1} dq^4 = \mathbf{g}(\mathcal{V}_1) =$$

$$(2\rho \cos u \sin q^2) dq^1 + (2(I+J) \sin u) dq^2 + (2\rho \cos u \cos q^2) dq^3 + (2J \sin u) dq^4$$

Let us remind that the projection matrix P_I^* is computed as (see (5.43))

$$(P_I^*)^r{}_s = \sum_{k=1}^{4} \frac{g^{r,k} \Omega_{k,1} \Omega_{s,1}}{\sum_{a,b=1}^{4} g_{a,b} V^a{}_1 V^b{}_1}.$$

Let us set $\Delta_1 = I + J + \rho^2 + (-I - J + \rho^2) \cos 2u$, $\Delta_2 = (I+J) \sin q^2 + J \cos q^2 \cos u \sin u$, $\Delta_3 = -(I+J) \cos q^2 + J \cos u \sin q^2 \sin u$, $\Delta_4 = -I - 3J + 5IJ^2 + 4J^3 - \rho^2$.

Then, trivial computations make the four differential equations in (5.44) explicit:

$$\begin{cases} \dot{q}^1 = 2\rho \cos u \sin q^2 \cdot \xi - \frac{J\rho \sin q^2 \sin 2u}{2\Delta_0} \cdot \dot{u} \\ \dot{q}^2 = 2 \sin u \cdot \xi - \frac{J \sin^2 u}{\Delta_0} \cdot \dot{u} \\ \dot{q}^3 = 2\rho \cos q^2 \cos u \cdot \xi - \frac{J\rho \cos q^2 \sin 2u}{2\Delta_0} \cdot \dot{u} \\ \dot{\xi} = -\sin 2u \left(\frac{(I+J-\rho^2)}{\Delta_1} + \frac{1}{2(\rho^2/\Delta_4 + \sin^2 u)} \right) \cdot \xi \dot{u} - \frac{2J\rho^2 \cos u}{\Delta_1^2} \cdot \dot{u}^2. \end{cases} \tag{6.40}$$

Remark 6.4.1 By implementing a rapidly oscillating control like, for instance, $u_\epsilon = \bar{u} + \epsilon \sin t/\epsilon$, one can simply verify that the motion of the Roller Racer's barycenter will approximate a forward line (if $\bar{u} = 0$) or a circle (if $\bar{u} \neq 0$, $|u| < \pi/2$.

This could be easily modified into a result of stabilization, for example in the case when the Roller Racer runs on a non horizontal plane (and gravity is present).

Let us remark that the force producing the motion (under the action of the control u_ϵ) is essentially due to the quadratic term —see Chapter

2—, which, in turn, would be zero if a non holonomic constraint were not acting on the system —see also Remark 5.3.3 . Notice that this last circumstance could be regarded as modeling the Roller Racer running on a ideally frictionless plane.

Remark 6.4.2 In view of Remark 6.4.1 it would be interesting to investigate the problem of obtaining the non holonomic constraint as a "limit" of larger and larger orthogonal friction acting on the toy's wheels.

Appendix A

Basics on Differential Manifolds

A.1 Differential manifolds

Let $\mathcal{M}$ be a set, let r be a positive integer, and let $(U_i, \phi_i)_{i\in I}$ be a collection of pairs such that

- for all $i \in I$, $U_i \subset \mathcal{M}$, and $\bigcup_{i\in I} U_i = \mathcal{M}$
- for all $i \in I$, ϕ_i is a injective map from U_i into $\mathbb{R}^{n_i}$, for some positive integer n_i
- $\phi_i(U_i)$ is an open subset
- for all pairs i, j such that $U_i \cap U_j \neq \emptyset$ the map

  $$\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$$

  is a C^r diffeomorphism.

Every collection $(U_i, \phi_i)_{i\in I}$ having the properties stated above is called an *atlas* for $\mathcal{M}$, and the pairs (U_i, ϕ_i) are called *charts*.

There is a unique *topology* on $\mathcal{M}$ such that every chart's domain U_i is an open subset, and every ϕ_i is a homeomorphism from U_i onto its image $\phi(U_i)$. It is easy to check that with respect to this topology a subset $A \subset \mathcal{M}$ is *open* if for every chart (U_i, ϕ_i) such that $U_i \cap A \neq \emptyset$, the image $\phi_i(U_i \cap A)$ is an open subset of $\mathbb{R}^{n_i}$.

In principle, for $i \neq j$, n_j and n_j need not to be equal. However, they are equal if U_i and U_j lie on the same connected component:

Lemma A.1.1 *Let n be a positive integer. Then the set $C(n)$ made of point $m \in \mathcal{M}$ such that for every chart (U_i, ϕ_i) at m*[1] *one has $n_i = n$ is closed and open. In particular, on every connected component of $\mathcal{M}$ the n_i's coincide.*

PROOF. [2] For every pair of indexes i, j such that $U_i \cap U_j \neq \emptyset$ one has $n_i = n_j$. Indeed the map

$$f \doteq \phi_i \circ \phi_j^{-1} : \phi_j(U_i \cap U_j) \to \phi_i(U_i \cap U_j)$$

is a C^r-isomorphism. Let $x \in \phi_j(U_i \cap U_j)$ and let $Df(x)(\cdot)$ denote the differential of f at x. In particular, $Df(x)$ is a linear isomorphism from $I\!R^{n_j}$ onto $I\!R^{n_i}$, so $n_j = n_i$.

Let n be a positive integer. If $C(n)$ is non-empty there exists $m \in \mathcal{M}$ such that $n_i = n$ for one —hence for every— chart (U_i, ϕ_i) at m. The set $C(n)$ is clearly open: if $m \in C(n)$, and (U_i, ϕ_i) is a chart at m, then $n_i = n$. Since (U_i, ϕ_i) is also a chart at every $m' \in U_i$, $U_i \subseteq C(n)$. Let us prove that $C(n)$ is closed as well. If $m \in A \doteq \mathcal{M} \backslash C(n)$ and (U_j, ϕ_j) is a chart at m, then $U_j \subseteq A$. In particular, A is open, so $C(n)$ is closed.

Let $U \subseteq \mathcal{M}$ be an open subset, let q be a positive integer, and let $\phi : U \to I\!R^q$ be a homeomorphism from U onto its image $\phi(U) \subseteq I\!R^q$. Let $i \in I$ be such that $U \cap U_i \neq \emptyset$. The pairs (ϕ, U), (U_i, ϕ_i) are called *compatible* if $\phi \circ \phi_i^{-1}$ is a C^r-isomorphism (from $\phi_i(U \cap U_i)$ onto $\phi(U \cap U_i)$).

Two atlases on $\mathcal{M}$ $(U_i, \phi_i)_{i \in I}$, $(U_j, \phi_j)_{i \in J}$ are called *compatible* if for every $i \in I$ and every $j \in J$ the charts (U_j, ϕ_j), (U_i, ϕ_i) are *compatible.* Clearly compatibility is an equivalence relation on the set of atlases on $\mathcal{M}$. An equivalence class of atlases is a *differential structure* on $\mathcal{M}$ of class C^r.

Definition A.1.2 *A set $\mathcal{M}$ together with a differential structure of class C^r is called a* **manifold of class** C^r*. If $(U_i, \phi_i)_{i \in I}$ is an atlas belonging to the given differential structure and there is n such that $n_i = n$ for all $i \in I$ then one says that* the manifold $\mathcal{M}$ has dimension n*, or, equivalently, that $\mathcal{M}$ is a n-manifold.*[3]

---

[1] We say that a chart (U_i, ϕ_i) is *at m* if $m \in U_i$.

[2] Actually, there would be no need of the differentiable structure. Indeed, thanks to the Invariance Domain Theorem this fact is true also when the maps $\phi_i \circ \phi_j^{-1}$ are homeomorphism.

[3] It is straightforward to check that n is well defined, i.e. it does not depend on the particular atlas used to define it. Actually it is easy to show that n depends only on $\mathcal{M}$: that is, if it is defined for a C^r differential structure on $\mathcal{M}$ it is defined also for all C^r differential structures on $\mathcal{M}$.

When speaking of a manifold, we shall tacitly always make the following assumption:

- If $m_1, m_2 \in \mathcal{M}$, $m_1 \neq m_2$, there are charts (U_1, ϕ_1), (U_2, ϕ_2) such that $m_1 \in U_1$, $m_2 \in U_2$, and $U_1 \cap U_2 \neq \emptyset$.

(This makes our manifold a Hausdorff topological space.)

Remark A.1.3 Let us point out that one can have different differential C^r structures on a set $\mathcal{M}$. For instance, if $\mathcal{M} = I\!R$ then the atlas $\mathcal{A}_1$ made of the one chart $(U_1, \phi_1) = (I\!R, id)$, where id denotes the identity map gives $I\!R$ the same differential structure as the one given by the atlas $\mathcal{A}_2$ made of the unique chart $(U_2, \phi_2) = (I\!R, \arctan)$. Also, the same structure is given by the union $\mathcal{A}_1 \cup \mathcal{A}_2$. Instead, the atlas $\mathcal{A}_3$ made of the unique chart $(U_3, \phi_3(t)) = (I\!R, t^3)$, which *gives a C^∞ differential structure* is not Cq equivalent to $\mathcal{A}_1$, for the map $s \to s^{\frac{1}{3}}$ is not even differentiable at $s = 0$. Of course $\mathcal{A}_1$ and $\mathcal{A}_3$ endow $I\!R$ with the same C^0 (=topological) structure. It is easy to construct atlases which give the same C^r structure but a different $C^{r'}$ structure on a set $\mathcal{M}$. For this purpose, observe that for every positive integer n and every open subset $A \subset I\!R^n$, any atlas on A made by only one chart (U, ϕ) *gives A a C^∞ structure as soon as ϕ is a homeomorphism*! For instance the square $\mathcal{M} =]-1, 1[\times]-1, 1[$ with the chart $\phi(x, y) = (x^{\frac{1}{3}}, y^{\frac{1}{5}})$ is a C^∞ (actually, analytic !) manifold...

When no otherwise specified, by *manifold* we shall mean a differential manifold of class (at least) C^1 with dimension $n > 0$.

Definition A.1.4 *Let $\mathcal{M}$, $\mathcal{M}'$ be manifolds of dimension n and d, respectively, let s be a non-negative integer such that $s \leq r$, and let $f : \mathcal{M} \to \mathcal{M}'$ be a map. Let $m \in \mathcal{M}$. The map f is said differentiable at m if, for every chart (U, ϕ) of $\mathcal{M}$ at m and every chart (V, ψ) of N at $f(m)$, the map $\psi \circ f \circ \phi^{-1}$ is differentiable at $\phi(m)$* [4]. *The map f is said differentiable if it is differentiable at each point of $\mathcal{M}$.*

Remark A.1.5 Clearly, the notion of differentiability depends on the differential structure. for instance, with reference to Remark A.1.3, the map $t \mapsto t^{\frac{1}{3}}$ is not differentiable from $(I\!R, \phi_1)$ into itself but it is differentiable as map from $(I\!R, \phi_3)$ into $(I\!R, \phi_1)$.

[4]Notice that $\psi \circ \phi^{-1}$ maps an open subset of $I\!R^n$ into $I\!R^d$ —with the usual Euclidean structure—, where the notion of differentiability coincides with the classical one.

A.1.1 The tangent bundle

Definition A.1.6 *Let $\mathcal{M}$ be a n-manifold and let us fix $\bar{m} \in \mathcal{M}$. If (U, ϕ), $(\tilde{U}, \tilde{\phi})$ are charts at m and $v, \tilde{v} \in I\!R^n$ we say that the triples (U, ϕ, v) and $(\tilde{U}, \tilde{\phi}, \tilde{v})$ are equivalent if*

$$\frac{\partial(\tilde{\phi} \circ \phi^{-1})}{\partial x}(\bar{x})(v) = \tilde{v},$$

where we have set $\bar{x} \doteq \phi(\bar{m})$. It is easy to check that this is indeed an equivalence relation. Let $[(U, \phi, v)]$ denote the class of equivalence of (U, ϕ, v). Every equivalence class will be called a **tangent vector at** *m. The set of tangent vectors at m will be called the* **tangent space of $\mathcal{M}$ at** *m, and will be denoted by $T_m\mathcal{M}$.*

$T_m\mathcal{M}$ inherits the structure of vector space from $I\!R^n$. Namely, if $\mathbf{v} = [(U, \phi, v)]$,$\mathbf{w} = [(\mathbf{U}, \phi, \mathbf{w})]$, $\alpha, \beta \in R$, we set

$$\alpha\mathbf{v} + \beta\mathbf{w} \doteq [(U, \phi, \alpha v + \beta w)].$$

Because of the linearity of the maps $\frac{\partial(\tilde{\phi}\circ\phi^{-1})}{\partial x}(\bar{x})(\cdot)$, this definition is in fact independent of the representatives of $\mathbf{v}$ *and* $\mathbf{w}$*.*

Let $\mathcal{M}$ be a n-manifold of class C^r, with $r > 1$. Let us consider the set

$$T\mathcal{M} \doteq \bigcup_{m \in \mathcal{M}} T_m\mathcal{M}.$$

We can endow $T\mathcal{M}$ with a C^{r-1} differential structure in the following way.

For every subset $U \subseteq \mathcal{M}$ let us define the subset $TU \subseteq TM$ by letting

$$TU \doteq \bigcup_{m \in U} T_m\mathcal{M}.$$

Let $(U_i, \phi_i)_{i \in I}$ be an atlas for $\mathcal{M}$. Let us define an atlas $(\hat{U}_i, \Phi_i)_{i \in I}$ for $T\mathcal{M}$ by setting

$$\hat{U}_i \doteq TU_i \qquad \Phi_i(m, \mathbf{v}) \doteq (x, v)$$

as soon as $\phi_i(m) = x$ and $\mathbf{v} = [(U_i, \phi_i, v)]$.

Clearly, each map Φ_i is a bijection from $\hat{U}_i$ onto $\phi_i(U_i) \times I\!R^n$. Moreover, for all pair i, j such that $U_i \cap U_j \neq \emptyset$, one has $\hat{U}_i \cap \hat{U}_j \neq \emptyset$. Furthermore, the transition map $\Phi_i \circ \Phi_j^{-1}$ is defined by

$$\Phi_i \circ \Phi_j^{-1}(x, v) = \left(\phi_i \circ \phi_j^{-1}(x)\,,\, \frac{\partial(\phi_i \circ \phi_j^{-1})}{\partial x} \cdot v \right),$$

which is clearly a diffeomorphism of class C^{r-1} from $\phi_j(U_j) \times I\!R^n$ onto $\phi_i(U_i) \times I\!R^n$ Let us remark that the maps $\hat{\phi}_i$ are linear in the second argument.

Fix $m \in M$ and let (ϕ, U) be a chart at m. If $x_1, \ldots, x_n$ denote the coordinates corresponding to the chart (ϕ, U) , the vectors

$$\frac{\partial}{\partial x^j} \doteq [(\phi, U, e_j)], \qquad j = 1, \ldots, n$$

—where e_j stands for the j-th element of the canonical base of $I\!R^n$— form a basis for the vector space $T_m\mathcal{M}$. In particular, $\Phi_i(m, \mathbf{v}) = (x, v)$ if and only if $x = \phi_i(m)$ and

$$\mathbf{v} = v^i \frac{\partial}{\partial x^i},$$

where we have adopted the so-called Einstein summation convention (which prescribes that if in an expression an index $i \in I$ appears twice we have to perform the summation of the expression for all $i \in I$.)

The collection $(TU_i, \Phi_i)_{i\in I}$ is clearly a C^{r-1} atlas on TM which makes $T\mathcal{M}$ a $2n$-manifold of class C^{r-1}. This manifold is called *the tangent bundle of* $\mathcal{M}$ [5]

Notational convention. If $f : \mathcal{M} \to \mathcal{Q}$ is a map between two manifolds $\mathcal{M}$ and $\mathcal{Q}$, $m \in \mathcal{M}$, and if $(U, x = \phi)$ and $(V, y = \psi)$ are charts at m and $f(m)$, respectively, respectively, let us define the map $f_{\phi,\psi} : I\!R^n \supseteq \phi(U \cap f^{-1}(V)) \to\to I\!R^d$ by setting

$$f_{\phi,\psi} \doteq \psi \circ f \circ \phi^{-1}.$$

Let us call this map the *representation of* f *in the coordinate charts* $(U, x = \phi)$ *and* $(V, y = \psi)$.

$\diamondsuit$

Definition A.1.7 *Let* $\mathcal{M}, \mathcal{Q}$ *be manifolds of dimension* n *and* d*, respectively, and let* $f : \mathcal{M} \to \mathcal{Q}$ *be a differentiable map. We define a map*

$$Df : T\mathcal{M} \to T\mathcal{Q}$$

by setting, for every $\bar{m} \in \mathcal{M}$ *and* $\mathbf{v} \in T_{\bar{m}}\mathcal{M}$,

$$Df(\bar{m}, \mathbf{v}) \doteq \big(f(\bar{m}), T_{\bar{m}} f(\mathbf{v})\big) \quad \big(\in T_{f(\bar{m})}\mathcal{Q}\big)$$

[5] If we performed the same construction starting with an atlas $(U'_{i'}, \phi'_{i'})_{i'\in I'}$ equivalent to $(U_i, \phi_i)_{i\in I}$ we would end up with an atlas $(\hat{U}'_{i'}, \hat{\phi}'_{i'\in I'})_{i'\in I'}$ for $T\mathcal{M}$ which is equivalent to $(\hat{U}_i, \hat{\phi}_i)_{i\in I}$. We leave the proof of this trivial fact as an exercise.

where, if $(U, x = \phi)$ and $(V, y = \psi)$ are charts at $\bar{m}$ and $f(\bar{m})$, respectively, and $\mathbf{v} = v^i \frac{\partial}{\partial x_i}$, $\bar{x} \doteq \phi(\bar{m})$, we have set

$$T_{\bar{m}} f(\mathbf{v}) = \left(\frac{\partial f^j_{\phi,\psi}}{\partial x^i}(\bar{x}) v_i \right) \frac{\partial}{\partial y_j}. \tag{A.1}$$

The function Df is called the tangent map of f, *and the linear function*

$$T_{\bar{m}} f(\cdot) : T_m \mathcal{M} \to T_n \mathcal{Q}$$

is called the tangent map to f at m. *Sometimes f_* is also called the* derivative (*or* the differential) *of f .*

Notice that (A.1) gives the coordinate expression of $T_{\bar{m}} f(m)(\mathbf{v}) = w^j \frac{\partial}{\partial y_j}$:

$$w^j = \frac{\partial f^j_{\phi,\psi}}{\partial x^i} v^i \qquad \forall j = 1, \dots, d$$

The proof that the definition of $T_{\bar{m}} f(\mathbf{v})$ is independent of the choices of the charts at m and $f(m)$ is trivial, and we let it as an exercise.

Let us observe that the rank of the Jacobian matrix $\frac{\partial f_{\phi,\psi}}{\partial x}(\bar{x})$ is in fact independent of the chosen charts. We call this number the *rank of $T_{\bar{m}} f$*. If the rank of $T_{\bar{m}} f$ is equal to k for every $\bar{m} \in M$, we say that Df has *constant rank equal to k.*

A.2 Submanifolds

Definition A.2.1 *Let $\mathcal{M}, \mathcal{Q}$ be manifolds of dimension n and d, respectively, and let $f : \mathcal{M} \to \mathcal{Q}$ be a differentiable map.*

- *f is called an* immersion *if, for every $m \in M$, $T_{\bar{m}} f$ is injective, i.e., if Df has constant rank equal to n.*
- *f is called a* submersion *if, for every $m \in M$, $T_{\bar{m}} f$ is surjective, i.e., if Df has constant rank equal to d.*

The following Lemma characterizes immersions and submersions as having special local expression when composed with suitable charts.

Lemma A.2.2 *A function $f : \mathcal{M} \to \mathcal{Q}$ is a submersion if and only if*

(A) *$d \leq n$, and, for every $m \in \mathcal{M}$, there exists a chart $(U, x = \phi)$ at m such that, for every chart $(V, y = \psi)$ at $f(m)$, $f_{\phi,\psi}$ coincides with the projection on the first d components, i.e.,*

$$f_{\phi,\psi}(x_1, \cdots, x_d, \ldots, x_n) = (x_1, \cdots, x_d) \qquad \forall (x_1, \cdots, x_d, \ldots, x_n) \in \phi(U).$$

In particular, every submersion is an open map, *i.e. it maps open sets into open sets.*

PROOF Let us assume **(A)**. Then, for every $m \in \mathcal{M}$, f is a submersion because

$$rank T_{\bar{m}} f = rank \, \frac{\partial f_{\phi,\psi}}{\partial x}(\bar{x})$$

$$= rank \begin{pmatrix} \mathbf{1}_d & 0_{d,n-d} \\ & \\ 0_{n-d,d} & 0_{n-d,n-d} \end{pmatrix} = d,$$

where $\bar{x} \doteq \phi(m)$, $\mathbf{1}_d$ stands for the unit matrix of dimension d, and, for all positive integers i, j, $0_{i,j}$ denotes the $i \times j$ zero matrix. Conversely, let f be a submersion at a point $m \in \mathcal{M}$. Let $(W, z = \xi)$ and $(V, y = \psi)$ be charts at m and $f(m)$ respectively. For every element $z \in I\!R^n$ let us set $z^I = (z^1, \ldots, z^d)$, and $z^{II} = (z^{d+1}, \ldots, z^n)$. Without loss of generality, we can assume that the first d columns of the Jacobian matrix $\frac{\partial f_{\xi,\psi}}{\partial z}(\bar{x})$ form a non-singular matrix. We use $\frac{\partial f_{\xi,\psi}}{\partial z^I}(\bar{x})$ to denote this matrix. Let us consider the map

$$k : \xi(W) \to I\!R^n$$

defined by

$$k(z) = \left(k^1(z), k^{II}(z)\right) = \left(f_{\xi,\psi}(z), z^{II}\right)$$

Since the matrix $\frac{\partial f_{\xi,\psi}}{\partial z^I}(\bar{x})$ is non singular, the matrix $\frac{\partial k}{\partial z}(\bar{x})$ is non-singular as well, so, by the Inverse Map Theorem, there exists an open neighborhood $Z \subseteq \xi(W) (\subseteq I\!R^n)$ of $\bar{x}$ such that $k(Z) \subseteq I\!R^n$ is open and the restriction $k_{|Z}$ is a diffeomorphism from Z onto $k(Z)$. Hence the pair $(U, x = \phi)$ defined by

$$U \doteq \xi^{-1}(Z) \qquad \phi \doteq k \circ \xi : U \to \phi(U) \subseteq I\!R^n.$$

is a chart. Moreover, it verifies **(A)**. Indeed, one has

$$f_{\phi,\psi}(x) = f_{\xi,\psi} \circ k^{-1}(x) = k^I \circ k^{-1}(x) = x^I = (x_1, \cdots, x_d)$$

for all $x \in \phi(U)$.

♢

Definition A.2.3 *We call a chart $(U, x = \phi)$ on $\mathcal{M}$ like the one whose existence is stated in Lemma A.2.2* a chart adapted to the submersion f.

Lemma A.2.4 *A function $f : \mathcal{M} \to \mathcal{Q}$ is an immersion if and only if*

(B) *$n \leq d$, and, for every $m \in \mathcal{M}$, there exists a chart $(V, y = \psi)$ at $f(m)$ such that, for every chart $(U, x = \phi)$ at m, there is a neighborhood U' of $f(m)$ such that*

$$f_{\phi,\psi}(x_1, \cdots, x_d) = (x_1, \cdots, x_d, 0, \ldots, 0)$$

for all $(x_1, \cdots, x_d) \in \phi(U \cap U')$. In particular, every immersion f is locally injective*, i.e., for every $m \in M$, there exists a neighbourhood U of M such that the restriction $f : U \to \mathcal{Q}$ is injective.*

Proof For any vector $v = (v^1, \ldots, v^d) \in \mathbb{R}^d$, let us set $v^I \doteq (v^1, \ldots, v^n)$ and $v^{II} \doteq (v^{n+1}, \ldots, v^d)$. Let $(U, x = \phi)$ be a chart of $\mathcal{M}$ at m, let $(W, z = \xi)$ be a chart of $\mathcal{M}'$ at $m' = f(m)$, and let us set $\bar{x} \doteq \phi(m)$. Let us consider the map

$$x \mapsto f^I_{\phi,\xi}(x)$$

from $\phi(U)$ into $\mathbb{R}^d$.

Clearly, it is not restrictive to assume that the Jacobian matrix

$$\frac{\partial f^I_{\phi,\xi}}{\partial x}(\bar{x})$$

is non-singular. Then, by the Inverse Mapping Theorem, there exists an open neighborhood A of $\bar{x}$ such that $f^I_{\phi,\xi}(A) \subset \mathbb{R}^n$ is open and the restriction

$$f^I_{\phi,\xi} : A \to f^I_{\phi,\xi}(A)$$

is a diffeomorphism of class C^r. Let

$$\eta : \tilde{f}^I_{\phi,\xi}(A) \to A$$

denote the inverse map, and let us define the map $\psi = (\psi^I, \psi^{II})$ by setting

$$\psi^I \doteq \eta \circ \xi^I \qquad \psi^{II} \doteq \xi^{II} - f^{II}_{\phi,\xi} \circ \eta \circ \xi^I$$

For every $y \in \xi(W \cap V)$ one has,

$$\psi \circ \xi^{-1}(y) = \Big(\eta(y^I)\,,\, y^{II} - f^{II}_{\phi,\xi} \circ \eta \circ \xi^I \circ \xi^{-1}(y)\Big) = \Big(\eta(y^I)\,,\, y^{II} - f^{II}_{\phi,\xi} \circ \eta(y^I)\big),$$

so the Jacobian matrix of $\psi \circ \xi^{-1}$ is nonsingular at any $y \in \xi(W \cap V)$, for the Jacobian matrix of η is nonsingular at any $y \in f^I_{\phi,\xi}(A)$. It follows that the map ψ is a homeomorphism from $V \doteq W \cap \xi^{-1}(f^I_{\phi,\xi}(A))$ onto its image, and $(V, y = \psi)$ is a chart at $f(m)$.

Finally, for every $x \in \phi(A)$, one has

$$f_{\phi,\psi}(x) = (\psi^I, \psi^{II}) \circ f \circ \phi^{-1}(x) = (x, 0),$$

which concludes the proof.

$\diamondsuit$

Definition A.2.5 *We call a chart $(V, y = \phi)$ on $\mathcal{Q}$ like the one of Lemma A.2.4 a chart* adapted *to the immersion f.* [6]

Definition A.2.6 (Submanifolds). *Let r, s be positive integers such that $s \leq r$. A subset $\mathcal{Q}$ of a differential manifold $\mathcal{M}$ of class C^r is called a* submanifold of class C^s *if*

i) $\mathcal{Q}$ is a manifold of class C^s, and

ii) the inclusion map $i : \mathcal{Q} \to \mathcal{M}$ is am immersion of class C^s.

We shall say that $\mathcal{Q}$ is a submanifold *(without additional specification) if it is a submanifold of class C^r.*

Definition A.2.7 (Embedded submanifolds). *A submanifold $\mathcal{Q} \subseteq \mathcal{M}$ (of any class) is called* embedded *if the manifold topology of $\mathcal{Q}$ coincides with the relative topology induced by the topology of $\mathcal{M}$. A submanifold $\mathcal{Q} \subseteq \mathcal{M}$ (of any class) is called* locally embedded *if for any point $q \in \mathcal{Q}$ there exist a neighborhood A_q of q (in $\mathcal{Q}$!) such that A_q is an embedded submanifold.*

Examples The image $\mathcal{Q}_1 = c(]0, 2\pi[)$ of the map $c :]0, 2\pi[\to I\!R^2$, $c(s) \doteq (\sin s, \sin(2s))$ can be given the structure a one-dimensional submanifold of $I\!R^2$of class C^∞ [7]But $\mathcal{Q}$ is not embedded: for instance, $c(]\pi - 1, \pi + 1[)$, which is a neighborhood of $q = (0, 0)$ in the topology of $\mathcal{Q}$, does not contain any neighborhood of q in the topology induced by the topology of $I\!R^2$.

[6]0f course, 0 can be replaced with any other constant $a \in I\!R^{n-d}$.

[7]We mean that $\mathcal{Q}_1$ is endowed with the one-chart atlas $\mathcal{A} = \{\mathcal{Q}_1, c^{-1}\}$. Of course one can consider different, non equivalent atlases. For instance, if $c' \doteq c(s) \doteq (\sin s, -\sin(2s))$, $\mathcal{A}' \doteq \{\mathcal{Q}_1, c'^{-1}\}$ is not equivalent to $\mathcal{A}$.

On the other hand $\mathcal{Q}_2 = \{(x, y) \in I\!R^2 \quad x^2 + y^2 = 1\}$ is an embedded submanifold of class C^∞. Finally $\mathcal{Q}_3 \doteq c[0, \pi]$ cannot be given the structure of a C^1 submanifold .(Why? Prove that $\mathcal{Q}_3$ is a Lipschitz embedded submanifold[8])

Proposition A.2.8 (Characterization of embedded submanifolds) *Let $\mathcal{M}$ be a manifold of class C^r and consider a subset $\mathcal{Q} \subseteq \mathcal{M}$. Then the following conditions are equivalent:*

1) $\mathcal{Q}$ is an embedded submanifold of class C^r and dimension $n - d \leq n$,

2) for every $q \in \mathcal{Q}$ there exists a chart (U, ϕ) and open subsets $A_1 \subset I\!R^d, A_2 \subset I\!R^{n-d}$ such that, $m(U) = A_1 \times A_2$, and, if $x(q) = (a_1, a_2)$, then

$$\mathcal{Q} \cap U = \phi^{-1}(\{a_1\} \times A_2)$$

Charts like the one in condition 2) are called adapted to the (embedded) submanifold $\mathcal{Q}$.

Remark A.2.9 *Clearly, if $\mathcal{Q}$ is an embedded submanifold, the collection of pairs*

$$\left\{ \left(U_i \cap \mathcal{Q}, \phi_{i|_{U_i \cap \mathcal{Q}}} \right) \quad | \quad (U, \phi) \text{ is adapted to } N \right\}$$

is an atlas for the manifold N.

Embedded submanifolds are often obtained as pre-image of submersions or as images of suitable restrictions of immersions. This is made precise by the following result:

Theorem A.2.10 *Let $\mathcal{M}$ $\mathcal{M}'$ be manifolds of class C^r, with $r \geq 1$, and let n and d be the dimensions of $\mathcal{M}$ and $\mathcal{M}'$, respectively. Let*

$$f : \mathcal{M} \to \mathcal{M}'$$

be a differentiable mapping of class C^r. Then:

1) if f is a submersion, for each $m' \in f(M)$, $f^{-1}(m')$ is an embedded submanifold of $\mathcal{M}$ of class C^r and dimension $n - d$;

[8] A Lipschitz manifold is a manifold whose transition maps are locally Lipschitz. Clearly, all manifolds of class C^r with $r \geq 1$, are Lipschitz. A Lipschitz submanifold of a $\mathcal{M}$ is a subset $\mathcal{Q} \subset \mathcal{M}$ with the structure of a Lipschitz manifold and such that the immersion $i : \mathcal{Q} \to \mathcal{M}$ is locally injective and locally Lipschitz.

2) if f is an immersion, then the image $f(M)$ is a locally embedded submanifold of class C^r.

Proof Let us prove 1). Let $m' \in \mathcal{M}'$ and let us set $\mathcal{Q} \doteq f^{-1}(m')$. Let us choose $m \in \mathcal{Q}$, and let $(V, y = \psi)$ be a chart at m'. By Lemma **??** there exists a chart $(U, x = \phi)$ at m such that

$$f_{\phi,\psi}(x^1, \ldots, x^n) = (x_1, \ldots, x^d)$$

for all $(x^1, \ldots, x^n) \in \phi(U)$. Of course, it is not restrictive to assume that there exist open subsets $A_1 \subset I\!R^d$, $A_2 \subseteq I\!R^{n-d}$ such that $\phi(U) = A_1 \times A_2$ and $\psi(V) = A_1$ and $\phi(m) = (0, 0) \in A_1 \times A_2$. Hence

$$\phi(\mathcal{Q} \cap U) = \{0\} \times A_2,$$

which, by Proposition A.2.8, implies that $\mathcal{Q}(= f^{-1}(m'))$ is an embedded submanifold od dimension $n - d$.

Let us prove [2)]. Let $m \in \mathcal{M}$ and let us set $m' \doteq f(m)$. Let $(U, x = \phi)$ be a chart at m. By Lemma A.2.4 there is a chart $(V, y = \psi)$ and an open neighborhood $U' \subseteq U$ of m such that

$$f_{\phi,\psi}(x_1, \cdots, x_d) = (x_1, \cdots, x_d, 0, \ldots, 0) \tag{A.2}$$

for all $(x_1, \cdots, x_d) \in A_1 \doteq \phi(U \cap U')$. It is not restrictive to assume that there exists an open subset $A_2 \in I\!R^{d-n}$ such that $(0, 0) \in A_1 \times A_2$, $\psi(m') = (0, 0)$, and $\psi(V) = A_1 \times A_2$. Hence, by (A.2),

$$\psi(f(U \cap U') \cap V) = \psi \circ f \circ \phi^{-1}(\phi(U \cap U') \cap V) = A_1 \times \{0\},$$

which, by Proposition A.2.8 implies that $f(U \cap U')$ is an embedded submanifold.

$\diamondsuit$

Theorem A.2.10 can be easily generalized to the following result, sometimes referred to as the *constant rank theorem*

Theorem A.2.11 (Constant rank theorem) *Let $\mathcal{M}$ $\mathcal{M}'$ be manifolds of class C^r, with $r \geq 1$. Let $f : \mathcal{M} \to \mathcal{M}'$ be a differentiable mapping of class C^r such that Df has constant rank k. Then:*

- *for each $m' \in f(M)$ $f^{-1}(m')$ is an embedded submanifold of $\mathcal{M}$ of class C^r and dimension $n - k$, where n is the dimension of $\mathcal{M}$;*
- *for each $m \in \mathcal{M}$ there is a neighborhood U_m of m such that the image $f(U_m)$ is an embedded submanifold of $\mathcal{M}'$ of dimension k.*

A.3 Vector bundles

Definition A.3.1 *Let r, s be non-negative integers. A s-dimensional* vector bundle *of class C^r is a triple $(E, \mathcal{M}, \pi)$, where E is a manifold of class C^r (the* total space*), $\mathcal{M}$ is a manifold (the* base manifold*, or simply the* base*) and $\pi : E \to \mathcal{M}$ is a surjective map (the* projection*), verifying the following conditions:*

i) for every $m \in \mathcal{M}$, the fiber *$E_m \doteq \pi^{-1}(m)$ has a structure of vector space;*

ii) for every $m \in \mathcal{M}$, there is a neighborhood U of m and a diffeomorphism $\Phi : \pi^{-1}(U) \to U \times I\!R^s$ of class C^r —called a local trivialization *of E— such that,*

$$Pr_1 \circ \Phi(e) = \pi(e),$$

for all $e \in \pi^{-1}(U)$, where Pr_1 denotes the projection of $U \times I\!R^s$ onto the first factor. Moreover, for every $m \in U$, the restriction of Φ to the fiber E_m, $\Phi : E_m \to \{m\} \times I\!R^s$ is a linear isomorphism.

The first example of fiber bundle is the tangent bundle $T\mathcal{M}$ of a $n-$*dimensional* manifold $\mathcal{M}$ of class C^r, with $r \geq 2$, which is in fact a n-dimensional fiber bundle of class C^{r-1} as soon we let the total space, the base, and the projection be $T\mathcal{M}$, $\mathcal{M}$ and $\pi_{T_\mathcal{M}}$, the latter denoting the canonical projection of $T\mathcal{M}$ onto $\mathcal{M}$. For each $m \in \mathcal{M}$, if $(U, x = \phi)$ is a chart at m, $\Phi \doteq \big(\pi_{T_\mathcal{M}}, \Phi\big)$ is clearly a local trivialization from $(\pi_{T_\mathcal{M}})^{-1}(U)$ onto $U \times I\!R^n$.

Definition A.3.2 (Sections) *Let $(E, \mathcal{M}, \pi)$ be a fiber bundle. A map $s : \mathcal{M} \to E$ is called a* section of $(E, \mathcal{M}, \pi)$ *(shortly,* of E*) if*

$$\pi \circ s(m) = m$$

for all $m \in \mathcal{M}'$. In particular, a section of the tangent bundle TM is called a vector field.

Starting from the tangent bundle $T\mathcal{M}$, we can construct other fiber bundles on the base $\mathcal{M}$, the *tensor bundles.* Let us begin with recalling the notion of tensor on a vector space.

A.3.1 Tensor calculus on vector spaces

. Let n be a positive integer, and let V be a n-dimensional real vector space. We use V^* to denote the dual space of V, that is the n-dimensional vector space of linear functionals from V into $I\!R$ —which are also called *covectors*, or *linear forms*. If $(\alpha, v) \in V^* \times V$ let us set

$$< \alpha, v > \doteq \alpha(v).$$

$< \cdot, \cdot >$ is called the *natural pairing* of V and V^*. We identify $(V^*)^* = V^{**}$ with V via the pairing $\prec v, \alpha \succ \doteq < \alpha, v >$.

Let r be a positive integer. A *covariant r-tensor* on V is a multilinear mapping

$$G : \underbrace{V \times \cdots \times V}_{r \text{ times}} \to I\!R$$

The vector space of r-covariant tensors will be denoted by $T^r_0(V)$

Let s be a positive integer.A *contravariant s-tensor* on V is a multilinear mapping

$$G : \underbrace{V^* \times \cdots \times V^*}_{s \text{ times}} \to I\!R$$

The vector space of s-contravariant tensors will be denoted by $T^0_s(V)$ (Notice that $T^s_0(V) = T^0_s(V^*)$).

A *mixed tensor of type* $\begin{pmatrix} r \\ s \end{pmatrix}$ (i.e.,*r-covariant and s-contravariant*) is a multilinear mapping

$$G : \underbrace{V^* \times \cdots \times V^*}_{s \text{ times}} \times \underbrace{V \times \cdots \times V}_{r \text{ times}} \to I\!R$$

The vector space of tensors of type $\begin{pmatrix} r \\ s \end{pmatrix}$ will be denoted by $T^r_s(V)$. Clearly' one has $T^0_1(V) = V$ (via the identification $V^{**} = V$, $T^1_0(V) = V^*$. Moreover, one can identify $T^1_1(V)$ with $End(V)$, the space of *endomorphisms* on V[9]: it is sufficient to identify any $f \in End(V)$ with the tensor $L_f \in T^1_1(V)$ defined by $L_f(\alpha, v) \doteq < \alpha, f(v) > (\alpha \circ f(v))$.

[9]An endomorphism is an linear mapping from V into itself.

Given two tensors $F \in T^r_s(V), G \in T^p_q(V)$ let us define the tensor $F \otimes G \in T^{r+p}_{s+q}(V)$ by setting

$$F \otimes G(\alpha^1, \ldots, \alpha^{s+q}, v_1, \ldots, v_{r+p}) \doteq$$

$$F(\alpha^1, \ldots, \alpha^s, v_1, \ldots, v_r)G(\alpha^{s+1}, \ldots, \alpha^{s+q}, v_{r+1}, \ldots, v_{r+p})$$

The tensor $F \otimes G$ is called the *tensor product* of F and G. In particular, if $\{v_1, \ldots, v_n\}$ is a basis of V and $\{v^1, \ldots, v^n\}$ is the dual basis of V^*, for all ordered s-tuples $i_1, \ldots, i_s$ and all ordered r-tuples $j_1, \ldots, i_r$ of elements of $\{1, \ldots, n\}$ one can consider the tensor products

$$v_{i_1} \otimes \cdots \otimes v_{i_s} \otimes v^{j_1} \otimes \cdots \otimes v^{j_r}.$$

It is trivial to verify that the so-obtained n^{s+r} tensors (of type $\begin{pmatrix} r \\ s \end{pmatrix}$) form a basis for the vector space $T^r_s(V)$. In particular, $T^r_s(V)$ has dimension equal to n^{s+r}. In fact, for every tensor $F \in T^r_s(V)$, one has

$$F = F^{i_1,\ldots,i_s}_{j_1,\ldots,i_r} v_{i_1} \otimes \cdots \otimes v_{i_s} \otimes v^{j_1} \otimes \cdots \otimes v^{j_r}$$

where the components $F^{i_1,\ldots,i_s}_{j_1,\ldots,i_r}$ are defined as follows:

$$F^{i_1,\ldots,i_s}_{j_1,\ldots,i_r} \doteq F(v^{j_1}, \ldots, v^{j_r}, v_{i_1}, \ldots, v_{i_s}) .$$

Proposition A.3.3 (Change of variables) *Let us consider two basis,*

$$\mathcal{B} = \{v_1, \ldots, v_n\} \quad , \quad \tilde{\mathcal{B}} = \{\tilde{v}_1, \ldots, \tilde{v}_n\}$$

and the corresponding dual bases,

$$\mathcal{B}^* = \{v^1, \ldots, v^n\} \quad , \quad \tilde{\mathcal{B}}^* = \{\tilde{v}^1, \ldots, \tilde{v}^n\}$$

. Let $F \in T^r_s(V)$ and let $F^{i_1,\ldots,i_s}_{j_1,\ldots,i_r}$ and $\tilde{F}^{i_1,\ldots,i_s}_{j_1,\ldots,i_r}$ be the components of F in the bases $\mathcal{B}, \mathcal{B}^*$ and $\tilde{\mathcal{B}}, \tilde{\mathcal{B}}^*$, respectively, that is

$$F = F^{i_1,\ldots,i_s}_{j_1,\ldots,i_r} v_{i_1} \otimes \cdots \otimes v_{i_s} \otimes v^{j_1} \otimes \cdots \otimes v^{j_r} =$$

$$\tilde{F}^{i_1,\ldots,i_s}_{j_1,\ldots,i_r} \tilde{v}_{i_1} \otimes \cdots \otimes \tilde{v}_{i_s} \otimes \tilde{v}^{j_1} \otimes \cdots \otimes \tilde{v}^{j_r}.$$

Let $A = (A_j^k)$ be the real, nonsingular $n \times n$ matrix —where the upper and lower indexes be regarded as a column and row indexes, respectively— such that

$$\tilde{v}_j = A_j^k v_k \qquad \forall j = 1, \dots, n.$$

Let $\hat{A} = \hat{A}_j^k$ denote the transposed of A, namely $\hat{A}_j^k \doteq A_k^j$, and let $C = C_j^k$ denote the inverse of $\hat{A}$. Then

$$\tilde{F}_{j_1,\dots,i_r}^{i_1,\dots,i_s} = C_{k_1}^{i_1} \dots C_{k_s}^{i_s} \hat{A}_{j_1}^{h_1} \dots \hat{A}_{j_s}^{h_r} F_{h_1,\dots,h_r}^{k_1,\dots,k_s}.$$

Alternating forms

For every non negative integer $k \leq$ let us consider the subspace $\Lambda^k(V) \subset T_0^k(V)$ of those multinear maps

$$\omega : \underbrace{V \times \cdots \times V} \to I\!R$$

(i.e. k-covariant 0-contravariant tensors) which are *alternating* (equivalently, *anti-symmetric*). In general, a tensor F is alternating if it change sign whenever two arguments are interchanged. With obvious meaning of the notation, this means that:

$$F(\dots, a, \dots, b, \dots) = -F(\dots, b, \dots, a, \dots)$$

These tensors of $\Lambda^k(V)$ called *(exterior) k-forms*. Clearly $\Lambda^1(V) = V^*$.

Let us just define the *wedge product* $\wedge$ on external forms. The wedge product is bilinear and associative. Moreover r is a non-negative integer, and $\omega_1, \dots, \omega_r$ are 1-forms, the product $\omega_1 \wedge \cdots \wedge \omega_r$ is a $r-$ form defined by setting

$$\omega_1 \wedge \cdots \wedge \omega_r(v_1, \dots, v_r) = det\left(\langle \omega_i, v_j \rangle\right)_{i,j=1,\dots r}$$

The extension to all forms can be made by linearity, by observing that

- If $v^1, \dots, v^n$ is a basis of V^*, for every $k \leq n$, the $\begin{pmatrix} n \\ k \end{pmatrix}$ elements

  $$v^{i_1} \wedge \cdots \wedge v^{i_k} \qquad i_1 < \cdots < i_k$$

  form a basis of $\Lambda^k(V)$, which in particular has dimension $\begin{pmatrix} n \\ k \end{pmatrix}$

In particular, let us observe that

- If $\alpha, \beta \in V^*(= \Lambda^1(V))$ one has

$$\alpha \wedge \beta = -\beta \wedge \alpha = \alpha \otimes \beta - \beta \otimes \alpha.$$

- If α is a k-form and β is a h-form, and $k + h \leq n$, then $\alpha \wedge \beta$ is a $k + h$-form, and

$$\alpha \wedge \beta = (-1)^{k+h+1} \beta \wedge \alpha.$$

Symplectic forms

A *symplectic* form on a vector space$V$ is a non-degenerate 2-form on V. We recall that a 2-form ω is non-degenerate if, for a given $v \in V$,

$$\omega(v, w) = 0 \quad \forall w \in V \Rightarrow v = 0.$$

Using non-degeneracy and skew-symmetry one can show that the existence of a symplectic form implies that *the dimension of* V *is even.*

A symplectic form establishes a isomorphism $S : V \to V^*$, which is defined by setting, for all $v, w \in V$,

$$< S(v), w > \doteq \omega(v, w).$$

Clearly, S is linear. Moreover, since ω is non-degenerate, one has

$$v \in kerS \Leftrightarrow S(v) = 0 \Leftrightarrow \omega(v, w) = 0 \quad \forall w \in V \Leftrightarrow v = 0,$$

so that S is injective, which implies that that S is an isomorphism. The map S^{-1} is used to define *Hamiltonian vector fields* on the cotangent bundle of a manifold (see below).

A.3.2 Tensor bundles

Let k, l be non-negative integers and let us consider the set

$$T^l_k\mathcal{M} \doteq \bigcup_{m \in \mathcal{M}} T^l_k(T_m\mathcal{M})$$

Similarly to the case of the tangent bundle, this set can be endowed with a structure of fiber bundle as follows ...... $T^l_k\mathcal{M}$ is called the *bundle of* $\begin{pmatrix} l \\ k \end{pmatrix}$-*tensors on* M. Notice that this bundle (with base $\mathcal{M}$) have dimension n^{l+k}. So, if $\mathcal{M}$ is a manifold of class C^r, with $r \geq 1$, as a manifold $T^l_k\mathcal{M}$ has dimension $n + n^{l+k}$.

Also, we shall set

$$\Lambda^k\mathcal{M} \doteq \bigcup_{m\in\mathcal{M}} \Lambda^k(T_m\mathcal{M})$$

and endow it of a structure of fiber bundle in the usual way. Notice that this bundle (with base $\mathcal{M}$) have dimension $\left(\begin{array}{c} n \\ k \end{array}\right)$). So, if $\mathcal{M}$ is a manifold of class C^r, with $r \geq 1$, the manifold $\Lambda^k\mathcal{M}$ has dimension $\left(\begin{array}{c} n \\ k \end{array}\right)$.

A.3.3 Some examples of sections of tensor bundles

We have already given the general notion of section of a fiber bundle. In particular, sections of tensor bundles on $\mathcal{M}$ are called also *tensor fields on* $\mathcal{M}$. The sections of TM are usually called **vector fields**, while the sections of T^*M are called **differential 1-form**on $\mathcal{M}$, or, also, **covector fields**. The sections of $\Lambda^k\mathcal{M}$ are called **differential k-forms** $\mathcal{M}$. (For brevity, sometimes they are simply called k-forms.)

Since tensor fields are mappings between differentiable manifolds it is clear what one means that a tensor field is of class C^r. Notice, in particular, that id $\mathcal{M}$ is a manifold of class C^s then a tensor field can be at most of class C^{s-1}.

Let us see some special cases of tensor fields that are of common use.

Vector fields

The vector fields are the sections of the tangent bundle. They are connected with the notion of ordinary differential equation (see below).

Riemannian metrics

A 2-covariant tensor field g is called a *Riemannian metric* if for every $m \in \mathcal{M}$ $g(m)$ is non-degenerate, symmetric, and positive definite. This means, respectively,

$$g(m)(v,w) = 0 \qquad \forall w \in T_m\mathcal{M} \to v = 0$$

$$g(m)(v,w) = g(m)(w,v) \qquad \forall v,w \in T_m\mathcal{M},$$

and

$$g(m)(v,v) > 0 \qquad \forall v \in T_m\mathcal{M}.$$

In other words, for every $m \in \mathcal{M}$ $g(m)$ is a *scalar product* on the vector space $T_m\mathcal{M}$. A manifold of class C^r $(r \geq 1)$ equipped with a Riemannian metrics is called a Riemannian manifolds.

Using partition of unity [10] one can prove that on every manifold there exists one Riemannian metric. The metric is said *flat* if there is exists atlas $\mathcal{A}$ such that on each chart $(U, x = \phi)$ of $\mathcal{A}$ one has

$$g(m) = g_{ij} dx_i \otimes dx_j \qquad \forall m \in U$$

with all the components g_{ij}'s constant (Of course the g_{ij}' can change with the chart).

If $(\mathcal{M}, g^{\mathcal{M}})$ and $(\mathcal{Q}, g^{\mathcal{Q}})$ are Riemannian manifolds, and $F : \mathcal{M} \to \mathcal{Q}$ is a diffeomorphism, f is called an *isometry* if $F_* g^{\mathcal{Q}} = g^{\mathcal{M}}$, $F_* g^{\mathcal{Q}}$ is *the pull-back of* $g^{\mathcal{Q}}$, that is the 2-covariant tensor on $\mathcal{M}$ defined by

$$F_* g^{\mathcal{Q}}(m)(v, w) \doteq g^{\mathcal{Q}}(F(m))\Big(T_m F(v), T_m F(w)\Big)$$

for all $m \in \mathcal{M}$ and all $v, w \in T_m \mathcal{M}$.

Riemannian Geometry investigates the properties of Riemannian manifolds up to isometries. We refer the interested reader to the many existing textbooks on the subject(see e.g. []).

The *duality* tensor field.

Let $\mathcal{M}$ be any manifold. There exists a tensor $\mathcal{I} \in \mathcal{T}_1^1(\mathcal{M})$ such that for *every* chart $(U, x\phi)$ one has

$$\mathcal{I}(m) = \sum_{1=1}^{n} dx^i \otimes \frac{\partial}{\partial x_i}$$

In others word, in all coordinate systems, $(\mathcal{I}_i^j)$ coincides with the identity matrix. This can be defined directly (using any atlas and verifying that if a tensor as components $(\mathcal{I}_i^j)$ on a chart, then it has the same components with the other charts on the intersection of the respective domains), or by defining Id intrinsically: in fact it is sufficient to set

$$\mathcal{I}(m)(\alpha, v) \doteq \langle \alpha, v \rangle \qquad \forall \alpha \in T^* \mathcal{M}, v \in T\mathcal{M}$$

where, as usual, $\langle \cdot, \cdot \rangle$ denotes the duality pairing.

---

[10] A *partition of unity* is....

The Lieuville 1-form on $T^*\mathcal{M}$

Let $\mathcal{M}$ be any manifold class C^2. There exists a one 1-form θ on $T^*\mathcal{M}$ —that is $\theta \in T^*(T^*\mathcal{M})$— such that for *every* chart $(U, x = \phi)$ on M one has

$$\theta(m,p) = p_i dx^i \qquad \forall (m,p) \in T^*U, \tag{A.3}$$

where the p_i are the components of $p \in T^*_m\mathcal{M}$ with respect to the basis $dx^1, \dots, dx^n$. The form θ is called the *Lieuville form.* Let us remark that this is a form *on* $T^*\mathcal{M}$, so (A.3) means that, in the basis $dx^1, \dots, dx^n, dp^1, \dots, dp^n$ of $T^*_{m,p}(T^*\mathcal{M})$, the covector $\theta(m,p) \in T^*_{m,p}(T^*\mathcal{M})$ has components $(p_1, \dots, p_n, 0, \dots, 0)$.

The existence of such a form can be verified directly (by testing that it transforms in the right way with the bundle- transition maps).

However it is usually defined intrinsically as follows:

- Call $pi_{T^*\mathcal{M}}$ the canonical projection

The symplectic 2-form on $T^*\mathcal{M}$

Let $\mathcal{M}$ be any manifold class C^2. There exists a 2-form ω on $T^*\mathcal{M}$ —that is $\omega \in \Omega^2(\mathcal{M})$— such that for *every* chart $(U, x = \phi)$ on M one has

$$\omega(m,p) = dx^i \wedge dp_i \qquad \forall (m,p) \in T^*U, \tag{A.4}$$

where the p_i are the components of $p \in T^*_m\mathcal{M}$ with respect to the basis $dx^1, \dots, dx^n$. The form ω is called the *canonical symplectic form on $T^*\mathcal{M}$.*

A.4 Ordinary Differential Equations

Definition A.4.1 *Let $\mathcal{M}$ be a manifold of class C^r, with $r \geq 1$. A* vector field *on $\mathcal{M}$ is a section of $(T\mathcal{M}, \mathcal{M}, \pi_{T\mathcal{M}})$, where $\pi_{T\mathcal{M}}$ denotes the canonical projection of $T\mathcal{M}$ onto $\mathcal{M}$. In other words, f is a vector field if and only if f is a map $f : \mathcal{M} \to TM$ verifying*

$$f(m) \in \{m\} \times T_m\mathcal{M}$$

for all $m \in \mathcal{M}$.[11]

A time-dependent vector field *on $\mathcal{M}$ is a map is a map $f : I \times \mathcal{M} \to TM$ such that i) I is a real interval, and ii) for every $t \in I$, $f(t, \cdot)$ is a vector field on $\mathcal{M}$.*

[11] Sometimes the notation $f(m)$ is used to denote what is in fact, the second component of $f(m)$ —i.e., an element of $T_m\mathcal{M}$. Yet, in general the real meaning is unambiguous because of the context.

We wish to give an chart-independent notion for a differential equation defined by a (possibly) time-dependent vector field. For this purpose, if I is a real interval, $m : I \to \mathcal{M}$ is a continuous map, and t is an interior point of I such that $m(\cdot)$ is differentiable at t, we use $\dot{m}(t)$ or $\frac{dm}{dt}(t)$ to denote the image of the vector $\frac{\partial}{\partial t} \in T_t I$ through the tangent map of $m(\cdot)$ at t. Namely, we let

$$\dot{m}(t) = \frac{dm}{dt}(t) \doteq m_*(t)\left(\frac{\partial}{\partial t}\right) \qquad (\in T_{m(t)}\mathcal{M})$$

Let us $(U, x = \phi)$ be a chart at $m(t)$, so that the map $x(s) \doteq \phi \circ m(s)$ can be defined in $]t-\delta, t+\delta[$ for a δ sufficiently small. Then $x(\cdot) :]t-\delta, t+\delta[\to I\!R^n$ is differentiable at $s = t$, and

$$\dot{m}(t) = \frac{dm}{dt}(t) = \frac{dx^i}{dt}(t)\frac{\partial}{\partial x_i}.$$

Therefore, if f is a time-dependent vector field defined on $I \times \mathcal{M}$, the notation

$$\dot{m}(t) = f(t, m(t)) \tag{A.5}$$

makes perfectly sense.

Let us observe that, if

$$f(m) = f^i(t, m)\frac{\partial}{\partial x_i}$$

with respect to the chart $(U, x = \phi)$, (A.5) is equivalent to

$$\frac{dx}{dt}(t) = f_\phi(t, x(t)) \tag{A.6}$$

where f_ϕ is the *the representation of f with respect to the chart* $(U, x = \phi)$, which is defined by

$$f_\phi(t, x) = \left(f^1(t, \phi^{-1}(x)), \cdots, f^n(t, \phi^{-1}(x))\right) \qquad \forall x \in \phi(U).$$

So, via composition with charts, most of the questions for ODE's can be reduced to analogous questions for ODE's on $I\!R^n$. On the other hand, there a lot of issues even in the Euclidean setting that have a chart invariant formulation. For instance the *transport of both vectors and covectors*, which are accounted by *variational* and *adjoint* equations (see below), can be expressed in a chart-independent way.

A.4.1 The Flow-Box Theorem

The Flow-box theorem says the if f is a locally Lipschitz, time-independent, vector field and $m \in M$ is a non-equilibrium point, i.e. $f(m) = 0$, then f *is constant in a neighborhood of* m.

Theorem A.4.2 *Let $\mathcal{M}$ be a manifold and let f be a time-independent vector field on $\mathcal{M}$, of class C^1. Let $m \in M$ be a point such that*

$$f(m) \neq 0.$$

Then there exists a chart $(U, x = \phi)$ near m such that

$$f_\phi(x) = \frac{\partial}{\partial x_1} \qquad \forall x \in \phi(U)$$ [12]

The theorem can be extended to locally Lipschitz vector fields. (see [41])

As it is known the thesis is not valid for more than one vector fields (with the same chart). In fact, according to the "multiple flow-box theorem", two vector fields can be rendered *constant* in a chart if and only if their Lie bracket (see below) is equal to zero.

A.5 Distributions, co-distributions, integral submanifolds

A *distribution* [13] Δ on a differentiable manifold M is a set-valued function $q \mapsto \Delta_q$ which maps a point $q \in M$ into a subspace Δ_q of the tangent space T_qM. (If Δ_q has constant dimension a distribution can also be regarded as a vector sub-bundle of the tangent bundle). If n is the dimension of M, a distribution Δ with constant dimension $k \leq n$ is called (completely) *integrable* if in a neighborhood of any point $q \in M$ one can find local coordinates $(x, y) = (x^1, \dots, x^k, y^1, \dots, y^{n-k})$ such that i) each level set $L_{\bar{q}} \doteq \{q \,|\, y(q) = y(\bar{q})\}$ is a k-dimensional submanifold of M, and ii) the tangent space to $L_{\bar{q}}$ at a point $q \in L_{\bar{q}}$ coincides with Δ_q. The sets L_q are called (local) *integral submanifolds* of the distribution Δ. Clearly, the

[12] Of course, the statement is still valid when $\frac{\partial}{\partial x_1}$ is replaced by any constant, non-zero, vector field in the open subset $\phi(U)$.

[13] This notion of *distribution* has nothing to do with the synonymous concept of Schwartz distributions.

question whether integral manifolds do exist is trivial when $k = 1$, for the problem reduces to a question of solutions' existence for an ODE. On the contrary, if $k > 1$, local integral submanifolds do not exist unless a geometrical condition, namely *involutivity*, is verified. As is well-known, the Frobenius Theorem characterizes local integrability by means of involutivity. We recall that a distribution Δ is called *involutive* if for every pair of fields (f, g) *belonging to* Δ,[14] the Lie bracket

$$[f, g] = Dg \cdot f - Df \cdot g$$

belongs to Δ as well.

A.5.1 Lie derivatives, Lie brackets, exterior derivatives

Let us begin by recalling the classical notions of Lie derivative and exterior derivative. [15].

Though the two definitions below are given in coordinate terms, it is trivial to verify they are invariant for local changes of coordinates. In fact, intrinsic definitions can be given as well.

Definition A.5.1 (LIE DERIVATIVES) *Let M be a manifold, let r, s be non-negative integers and let E be the (r, s)-type tensor bundle on M* [16]. *If $T : M \to E$ is a section of E (i.e. a (r, s)-type tensor field) of class C^1 and f is a vector field on M of class C^1, then the Lie derivative $L_f T$ of T along f is the (continuous) section of E defined as follows.*

Let (U, x) be a coordinate chart, and let

$$T(q) = T^{i_1,\dots,i_r}_{j_1,\dots,j_s}(q) \frac{\partial}{\partial x^{i_1}} \otimes \dots \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \dots dx^{j_s} \qquad \forall q \in U\,,$$

where: i)the multi-indexes $i_1, \dots, i_r$ and $j_1, \dots, j_s$ range, respectively, over all r-tuples and s-tuples of elements of $\{1, \dots, n\}$; ii)the real functions $T^{i_1,\dots,i_r}_{j_1,\dots,j_s}$ are of class C^1; and iii)the summation convention is adopted. Then the Lie derivative $L_f T$ is expressed on U by

$$L_f T = W^{i_1,\dots,i_r}_{j_1,\dots,j_s} \frac{\partial}{\partial x^{i_1}} \otimes \dots \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \dots dx^{j_s}\,, \qquad (A.7)$$

[14] We say that a vector field f belongs to Δ if $f(q) \in \Delta_q$ for every $q \in M$.

[15] Extensions of these objects to the case when the involved functions are locally Lipschitz can be found in [41]

[16] This means that, for each $q \in M$,

$$E_q = \underbrace{T_qM \otimes \dots \otimes T_qM}_{r \text{ times}} \otimes \underbrace{T_q^*M \otimes \dots \otimes T_q^*M}_{s \text{ times}}$$

where, for every value of the multi-indexes $(i_1,\dots,i_r)$ *and* $(j_1,\dots,j_s)$*, the map* $W^{i_1,\dots,i_r}_{j_1,\dots,j_s} : U \to I\!R$ *is defined by*

$$W^{i_1,\dots,i_r}_{j_1,\dots,j_s} \doteq f^c \frac{\partial T^{i_1\dots i_r}_{j_1\dots j_s}}{\partial x^c} - \frac{\partial f^{i_1}}{\partial x^c} T^{c\dots i_r}_{j_1\dots j_s} - \dots - \frac{\partial f^{i_r}}{\partial x^c} T^{i_1\dots i_{r-1}c}_{j_1\dots j_s} +$$

$$\frac{\partial f^c}{\partial x^{j_1}} T^{i_1\dots i_r}_{c\dots j_s} + \dots + \frac{\partial f^c}{\partial x^{j_s}} T^{i_1\dots i_r}_{j_1\dots j_{s-1}c} .$$

Definition A.5.2 (LIE BRACKET) *If* f *and* g *are vector field of class* C^1*, the vector field* $[f,g]$ *is defined as*

$$[f,g] \doteq L_f g.$$

$[f,g]$ *is called the Lie bracket of* f *and* g*. In particular, in coordinates* $(x^1,\dots,x^n)$ *one has*

$$[f,g]^i = \frac{\partial g^i}{\partial x^r f^r} - \frac{\partial f^i}{\partial x^r g^r} .$$

We remind that the skew-symmetry of the Lie derivative, $[f,g] = -[g,f]$, from which it follows idempotency: $[f,f] = 0$ for all vector fields f of class C^1 (but see [41] for the Lipshcitz case). Let us also recall the Jacobi's identity, $[f,[g,h]] + [h,[f,g]] + [g,[h,f]] = 0$, valid for all triples f,g,h.

Let n be a positive integer, let M be a n-dimensional manifold, and let $U \subset M$ be open. For every integer h such that $0 \le h \le n$ and every $q \in U$, let Λ^h_q denote the space of skew-symmetric, h-linear forms, on $(T_qM)^h$. Let $\Lambda^h(U)$ be the (U-based) corresponding vector bundle, and, for every $r = 0,1$ let us use

$$\Omega^h_r(U)$$

to denote the set of sections of $\Lambda^h(U)$ that are of class C^r. Namely, $\Omega^h_r(U)$ is the set of h*-forms* on U that are of class C^r. In addition, we use

$$\Omega^h_{0,1}(U) \quad \big(\subset \Omega^h_0(U)\big)$$

to denote the set of h-forms defined on U that are of class $C^{0,1}$, also called *locally Lipschitz h-forms.* In particular, $\Omega^0_1(U)$ and $\Omega^0_{0,1}(U)$ denote the set of real functions defined on U that are, respectively, continuously differentiable and locally Lipschitz.

Let us recall the definition of exterior derivative for a h-form of class C^1.

Definition A.5.3 (Exterior derivative) *Let h be an integer such $0 \leq h \leq n-1$, and let $\omega \in \Omega_1^h(U)$. The exterior derivative $d\omega$ of ω is a $h+1$-form of class C^0 defined as follows: if*

$$\omega(q) = \sum_{\sigma} c_{\sigma_1,\ldots,\sigma_h} dx^{i_1} \wedge \ldots dx^{i_h} \quad {}^{17}$$

is the local expression of ω on a coordinate chart (U', x), then

$$d\omega(q) = \sum_{\sigma} \sum_{r=1}^{n} \frac{\partial c_{i_1,\ldots,i_h}}{\partial x^r} dx^r \wedge dx^{\sigma_1} \wedge \ldots dx^{\sigma_h}, \tag{A.8}$$

for all $q \in U'$.

A.5.2 Distributions and codistributions

For any finite subset $\{v_1, \ldots, v_r\}$ of a real vector space V, let us use $span\{v_1, \ldots, v_r\}$ to denote the linear subspace generated by $\{v_1, \ldots, v_r\}$.

Definition A.5.4 *Let n, k be non-negative integers such that $k \leq n$, and let M be a n-dimensional manifold. By a k-dimensional* distribution *of class C^1 we mean a subset $\Delta \subseteq TM$ such that, for every $\bar{q} \in M$,*

i) $\Delta_{\bar{q}} \doteq \Delta \cap T_{\bar{q}}M$ *is a linear subspace of $T_{\bar{q}}M$ of dimension k,*

ii) there is a neighborhood U of $\bar{q}$ and vector fields $f_1, \ldots, f_k$ of class C^1, defined on U verifying

$$\Delta_q = span\{f_1(q), \ldots, f_k(q)\}$$

for all $q \in U$. Any such set of vector fields is called a local frame *of class C^1 for Δ.*

Definition A.5.5 *Let n, h be non-negative integers such that $h \leq n$, and let M be a n-dimensional manifold. By a h-dimensional* codistribution *of class C^1 we mean a subset $\Theta \subseteq T^*M$ such that, for every $\bar{q} \in M$,*

i) $\Theta_{\bar{q}} \doteq \Delta \cap T^*_{\bar{q}}M$ *is a linear subspace of $T^*_{\bar{q}}M$ of dimension h,*

[17] Of course the coefficients $c_{\sigma_1,\ldots,\sigma_h}$ are functions of class C^r, and the summation is performed over all strictly increasing h-tuples $\sigma = (\sigma_1, \ldots, \sigma_h)$ with values in $\{1, \ldots, n\}$.

ii) there is a neighborhood U of $\bar{q}$ and 1-forms $\omega^1, \ldots, \omega^h$ of class C^1, defined on U verifying

$$\Theta_q = span\left\{\omega^1(q), \ldots, \omega^h(q)\right\}$$

for all $q \in U$. Any such set of 1-forms is called a local frame *of class C^1 for Θ.*

We shall be mainly concerned with distributions [resp. codistributions] of class $C^{0,1}$, which we also call Lipschitz distributions [resp. codistributions].

Definition A.5.6 *If Δ is a distribution on a manifold M, and f is a vector field defined on a subset $M' \subseteq M$, we shall say that f belongs to Δ if, for every $q \in M'$, $f(q) \in \Delta_q$. In a similar way, we define the notion of a 1-form belonging to a codistribution.*

Definition A.5.7 *Let M and Δ as in Definition A.5.4. The $(n-k)$-dimensional codistribution $\Delta^{\mathfrak{t}}$ defined by*

$$\Delta_q^{\mathfrak{t}} = \{\omega \in T^*M \quad | \quad \langle \omega, v\rangle = 0 \;\; \forall v \in \Delta_q\} \qquad \forall q \in M$$

is called the annihilating codistribution of Δ .

Remark A.5.8 By the Implicit Function Theorem, a distribution Δ is of class C^1 if and only if the codistribution $\Delta^{\mathfrak{t}}$ is of class C^1.

A.5.3 Involutivity, commutativity, and integrability

Definition A.5.9 *(Involutivity of distributions) Let n, k be non-negative integers such that $k \le n$, let M be a n-dimensional manifold, and let Δ be a k-dimensional distribution on M of class C^1. We say that Δ is* involutive *if, for every pair of vector fields f and g (of class C^1) belonging to Δ, one has*

$$[f, g](q) \in \Delta_q \qquad \forall q \in M. \tag{A.9}$$

Definition A.5.10 *(Involutivity of families of vector fields) Let n be a non-negative integer, and let M be a n-dimensional manifold. Let $U \subseteq M$ be an open subset and let $\mathcal{V}$ be a family of vector fields of class C^1 on U. We say that $\mathcal{V}$ is* involutive *if, for every pair $f, g \in \mathcal{V}$ and every $q \in M$, one has*

$$[f, g](q) \in span\{h(q) \,|\, h \in \mathcal{V}\}. \tag{A.10}$$

Let us give the notion of *commutativity* for vector fields.

Definition A.5.11 *(Commutativity of vector fields) Let n be a non-negative integer let M be a n-dimensional manifold. Let $U \subseteq M$ be an open subset and let f and g vector fields on U. We say that f and g* commute *if*

$$[f,g](q) = \{0\} \quad \forall q \in U\,. \tag{A.11}$$

Finally, let us give the notion of integrability for distributions.

Definition A.5.12 *(*Integrability*) Let n, k be non-negative integers such that $k \leq n$. Let Δ be a k-dimensional distribution of class C^r ($r \geq 1$) on a n-dimensional manifold M (of class C^{r+1}). One says that Δ is* completely integrable *if for each $\in M$ there exist a neighborhood U of q, open subsets $X \subseteq I\!R^k, Y \subseteq I\!R^{n-k}$, and a coordinate chart $(U,(x,y))$ of class C^r verifying $(x,y)(U) = X \times Y$, such that*

$$T_{\bar{q}}\Big((x,y)^{-1}\Big)(X \times \{\bar{y}\}) = \Delta_{\bar{q}},$$

for every $\bar{q} \in U$, where $\bar{y}$ is such that $(x,y)(\bar{q}) = (\bar{x},\bar{y})$, for a suitable $\bar{x}$.

The sets $(x,y)^{-1}(X \times \{\bar{y}\})$ are called local leaves *of Δ, and the family of such leaves is sometimes referred as the* foliation *corresponding to Δ.*

$\diamondsuit$

Remark A.5.13 *The C^r-regularity of the submanifolds $(x,y)^{-1}(X \times \{\bar{y}\})$ is guaranteed by the Implicit Function Theorem. (See [41] for the more involved case of Lipschitz distributions).*

$\diamondsuit$

A.5.4 Frobenius Theorem

We are going to state Frobenius theorem in the form of three conditions, each of which is equivalent to complete integrability. The first two of these conditions involve vector fields and their Lie brackets. The third one deals only with the forms spanning the annihilating distribution and their exterior derivatives. Finally, the fourth one involves both forms and vector fields (and the Lie derivative of the former along the latter).

Theorem A.5.14 *Let n, k be non-negative integers such that $k \leq n$, and let Δ be a k-dimensional distribution of class C^r $(r \geq 1)$ on a n-dimensional manifold M (of class C^{r+1}. Then the following conditions are equivalent:*

1) *Δ is involutive.*

2) *For every $\bar{q} \in M$ there exists an open neighborhood $U \subseteq M$ of $\bar{q}$ local frame $\{g_1, \ldots, g_k\}$ of Δ such that for every $i, j = 1, \ldots, k$, g_i and g_j commute, i.e., $[g_i, g_j] = 0$ on U.*

3) *If $U \subseteq M$ and $\{\omega^1, \ldots, \omega^{n-k}\}$ are an open subset and a frame of Δ^t on U, respectively, then, for every $\alpha = 1, \ldots, n-k$ and for every $q \in U$, one has*

$$(d\omega^\alpha)(q) = \theta_1 \wedge \omega^1(q) + \cdots + \theta_{n-k} \wedge \omega^{n-k}(q),$$

for suitable one-forms $\theta_1, \ldots, \theta_{n-k}$.

4) *If f is a vector field belonging to Δ and ω is 1-form (of class C^r) belonging to Δ^t, then, for every $q \in M$, one has*

$$(L_f \omega)(q) \in \Delta^t(q) \ .$$

5) *Δ is completely integrable.*

A non smooth extension of this Theorem can be found e.g. in [41].

Appendix B

Lagrangian and Hamiltonian type equations

The aim of this Appendix –which is formally disconnected from the remaining part of these notes– is to recall some very basic facts on Lagrangian and Hamiltonian equations. We do not even mean to give an "introduction" to Lagrangian or Hamiltonian Mechanics. Of course there is plenty of textbooks ant treatises, some of them quite celebrated. Rather, these few pages might be regarded as a *appetizer* for much more substantial tools of Mechanics, some of which are involved in the previous chapters. Everything is here treated on Euclidean spaces, and no configuration manifolds or their tangent and cotangent bundle show up. Still, the main focus lies on coordinate-invariance.

B.1 Parametrized Fenchel-Legendre transforms

Let V be a finite-dimensional real vector space, and let V^* the dual space of V, and let $L : V \to I\!R$ be a function.

The function $L^* : V^* \to I\!R \cup \{+\infty\}$ defined by

$$L^*(p) \doteq \sup_{v \in I\!R^n} \{p \cdot v - L(v)\}$$

is called the Legendre-Fenchel transform of L. Being defined as the sup of a family of linear functions L^* is a convex map. If L is strictly convex then L^* is strictly convex as well. It is well-known that if L is lower bounded, the map

$$(L^*)^* : V \to I\!R \cup \{+\infty\}$$

is the convex envelope of L, that is

$$(L^*)^* \geq f \qquad \forall f : V \to I\!R \quad \text{convex}, f \leq L$$

In particular, if L is lower bounded, $(L^*)^*$ is real valued, and $L = (L^*)^*$ is and only if L is convex. Moreover, if

If Γ is a set and $L : \Gamma \times V \to I\!R$ is a map, when there will be no danger of confusion we shall use L^* to denote the map Fenchel-Legendre transform made with respect to the variable $v \in V$, namely

$$L^*(\gamma, p) \doteq \sup_{v \in I\!R^n} \Big\{ p \cdot v - L(\lambda, v) \Big\}.$$

Theorem B.1.1 *Let $L : V \to I\!R$ be strictly convex and let us set $H = L^*$. Then*

i) For every $p \in (I\!R^n)^*$, there is a unique $v = v^H(p)$ such that

$$H(p) = p \cdot v^H(p) - L(v^H(p)).$$

Furthermore, if $H(\cdot)$ is differentiable at p, then

$$v^H(p) = \frac{\partial H}{\partial p}(p).$$

ii) For every $v \in I\!R^n$, there is a unique $p = p^L(v)$ such that

$$L(v) = p^L(v) \cdot v - H(p^L(v)).$$

Furthermore, if $L(\cdot)$ is differentiable at v, then

$$p^L(v) = \frac{\partial L}{\partial v}(v).$$

iii) Both p^L and v^H are bijective, and

$$v^H = (p^L)^{-1}.$$

$\diamondsuit$

Let I be an interval , let n a positive integer, and let $\mathcal{O} \subset I\!R^n$. $L : I \times \mathcal{O} \times V \to I\!R$ is a map, we shall use L^* to denote the map Fenchel-Legendre transform made with respect to the variable $v \in V$, namely

$$L^*(t, x, p) \doteq \sup_{v \in I\!R^n} \Big\{ p \cdot v - L(t, x, v) \Big\}.$$

Moreover, setting $H \doteq L^*$, for every $(t,x) \in I \times \mathcal{O}$, we shall use $v^H_{t,x}$ and $p^L_{t,x}$ to denote $v^{H(t,x,\cdot)}$ and $p^{L(t,x,\cdot)}$, respectively.

Hence, in view of Theorem B.1.1, for every $(t,x) \in I \times \mathcal{O}$, one has

$$H(t,x,p) = p \cdot v^H_{t,x}(p) - L(t,x,v^H_{t,x}(p)) \qquad L(t,x,v) = p^L(t,x,v)\cdot v - H(t,x,p^L(v))$$

Furthermore, if $H(t,x,\cdot)$ is differentiable at p, then

$$v^H_{t,x}(p) = \frac{\partial H}{\partial p}(t,x,p)\,,$$

and if $L(t,x,\cdot)$ is differentiable at v, then

$$p^L_{t,x}(v) = \frac{\partial L}{\partial v}(t,x,v)\,.$$

Theorem B.1.2 *Let the map L be continuous and strictly convex in v. If for every (t,v) the map $x \mapsto L(t,x,v)$ is differentiable, then for every (t,p) the map $x \mapsto H(t,x,p)$ is differentiable as well, and*

$$\frac{\partial L}{\partial x}(t,x,v) = -\frac{\partial H}{\partial x}(t,x,p^L_{t,x}(v)) \qquad \forall (t,x,v) \in I\!R \times I\!R^n \times I\!R^n.$$

A case interesting Mechanics is the one when there exist maps $U = U(t,x)$ and $T = T(t,x,v)$ such that

$$L(t,x,v) = -U(t,x) + T(t,x,v), \tag{B.1}$$

with $v \mapsto T(t,x,v)$ strictly convex for every (t,x). Then

$$H(t,x,p) \doteq L^*(t,x,p) = U(t,x) + T^*(t,x,p).$$

In particular, if

$$T(t,x,v) = \frac{1}{2} v^{\mathtt{t}} G(t,x) v + A(t,x) v + \frac{1}{2} B(t,x),$$

where, for every (t,x), $G(t,x)$ $A(t,x)$, and $B(t,x)$ are a $n \times n$, positive definite matrix, a $1 \times n$ covector, and a positive real number, respectively, then

$$H = U(t,x,v) + T^*(t,x,p) = U(t,x,v) + \frac{1}{2} p^{\mathtt{t}} G^{-1} p - A G^{-1} p^{\mathtt{t}} + \frac{1}{2} A G^{-1} A^{\mathtt{t}}$$

From Theorem B.1.2 we obtain

Corollary B.1.3 *If G, A, B are differentiable in x, then ($x \mapsto G^{-1}$ is differentiable as well, and)*

$$\frac{\partial L}{\partial x}(t,x,v) = -\frac{\partial U}{\partial x} + \frac{1}{2}p\frac{\partial G^{-1}}{\partial x}p - \frac{\partial AG^{-1}}{\partial x}p + \frac{1}{2}\frac{\partial AG^{-1}A^t}{\partial x}$$

for all $(t,x,v) \in I\!R \times I\!R^n \times I\!R^n$, where we have set $p = p^L_{t,x}(v)$

B.2 Lagrangian and Hamiltonian type equations

Let $L = L(t,x,v), F = F(t,x,v)$ be maps defined on $I \times \mathcal{O} \times I\!R^n$, where $\mathcal{O}$ is an open subset of $I\!R^n$. Let us define the Lagrangian type equation corresponding to the the *Lagrangian* L and the *force* F as the following (implicit) second order ODE:

$$\frac{d}{dt}\frac{\partial L}{\partial v}(t,x,\dot{x}) = \frac{\partial L}{\partial x}(t,x,\dot{x}) + F(t,x,\dot{x}) \tag{B.2}$$

If $H = H(t,x,p)$ and $\check{F} = \check{F}(t,x,p)$ are with domain equal to $I \times \mathcal{O} \times (I\!R^n)^*$ let us define the Hamiltonian type equation corresponding to the the *Hamiltonian* L and the *force* $\check{F}$ as the first order ODE:

$$\begin{cases} \dot{x} = \frac{\partial H}{\partial p}(t,x,p) \\ \\ \dot{p} = -\frac{\partial H}{\partial x}(t,x,p) + \check{F}(t,x,p) \end{cases} \tag{B.3}$$

Remark B.2.1 Of course one needs that the derivatives involved in these equations exist in order they can be well defined, though various generalizations can be considered (see.e.g. [**?**], ..., and Remark **??**). Moreover, existence and uniqueness results on ODE's can be applied at various levels of generality. We shall discuss this aspect of the question later, and in the following we shall assume sufficient regularity in order to retain a higher clarity in establishing the relation between (B.2) and (B.3).

Using $\mathbf{I}_n$ to denote the $n \times n$ unit matrix, and setting

$$\mathbf{J} \doteq \begin{pmatrix} 0_n & \mathbf{I}_d \\ \\ -\mathbf{I}_n & 0_d \end{pmatrix}$$

one can write the Hamiltonian type equation (B.3) as

$$\begin{pmatrix} \dot{x} \\ \\ \dot{p} \end{pmatrix} = \mathbf{J} \cdot \nabla H(t,q,p) + \begin{pmatrix} 0 \\ \\ \check{F}(t,x,p) \end{pmatrix}, \tag{B.4}$$

where $\nabla H(t,q,p)$ denotes the *phase-space gradient* of H, that is, the column vector whose entries are the partial derivatives with respect to $(q^1,\ldots,q^n,p_1,\ldots,p_n)$. The vector field

$$X_H \doteq \begin{pmatrix} \frac{\partial H}{\partial p} \\ \\ -\frac{\partial H}{\partial x} \end{pmatrix} = \mathbf{J}\cdot\nabla H$$

is called *the Hamiltonian vector field corresponding to* H. Notice that X_H is in fact a vector field, possibly dependent on time.

The following result establishes the relationship between Lagrangian-type and Hamilton-type equations, when the Lagrangian is strictly convex (in v) and $H = L^*$.

Theorem B.2.2 *Let L be function of class C^2, strictly convex in v for every $(t,x)\in I\times\mathcal{O}$. Let F be a continuous map. Let $J\subset I$ be a subinterval and let $x : J\to\mathcal{O}$ be a solution of (B.2) of class C^2. If we define H and $\check{F}$ by setting*

$$H(t,x,p)\doteq L^*(t,x,p)\qquad \check{F}(t,x,p)\doteq F(t,x,v^L_{t,x}(p)),$$

then H is of class C^1 and the map

$$(x,p)(t)\doteq\Big(x(t),p^L_{t,x(t)}(\dot{x}(t))\Big) \tag{B.5}$$

is a solution of (B.3).

Conversely, let H be a map of class C^2. If $(x,p) : J\to\mathcal{O}\times\mathbb{R}^{n*}$ is a solution of class C^1 of (B.3), then the map $x(\cdot)$ is in fact of class C^2 and is a solution of (B.2), with

$$L(t,x,v)\doteq H^*(t,x,p)\quad F(t,x,p)\doteq\check{F}(t,x,p^H_{t,x}(p)).$$

Proof. Let $x : J\to\mathcal{O}$ a solution of (B.2) of class C^2 and let us set $(x,p)(\cdot)$ be the map defined in (B.5). By Theorem B.1.1 we obtain, for every $t\in J$,

$$p(t)=\frac{\partial L}{\partial v}\Big(t,x(t),\dot{x}(t)\Big),\quad \dot{x}(t)=v^H_{t,x}(p(t))=\frac{\partial H}{\partial p}\Big(t,x(t),p(t)\Big) \tag{B.6}$$

so that

$$\frac{\partial L}{\partial x}\Big(t,x(t),\dot{x}(t)\Big)+F(t,x(t),\dot{x}(t))=\frac{d}{dt}\left(\frac{\partial L}{\partial v}\Big(t,x(t),\dot{x}(t)\Big)\right)=\dot{p}(t)\quad\forall t\in J.$$

Therefore, by Theorem B.1.2 and the definition of $\check{F}$, we obtain

$$\dot{p}(t)=-\frac{\partial H}{\partial x}\Big(t,x(t),p(t)\Big)+\check{F}\Big(t,x(t),p(t)\Big). \tag{B.7}$$

By (B.6) and (B.7), the first part of the theorem is proved.

Now let $(x,p) : J \to \mathcal{O} \times I\!R^{n*}$ be a solution of class C^1 of (B.3). Since $L^* = (H^*)^* = H$, by Theorem B.1.1 we have, for every $t \in J$,

$$\dot{x} = \frac{\partial H}{\partial p}\Big(t, x(t), p(t)\Big) = v^H_{t,x(t)}(p(t)).$$

so that, in particular, $x(\cdot)$ is of class C^2. In addition,

$$p(t) = \left(v^L_{t,x(t)}\right)^{-1}(\dot{x}(t)) = \left(p^L_{t,x(t)}\right)(\dot{x}(t)) = \frac{\partial L}{\partial v}\Big(t, x(t), \dot{x}(t)\Big)$$

Hence, by Theorem B.1.2 and the definition of $\check{F}$, we obtain

$$\tfrac{d}{dt}\left(\tfrac{\partial L}{\partial v}(t, x(t), \dot{x}(t))\right) = \dot{p}(t) = -\tfrac{\partial H}{\partial x}\Big(t, x(t), p(t)\Big) + \check{F}(t, x(t), p(t)) =$$

$$-\tfrac{\partial L}{\partial x}\Big(t, x(t), p(t)\Big) + F(t, x(t), \dot{x}(t),$$

namely $x(\cdot)$ is a solution of (B.2).

B.2.1 Invariance by coordinates changes

A striking property of both Lagrangian type equations and Hamiltonian type equations is that they are *invariant with respect to changes of coordinates*, as it will be made precise in Theorem **??**. In particular, this is essential in Mechanics, e.g. when one derives the equations in the presence of ideal constraints. Of course, this invariance says that there are intrinsic equations of which Lagrangian type equations and Hamiltonian type equations are the coordinate expressions, and that these equations *can be written on manifolds.* [1]

Invariance for Lagrangian type equations

Let $\mathcal{O}' \subseteq \mathcal{O}$ be an open subset, and let $y : J \times \mathcal{O} \to \mathcal{O}'$ be a *time dependent change of coordinates* . More precisely, y is of class C^2, and for every t, the map $y(t, \cdot)$ is a diffeomorphism with an inverse of class C^2. In particular,

[1] An intrinsic approach is part of well established theories, respectively the *geometry of sprays* and symplectic geometry. We refer the reader interested into an account of these geometrical theories to the vast specialized literature.

this implies that the Jacobian matrix $\frac{\partial y}{\partial x}(t,x)$ is non-singular at each $x \in \mathcal{O}$. Let us use $x(t,\cdot)$ to denote the inverse of $y(\cdot,t)$

Let $L, F : I \to \mathcal{O} \times I\!R^n$ be continuous functions and let L be of class C^2, and, for all $(t,y,w) \in J \times \mathcal{O}' \times I\!R^n$, let us set

$$\tilde{L}(t,y,w) \doteq L\Big(t, x(t,y), \frac{\partial x}{\partial y} \cdot w\Big), \qquad \tilde{F}(t,y,w) \doteq \frac{\partial x}{\partial y} \cdot F\Big(t, y, \frac{\partial x}{\partial y} \cdot w\Big).$$

Proposition B.2.3 *Let $J \subset I$ be a subinterval and let $\hat{x} : J \to \mathcal{O}$ be a solution of the corresponding Lagrangian type equation*

$$\frac{d}{dt}\frac{\partial L}{\partial v}(t,x,\dot{x}) = \frac{\partial L}{\partial x}(t,x,\dot{x}) + F(t,x,\dot{x})$$

Then the map

$$\hat{y} : J \to \mathcal{O}' \qquad \hat{y}(t) \doteq y(t, \hat{x}(t))$$

verifies the Lagrangian type equation

$$\frac{d}{dt}\frac{\partial \tilde{L}}{\partial w}(t,y,\dot{y}) = \frac{\partial \tilde{L}}{\partial y}(t,y,\dot{y}) + \tilde{F}(t,y,\dot{y}).$$

Proof. For every $t \in J$ and for all $r = 1, \dots, N$, one has

$$\frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial w^r}\right) = \frac{d}{dt}\left(\frac{\partial x^s}{\partial y^r} \cdot \frac{\partial L}{\partial v^s}\right) =$$

$$\frac{\partial x^s}{\partial y^r} \cdot \frac{d}{dt}\left(\frac{\partial L}{\partial v^s}\right) + \frac{\partial L}{\partial v^s}\left(\frac{\partial^2 x^s}{\partial t \partial y^r} + \frac{\partial^2 x^s}{\partial y^j \partial y^r} \cdot \dot{\hat{y}}^j(t)\right)$$

and

$$\frac{\partial \tilde{L}}{\partial y^r} = \frac{\partial x^s}{\partial y^r} \cdot \frac{\partial L}{\partial x^s} + \frac{\partial L}{\partial v^s}\left(\frac{\partial^2 x^s}{\partial t \partial y^r} + \frac{\partial^2 x^s}{\partial y^j \partial y^r} \cdot \dot{\hat{y}}^j(t)\right)$$

where we mean that functions of (t,x,v), (t,y,w), (t,x), and (t,y) are calculated in $(t, \hat{x}(t), \dot{\hat{x}}(t))$, $(t, \hat{y}(t), \dot{\hat{y}}(t))$, $(t, \hat{x}(t))$, and $(t, \hat{y}(t))$, respectively, and we have used the relation $\dot{\hat{x}}(t) = \frac{\partial x}{\partial t} + \frac{\partial x}{\partial y^\ell} \cdot \dot{\hat{y}}^\ell(t)$.

◊

Invariance for Hamiltonian type equations

An analogous result holds true for Hamiltonian type equations, even though the new Hamiltonian is not merely the old one composed with the inverse transformation $x(\cdot, y)$ (and theone induced on the adjoint variables). Here we shall give an ad hoc proof of this invariance. In order to better understand this invariance — which finds its proper environment in the classical theory of canonical and symplectic transformations— the interested reader can refer to classical text-books.

Let $\mathcal{O}' \subset \mathcal{O}$ be an open subset and let $y : I \times \mathcal{O} \to \mathcal{O}'$ a time-dependent change coordinates with the same properties as the one considered above. For every $t \in I$ let $x(t, \cdot)$ denote the inverse of $y(t, x)$. Let $H, \check{F} : I \to \mathcal{O} \times (I\!R^n)^*$ be continuous functions and, in addition, let H be of class C^1. Define $\tilde{H}$ and $\tilde{\check{H}}$ by setting for all $(t, y, \pi) I \times \mathcal{O} \times (I\!R^n)^*$

$$\tilde{H}(t, y, \pi) \doteq \pi \cdot \frac{\partial y}{\partial t}(t, x(t, y)) + H\left(t\,,\, x(t, y)\,,\, \pi \cdot \left(\frac{\partial y}{\partial x}\right)^{\mathtt{t}} (x(t, y))\right)$$

$$\tilde{\check{F}}(t, y, \pi) \doteq \frac{\partial x}{\partial y}(t, y) \cdot \check{F}\left(t\,,\, x(t, y)\,,\, \pi \cdot \left(\frac{\partial y}{\partial x}\right)^{\mathtt{t}} (x(t, y))\right)$$

Proposition B.2.4 *Let $J \subset I$ be a subinterval and let $(\hat{x}, \hat{p}) : J \to \mathcal{O} \times (I\!R^n)^*$ be a solution of the Hamiltonian type equations*

$$\begin{cases} \dot{x} = \frac{\partial H}{\partial p}(t, x, p) \\ \\ \dot{p} = -\frac{\partial H}{\partial x}(t, x, p) + \check{F}(t, x, p) \end{cases} \tag{B.8}$$

Then the map $(\hat{y}, \hat{\pi}) : J \to \mathcal{O}' \times (I\!R^n)^*$ defined by

$$\Big(\hat{y}(t), \hat{\pi}(t)\Big) \doteq \left(y(t, \hat{x}(t))\,,\, \hat{p}(t) \cdot \left(\frac{\partial x}{\partial y}\right)^{\mathtt{t}} (t, \hat{y}(t))\right)$$

verifies the Hamilton type equation

$$\begin{cases} \dot{y} = \frac{\partial \tilde{H}}{\partial \pi}(t, y, \pi) \\ \\ \dot{\pi} = -\frac{\partial \tilde{H}}{\partial y}(t, y, \pi) + \tilde{\check{F}}(t, y, \pi)\,, \end{cases}$$

We shall derive this Proposition B.2.4 from a more general result , namely Corollary B.2.1 below. Let us recall the notion of *symplectic transformation*: a differentiable map $(y, \pi) = (y(x, p), \pi(x, p)$ from $\mathcal{O} \times I\!R^n$ onto $\mathcal{O}' \times I\!R^n$ is called symplectic if, for every $(x, p) \in \mathcal{O} \times I\!R^n$, one has

$$\frac{\partial(y, \pi)}{\partial(x, p)} \mathbf{J} \left(\frac{\partial(y, \pi)}{\partial(x, p)}\right)^{\mathtt{t}} = \mathbf{J}$$

Lemma B.2.5 *If the map $(y, \pi) = (y(x,p), \pi(x,p)$ is symplectic then for every differentiable map $H = H(t,x,p)$ one has*

$$X_{\tilde{H}} = \frac{\partial(y,\pi)}{\partial(x,p)} \cdot X_H$$

where

$$\tilde{H}(t,y,\pi) = H\left(t, x(y,\pi), \left(\frac{\partial(y,\pi)}{\partial(x,p)}\right)^t \cdot \pi\right)$$

(Of course, the matrix $\left(\frac{\partial(y,\pi)}{\partial(x,p)}\right)^t$ is calculated at $(x,p) = (x(y,\pi), p(y,\pi))$.)

Proof. Let the transformation $(x,p) \mapsto (y,\pi)$ be symplectic. Then

$$X_{\tilde{H}} = \mathbf{J} \cdot \nabla \tilde{H} = \mathbf{J} \cdot \left(\frac{\partial(y,\pi)}{\partial(x,y)}\right)^{\mathbf{t}} \cdot \nabla H = \frac{\partial(x,p)}{\partial(y,\pi)} \cdot \mathbf{J} \cdot \nabla H = \frac{\partial(x,p)}{\partial(y,\pi)} \cdot X_H$$

Remark B.2.6 *The converse of the implication in Lemma B.2.5 holds true as well.*

Corollary B.2.7 *Let $\mathcal{A}, \mathcal{A}' \subseteq I\!\!R^n \times (I\!\!R^n)^*$ be open subsets, let $I \subset I\!\!R$ be an interval, and let $\mathcal{A} \ni (x,p) \mapsto (y,\pi) \in \mathcal{A}'$ be a symplectic map. For given H and $\check{F}$ defined on $\mathcal{A}$, a map $(\hat{x},\hat{p}) : I \to \mathcal{A}$ is a solution of the Hamiltonian type equations*

$$\begin{pmatrix} \dot{x} \\ \\ \dot{p} \end{pmatrix} = X_H(t,x,p) + \begin{pmatrix} 0 \\ \\ \check{F}(t,x,p) \end{pmatrix}, \tag{B.9}$$

if and only if the map $(\hat{y},\hat{\pi}) : I \to \mathcal{A}'$ defined by

$$\Big(\hat{y}(t), \hat{\pi}(t)\Big) \doteq \Big(y\big(\hat{x}(t),\hat{p}(t)\big)\,,\, \pi\big(\hat{x}(t),\hat{p}(t)\big)\Big)$$

verifies the Hamilton type equation

$$\begin{pmatrix} \dot{y} \\ \\ \dot{\pi} \end{pmatrix} = X_{\tilde{H}(t,y,\pi)} + \begin{pmatrix} 0 \\ \\ \left(\frac{\partial x}{\partial y}\right)^t \check{F}(t, x(y,\pi), p(y,\pi)) \end{pmatrix}, \tag{B.10}$$

Proof of Proposition B.2.4 We begin we the case when the coordinate transformation is time-independent, so that we shall write $y(x)$ (and

$x(y)$ for its inverse). In view of Corollary it is sufficient to show that the transformation

$$(x,p) \mapsto \left(y(x), \left(\frac{\partial x}{\partial y}\right)^{\mathtt{t}} \cdot p\right)$$

is symplectic. This means we have to check that

$$\begin{pmatrix} \frac{\partial y}{\partial x} & 0_n \\ \left(\frac{\partial^2 x^\ell}{\partial y^j \partial y^r}\frac{\partial y^r}{\partial x^i}p_\ell\right)_{ji} & \left(\frac{\partial x}{\partial y}\right)^{\mathtt{t}} \end{pmatrix} \cdot \mathbf{J} \cdot \begin{pmatrix} \frac{\partial y}{\partial x} & \left(\frac{\partial^2 x^\ell}{\partial y^j \partial y^r}\frac{\partial y^r}{\partial x^i}p_\ell\right)_{ij} \\ 0_n & \left(\frac{\partial x}{\partial y}\right)^{\mathtt{t}} \end{pmatrix} = \mathbf{J}. \tag{B.11}$$

Indeed

$$\begin{pmatrix} \frac{\partial y}{\partial x} & 0_n \\ \left(\frac{\partial^2 x^\ell}{\partial y^j \partial y^r}\frac{\partial y^r}{\partial x^i}p_\ell\right)_{ji} & \left(\frac{\partial x}{\partial y}\right)^{\mathtt{t}} \end{pmatrix} \cdot \mathbf{J} \cdot \begin{pmatrix} \frac{\partial y}{\partial x} & \left(\frac{\partial^2 x^\ell}{\partial y^j \partial y^r}\frac{\partial y^r}{\partial x^i}p_\ell\right)_{ij} \\ 0_n & \left(\frac{\partial x}{\partial y}\right)^{\mathtt{t}} \end{pmatrix} =$$

$$\begin{pmatrix} 0_n & \frac{\partial y}{\partial x} \\ -\left(\frac{\partial x}{\partial y}\right) & \left(\frac{\partial^2 x^\ell}{\partial y^j \partial y^r}\frac{\partial y^r}{\partial x^i}p_\ell\right)_{ij} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial y}{\partial x} & \left(\frac{\partial^2 x^\ell}{\partial y^j \partial y^r}\frac{\partial y^r}{\partial x^i}p_\ell\right)_{ij} \\ 0_n & \left(\frac{\partial x}{\partial y}\right)^{\mathtt{t}} \end{pmatrix}$$

So (B.11) is verified as soon as

$$-\left(\frac{\partial x^i}{\partial y^k}\right) \cdot \left(\frac{\partial^2 x^\ell}{\partial y^j \partial y^r}\frac{\partial y^r}{\partial x^i}p_\ell\right) + \left(\frac{\partial^2 x^\ell}{\partial y^j \partial y^r}\frac{\partial y^r}{\partial x^i}p_\ell\right) \cdot \left(\frac{\partial x^i}{\partial y^k}\right)^{\mathtt{t}} = 0$$

for all $j,k = 1,\ldots,n$. This is straightforward, so the proof of the time-independent case is concluded.

In order to prove the Proposition in the general case, let us set (x^0, x), and let (p_0, p) denote the general variables of $I\!R^{1+n}$ and $(I\!R^{1+n})^*$, respectively and let us consider the Hamiltonian $H' : (I \times \mathcal{O}) \times (I\!R^{1+n})^* \to I\!R$ defined by

$$H'(x^0, x, p_0, p) = p_0 + H(x^0, x, p)$$

, and let us observe that if $(\hat{x}(\cdot), \hat{p}(\cdot))$ is a solution of (B.8) then, setting

$$x^0(t) \doteq t \qquad p_0(t) = -\int_{\bar{t}}^{t} \frac{\partial H}{\partial s}(s, \hat{x}(s), \hat{p}(s))ds$$

, the map $(\hat{x}^0, \hat{x}, \hat{p}_0, \hat{p})(t)$ solves the Hamiltonian type equation

$$\begin{cases} \dot{x}^0 = \frac{\partial H'}{\partial p_0} \\ \dot{x} = \frac{\partial H'}{\partial p} \\ \dot{p}_0 = -\frac{\partial H'}{\partial x^0} \\ \dot{p} = -\frac{\partial H'}{\partial x} + \check{F} \end{cases} \tag{B.12}$$

B.3 Newton's equations as Lagrangian and Hamiltonian type equations

In this section we write Newton equation in two equivalent forms, which we call Lagrangian and Hamiltonian, respectively. This might appear pointless and unnecessarily obscure, especially when we will use the Legendre-Fenchel transform to obtain the Hamiltonian form. Indeed the two forms of Newton equations differ just in the fact that in the Lagrangian form we use velocities, and in the Hamiltonian form we use *momenta*, namely the velocities multiplied by the corresponding masses. Actually this section has essentially the pedagogical goal of showing that the trivial of substituting velocities with velocities-times-masses is nothing the simplest instance of the process that in a more abstract context allows one to pass from the Lagrangian approach to the Hamiltonian one.

B.3.1 Newton's equations

Let us consider a finite number of mass-points $\{P_1, \ldots, P_N\}$ of mass $m_1, \ldots, m_N$, respectively. For every $\mathcal{S} = 1, \ldots, N$ let $x_{\mathcal{S}} = (x^1_{\mathcal{S}}, x^2_{\mathcal{S}}, x^3_{\mathcal{S}}) \in \mathbb{R}^3$ be the coordinates of $P_{\mathcal{S}}$ with respect to a given inertial frame. So, when no constraints are acting on the system, the state space is the vector space $(\mathbb{R}^3)^N$, which we identify with $\mathbb{R}^{3N}$ by setting

$$(x^1, x^2, x^3, x^4, \ldots, x^{3N-1}, x^{3N})^{\mathtt{t}} \doteq (x^1_1, x^2_1, x^3_1, \ldots, x^2_N, x^3_N)^{\mathtt{t}} ,$$

where the apex $^{|}dag$ denotes transposition. (We regard the elements of $\mathbb{R}^{3N}$ as column vectors.) Similarly, if $\{V_1, \ldots, V_N\}$ are the velocities of the points $\{P_1, \ldots, P_N\}$ and for every $\mathcal{S}$ let $(v^1_{\mathcal{S}}, v^2_{\mathcal{S}}, v^3_{\mathcal{S}})$ are the component of $V_{\mathcal{S}}$, we set

$$(v^1, v^2, v^3, v^4, \ldots, v^{3N-1}, v^{3N})^{\mathtt{t}} \doteq (v^1_1, v^2_1, v^3_1, \ldots, v^2_N, v^3_N)^{\mathtt{t}} .$$

Let us consider the $3N \times 3N$ *inertial matrix*

$$G \doteq \begin{pmatrix} m_1 & 0 & 0 & & & & & \\ 0 & m_1 & 0 & & & & & \\ 0 & 0 & m_1 & & & & & \\ & & \cdots\cdots & & & & & \\ & & \cdots\cdots & & & & & \\ & & & & m_N & 0 & 0 \\ & & & & 0 & m_N & 0 \\ & & & & 0 & 0 & m_N \end{pmatrix}$$

which allows us to define the *kinetic energy* $T : I\!R^{3N} \to I\!R$ as the positive definite quadratic form

$$T(v) = \frac{1}{2} v^{\mathsf{t}} G v = \frac{1}{2} \sum_{\mathcal{S}=1}^{N} \left(m_{\mathcal{S}}((v_{\mathcal{S}}^1)^2 + ((v_{\mathcal{S}}^2)^2 + ((v_{\mathcal{S}}^3)^2 \right)$$

For each $\mathcal{S} = 1, \ldots, N$, and every $(t, x, v) \in I\!R \times I\!R^{3N} \times I\!R^{3N}$, let

$$\tilde{\mathbf{F}}_{\mathcal{S}}(t, x, v) = (\tilde{F}_{\mathcal{S}}^1, \tilde{F}_{\mathcal{S}}^2, \tilde{F}_{\mathcal{S}}^3)(t, x, v)$$

denote the *force acting on the mass-point* $P_{\mathcal{S}}$ *at time* t *when the system is in the configuration* x *with velocity* v. Let us set

$$\tilde{\mathbf{F}} = (\tilde{\mathbf{F}}_1, \ldots, \tilde{\mathbf{F}}_N).$$

Component-wise, we have

$$\tilde{\mathbf{F}} = (\tilde{F}^1, \ldots, \tilde{F}^{3N})^{\mathsf{t}} : I\!R \times I\!R^{3N} \times I\!R^{3N} \to I\!R^{3N} \times I\!R^{3N},$$

where

$$\tilde{F}^{3(\mathcal{S}-1)+k} \doteq \tilde{F}_{\mathcal{S}}^k$$

for all $\mathcal{S} = 1, \ldots, N$, $k = 1, 2, 3$. For every $(t, x, v) \in I\!R \times I\!R^{3N} \times I\!R^{3N}$, we call $\mathbf{F}(t, x, v)$ *force acting at time* t *on the system when this has configuration* x *and velocity* v.

According to Newton's second law the evolutions of the points $P_{\mathcal{S}}$ verify the second order evolution equations

$$m_{\mathcal{S}} \ddot{x}_{\mathcal{S}} = \mathbf{F}_{\mathcal{S}}(t, x, v) \qquad SS = 1, \ldots, N\,, \tag{B.13}$$

which can be written as first order equations in the following form:

$$\begin{cases} \dot{x}_{\mathcal{S}} = v_{\mathcal{S}} \\ \\ \dot{v}_{\mathcal{S}} = \frac{\mathbf{F}_{\mathcal{S}}}{m_{\mathcal{S}}}(t, x, v) \qquad SS = 1, \ldots, N\,, \end{cases} \tag{B.14}$$

Equations (B.13) and (B.14) can be rewritten as

$$G\ddot{x} = \tilde{\mathbf{F}}(t, x, \dot{x}). \tag{B.15}$$

and

$$\begin{pmatrix} \dot{x} \\ \\ \dot{v} \end{pmatrix} = \begin{pmatrix} v \\ \\ G^{-1}\tilde{\mathbf{F}}(t, x, v) \end{pmatrix}, \tag{B.16}$$

respectively.

Of course existence and uniqueness properties for equation (B.16) —or, equivalently, (B.15)—depend the assumptions made on $\mathbf{F}$. For instance, if $\mathbf{F}$ is continuous in t and locally Lipschitz with respect to (x, v), uniformly with respect to t, then local existence and uniqueness of a solution to a Cauchy problem is guaranteed. Global solutions exist in the presence of growth assumptions —e.g. sublinearity in (x, v) — or geometrical assumptions — e.g. Nagumo's conditions. It is also possible to relax the assumption on the regularity in t by allowing to be merely measurable in t, as in the classical Caratheodory's conditions.

B.3.2 Newton's equations as Lagrangian type equations

Let $U : I\!R \times I\!R^{3N} \to I\!R$ be a differentiable real function defined on $I\!R^{3N}$ and let $\mathbf{F} : \mathbf{R} \times \mathbf{R^{3N}} \times \mathbf{R^{3N}} \to \mathbf{R^{3N}}$ be a force such that

$$\tilde{\mathbf{F}}(t, x, v) = -\frac{\partial U}{\partial q}(t, x) + \mathbf{F}(\mathbf{t}, \mathbf{x}, \mathbf{v}).$$

$U = U(t, x)$ is called a *potential* acting on the system, and we use $-\frac{\partial U}{\partial q}$ to denote its gradient with respect to x. Of course there exist infinitely many such decompositions, the more significative being suggested by the physical model. The extreme cases occur when U is identically equal to a constant, —that is, $\frac{\partial U}{\partial q}(t, x) = 0$ — or, conversely, when $\tilde{\mathbf{F}}(t, x, v) = 0$. In the latter case, if, in addition, U is time-independent, $\tilde{\mathbf{F}}$ is called a *conservative* force and U is the potential of $\tilde{F}$

The *Lagrangian* L *associated with the potential* U (and the matrix G) is defined by

$$LL(t, x, v) \doteq -U(t, x, v) + T(v), \tag{B.17}$$

By means of the Lagrangian L one can write Newton's equation (B.15) as a Lagrangian type equations:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial x} = \mathbf{F}. \tag{B.18}$$

The corresponding *Lagrangian* first order system is given by

$$\begin{pmatrix} \frac{dx}{dt} \\ \\ \frac{d}{dt}\frac{\partial L}{\partial v} \end{pmatrix} = \begin{pmatrix} G^{-1}\cdot\frac{\partial L}{\partial v} \\ \\ -\frac{\partial L}{\partial x} \end{pmatrix} + \begin{pmatrix} 0 \\ \\ \mathbf{F} \end{pmatrix} \tag{B.19}$$

Of course, system (B.19) can be made explicit in the variable v as well, by simply using the relation $v = G^{-1}\frac{\partial L}{\partial v}$. Indeed this implies implies $\frac{dv}{dt} = G^{-1}\frac{d}{dt}\frac{\partial L}{\partial v}$, so that (B.19) reduces to

$$\begin{pmatrix} \dot{x} \\ \\ \dot{v} \end{pmatrix} = \begin{pmatrix} G^{-1}\cdot\frac{\partial L}{\partial v} \\ \\ -G^{-1}\cdot\frac{\partial L}{\partial x} \end{pmatrix} + \begin{pmatrix} 0 \\ \\ G^{-1}\cdot\mathbf{F} \end{pmatrix}$$

which, of course, coincides with (B.16).

B.3.3 Newton's equations in Hamiltonian form

If we consider the momenta

$$p_{\mathcal{S}} \doteq m_{\mathcal{S}} v_{\mathcal{S}} \qquad ß + 1, \dots, N\,,$$

then Newton's equations of motion can be written as follows:

$$\begin{cases} \dot{x}_{\mathcal{S}} = \frac{p_{\mathcal{S}}}{m_{\mathcal{S}}} \\ \\ \dot{p}_{\mathcal{S}} = \check{\mathbf{F}}_{\mathcal{S}}(t,x,p), \end{cases}$$

where we have set $p = (p_1, \dots, p_{3N})$, with

$$p_{3\mathcal{S}+j} \doteq m_{\mathcal{S}} v^{3\mathcal{S}+j} \qquad \mathcal{S} = 1, \dots, N,\ j = 0, 1, 2,$$

and

$$\check{\mathbf{F}}_{\mathcal{S}}(t,x,p) \doteq \mathbf{F}_{\mathcal{S}}(t, s, \frac{p_1}{m_1}, \dots, \frac{p_{3N}}{m_N}.$$

In other words, one introduce s the linear isomorphism

$$p = v^{\mathsf{t}} G$$

from $I\!R^{3N}$ onto its dual space $(I\!R^{3N})^*$, so that Newton'equation is reduced to the first order differential system

$$\begin{pmatrix} \dot{x} \\ \\ \dot{p} \end{pmatrix} = \begin{pmatrix} G^{-1}p \\ \\ \mathbf{F}(t,x,G^{-1}p) \end{pmatrix} \tag{B.20}$$

Let us observe that

- these equations are Hamiltonian type equations with

$$H(t,x,p) \doteq U(t,x) + \frac{1}{2} p G_{-1} p^{\mathtt{t}} \qquad \check{F}(t,x,p) \doteq \mathbf{F}(t,x,pG^{-1}) \quad \text{(B.21)}$$

- In fact, *equations* (B.20)*are exactly the Hamiltonian type equations corresponding to the Newton's equation in the Lagrangian type form (B.18) as described in Theorem B.2.2.*

  Indeed, as observed in Subsection B.1, if L and H are defined as in **??** and B.21 one has

$$H = L^*$$

Bibliography

[1] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt, *Mathematical Aspects of Classical and Celestial Mechanics*, Third Edition, Springer-Verlag, 2006.

[2] J. P. Aubin and A. Cellina, *Differential Inclusions*, Springer-Verlag, 1984.

[3] J. Baillieul, *The geometry of controlled mechanical systems*, in Mathematical Control Theory, J. Baillieul & J. C. Willems, eds. Springer Verlag, New York, 1998, 322-354.

[4] A. M. Bloch, J. E. Marsden, and D. V. Zenkov, *Nonholonomic dynamics.* Notices Amer. Math. Soc. 52 (2005), 324–333.

[5] A. M. Bloch, *Nonholonomic Mechanics and Control*, Springer Verlag, 2003.

[6] A. Bressan, *On differential systems with impulsive controls*, Rend. Sem. Mat. Univ. Padova, 78, 227-236 (1987).

[7] A. Bressan, *Impulsive control of Lagrangian systems and locomotion in fluids*, Discr. Cont. Dynam. Syst. 20 (2008), 1-35.

[8] A. Bressan, and B. Piccoli, *Introduction to the Mathematical Theory of Control*, A.I.M.S. Series on Applied Mathematics, Vol.2, Springfield MO, 2007.

[9] A. Bressan and F. Rampazzo, *Impulsive control systems with commutative vector fields*, Journal of Optimization Theory and Applications, 71, 67-83 (1991).

[10] A. Bressan and F. Rampazzo, *Impusive control systems without commutativity assumption*, Journal of Optimization Theory and Applications, 81, 435-457 (1994).

[11] A. Bressan and F. Rampazzo, *Moving constraints as stabilizing controls in classical mechanics*, Arch. Rational Mech. Anal., 196-1, 97-141.

[12] A. Bressan and F. Rampazzo, *On differential systems with vector-valued impulsive controls*, Boll. Un. Mat. Ital., (7) 2-B, 641-656 (1988).

[13] A. Bressan and F. Rampazzo, *On systems with quadratic impulses and their application to Lagrangean mechanics*, SIAM J. Control 31 (1993), 1205-1220.

[14] Aldo Bressan, *Hyperimpulsive motions and controllizable coordinates for Lagrangian systems*, Atti Accad. Naz. Lincei, Mem. Cl. Sc. Fis. Mat. Natur., 19, (1991).

[15] Aldo Bressan, *On some control problems concerning the ski or swing*, Atti Accad. Naz. Lincei, Memorie, Serie IX, Vol. I, (1991), 147-196.

[16] Aldo Bressan and M. Motta, A class of mechanical systems with some coordinates as controls. A reduction of certain optimization problems for them. Solution methods. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Mem.* **9** ,Mat. Appl. 2-1 (1993), 5- 30.

[17] Aldo Bressan and M. Motta, *A class of mechanical systems with some coordinates as controls. A reduction of certain optimization problems for them. Solution methods.* Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Mem. 9, Mat. Appl. 2-1 (1993), 5- 30.

[18] F. Bullo and A. D. Lewis, *Geometric Control of Mechanical Systems*, Springer-Verlag, 2004.

[19] F. Cardin and M. Favretti, *Hyper-Impulsive Motion on Manifolds.* Dynamics of continuous, discrete, and impulsive systems 4(1), 1998, 1-21.

[20] J. R. Dorroh and G. Ferreyra, *A multi-state, multi-control problem with unbounded controls*, SIAM J. Control and Optimization, 32, 1322-1332 (1994).

[21] V. A. Dykhta, *Impulse trajectory extension of degenerated optimal control problems*, The Ljapunov functions methods and applications, P. Borne and V. Matrosov eds., J.C. Baltzer AG, Scientific Publishing Co., 103-109 (1990).

[22] O. Hajek, *Book review* on Bull. Am. Math Soc. 12-2, 272-279 (1975).

[23] V. Jurdjevic, *Geometric Control Theory.* Cambridge University Press, Cambridge, 1997.

[24] J. M. Lee, *Riemannian Manifolds. An introduction to Curvature*, Springer, 1997.

[25] M. Levi, *Geometry and physics of averaging with applications*, Physica D 132 (1999), 150-164.

[26] M. Levi, *Geometry of vibrational stabilization and some applications*, Int. J. Bifurc. Chaos 15 (2005), 2747-2756.

[27] M. Levi and Q. Ren, *Geodesics on vibrating surfaces and curvature of the normal family*, Nonlinearity 18 (2005), 2737-2743.

[28] W. S. Liu and H. J. Sussmann, *Limits of highly oscillatory controls and the approximation of general paths by admissible trajectories*, in Proc. 30-th IEEE Conference on Decision and Control, IEEE Publications, New York, 1991, 437-442.

[29] C. Marle, *Géométrie des systèmes mécaniques à liaisons actives*, in Symplectic Geometry and Mathematical Physics, 260–287, P. Donato, C. Duval, J. Elhadad, and G. M. Tuynman Eds., Birkhäuser, Boston, 1991

[30] B. M. Miller, *Conditions for the optimality in problems of generalized control. I. Necessary conditions for optimality*, Automation and Remote Control, 53, (1992).

[31] B. M. Miller, *Optimization of dynamic systems with a generalized control*, Automation and Remote Control, 50, (1989).

[32] B. M. Miller, *The generalized solutions of nonlinear optimization problems with impulse control*, SIAM J. Control and Optimization, 34, 1420-1440 (1996).

[33] B. M. Miller, *The generalized solutions of ordinary differential equations in the impulse control problems*, Journal of Mathematical Systems. Estimation and Control, 4, 385-388 (1994).

[34] M. Motta and F. Rampazzo, *Space-time trajectories of nonlinear systems driven by ordinary and impulsive controls*, Differential and Integral Equations, 8-2, 269-288 (1995).

[35] M. Motta and F. Rampazzo, *Dynamic programming for nonlinear systems driven by ordinary and impulsive controls*, SIAM J. Control and Optimization, (1996).

[36] M. Motta and F. Rampazzo, *Nonlinear systems with unbounded controls and state constraints: a problem of proper extension*, NoDEA - Nonlinear Differential Equations and Applications, 3, 191-21 (1996).

[37] M. Motta and F. Rampazzo, *The value function of a slow growth control problem with state constraints*, J. Math. Syst. Est. and Control, to appear (available electronic manuscript, short version in J. Math. Syst. Est. and Control 7, 375-378 (1997).

[38] M. Motta and F. Rampazzo, *State constrained control problems with neither coercivity nor L^1 bound on the controls*, Appl. Math. and Optim., to appear.

[39] H. Nijmejer and A.J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York, 1990.

[40] F. Pereira, G. Nunes Silva, V. Oliveirac, *Invariance for impulsive control systems*,Avtomat. i Telemekh., 2008:5, 57-71.

[41] F. Rampazzo, *Frobenius-type Theorems for Lipschitz Distributions.* Journal of Differential Equations 43, 2003, 270-300.

[42] F. Rampazzo, *Lecture Notes on Control and Mechanics.* Preprint.

[43] F. Rampazzo, *Lie brackets and impulsive controls: an unavoidable connection*, in Differential Geometry and Control, Proc. Sympos. Pure Math. (AMS, Providence)(1999), 279-296.

[44] F. Rampazzo *On Lagrangian systems with some coordinates as controls.* Atti Accad. Naz. dei Lincei, Classe di Scienze Mat. Fis. Nat. Serie 8, 82, 1988, 685-695.

[45] F. Rampazzo, *On the Riemaniann structure of a Lagrangian system and the problem of adding time-dependent constraints as controls*, Eur. J. Mech., A/Solids, Gauthier-Villars, 10, 405-431 (1991).

[46] F. Rampazzo and C. Sartori, *Hamilton-Jacobi-Bellman equations with fast gradient-dependence.* Indiana Univ. Math. J. 49 (2000), 1043-1078.

[47] F. Rampazzo and C. Sartori, *The minimum time function with unbounded controls*, J. Math. Syst. Est. and Control, to appear.

[48] B. L. Reinhart, *Differential geometry of foliations. The fundamental integrability problem.* Springer-Verlag, Berlin, 1983.

[49] B. L. Reinhart, *Foliated manifolds with bundle-like metrics.* Annals of Math. 69 (1959), 119-132.

[50] R. W. Rishel, *An extended Pontryagin principle for control systems whose control laws contain measures*, SIAM J. Control, 3, 191-205 (1965).

[51] A. V. Sarychev, *Nonlinear systems with impulsive and generalized function controls*, Proc. Conf. on Nonliner Synthesis, Sopron, Hungary, (1989).

[52] W. W. Schmaedeke, *Optimal control theory for nonlinear differential equations containing measures*, SIAM J. Control, 3, 231-279 (1965).

[53] A. N. Sesekin, *Impulse extension in the problem of the optimization of the energy functional*, Automation and Remote Control, (1993).

[54] A. N. Sesekin, *Nonlinear differential equations in the class of functions of bounded variation*, Automtion and Remote Control, (1990).

[55] G. N. Silva and R. B. Vinter, *Necessary Conditions for Optimal Impulsive Control Problems*, SIAM J. Control and Optim., SIAM J. Control and Optim., 35, (1997), 1829-1846.

[56] G. V. Smirnov, *Introduction to the Theory of Differential Inclusions*, American Mathematical Society, Graduate Studies in Mathematics, vol. 41 (2002).

[57] E. D. Sontag, *A Lyapunov-like characterization of asymptotic controllability*, SIAM J. Control & Optim. 21 (1983), 462-471.

[58] E. D. Sontag, *Mathematical Control Theory*, Springer Verlag, New York, 1990.

[59] H. J. Sussman, *On the gap between deterministic and stochastic ordinary differential equations*, Ann. of Probability, 6, 17-41 (1978).

[60] H. J. Sussmann, *Lie brackets, real analyticity and geometric control theory*, in Mathematical Control Theory, R.W. Brockett, R.S. Millmam and H.J. Sussmann eds. Birkhäuser, Boston Inc. (1983), 1-115.

[61] J. L. Synge and A. Schild, *Tensor Calculus*, Dover Publications, New York, 1978.

[62] R. B. Vinter and F. Pereira, *A Maximum Principle for Optimal Processes with Discontinuous Trajectories*, SIAM J. Control and Optim., 26, (1988), 205-229.

[63] P. Wolenski and S. Zabic, *A sampling method and approximation results for impulsive systems*, SIAM Journal on Control and Optimization, Volume 46, Issue 3, 2007, 983-998.