Università degli Studi di Padova

Dipartimento di Matematica
Corso di Laurea Magistrale in Matematica

Tesi di Laurea Magistrale

Length-Minimizing Curves on Sub-Riemannian Manifolds: Necessary Conditions Involving the End-Point Mapping

Francesco Palmurella
Matricola 1058434

Relatore
Roberto Monti

Controrelatore
Davide Vittone

26 Settembre 2014

Anno Accademico 2013/2014

Length-Minimizing Curves on Sub-Riemannian Manifolds: Necessary Conditions Involving the End-Point Mapping

Francesco Palmurella

Contents

Preface

The concept of "geodesic" or, rather, of curve that minimizes the length between two points in a given metric space is easily perceived by intuition and widely used also outside mathematical contexts. Due to its relevance (think just of its role in General Relativity) the topic is broadly studied and subject to intense analysis. Its mathematical formalizations are performed by means of variational principles or through differential equations - whose solution is in general rather implicit. In this thesis, we study such curves in what are called *sub-Riemannian manifolds*.

The basic idea that leads to the definition of a sub-Riemannian structure is the following: given a smooth manifold M and a smooth subbundle $\mathcal{D}$ of TM - equivalently called *smooth distribution* - one is interested in studying curves on M whose tangent vector is prescribed to lay on $\mathcal{D}$. Such curves are called *$\mathcal{D}$-horizontal* or just *horizontal* when $\mathcal{D}$ is understood. For example, given $X_1, \dots X_m$ smooth linearly independent vector fields in $\mathbb{R}^n$, one may want to study curves whose tangent vector lies in $\mathcal{D} = \mathrm{span}_{\mathbb{R}}\{X_1, \dots, X_m\}$. If $\mathcal{D}$ is bracket-generating, and G is a metric on $\mathcal{D}$, $(M, \mathcal{D}, G)$ is naturally a length space and it is possible to speak about length-minimizing curves. The limiting situation when $\mathcal{D} = TM$ coincides then with the Riemannian case. For introductory texts in sub-Riemmanian geometry together with problems and applications, we refer to [**1**], [**2**] and [**3**].

When formalizing the problem, the required basic regularity of a candidate length-minimizing horizontal curve $\gamma : [a, b] \to M$ connecting two fixed points is the Lipschitz regularity (some authors, for example Liu and Sussmann in [**4**] only require the absolute continuity of γ but there is no substantial difference in this choice). Now, one of the fundamental aspects of Riemannian geodesics is that they turn out to be always C^∞-smooth. This follows, for example, because they satisfy the geodesic differential equation, which is a classical ODE with smooth coefficients (see [**5**]) so it is natural to ask whether this is true also in sub-Riemannian manifolds. The question has turned to be surprisingly difficult to answer - in fact, at the present time there is no answer at all. On one hand, we do not know of any example of nonsmooth length-minimizing curve for a sub-Riemmanian manifold, on the other hand we have no general regularity theorems.

What really makes the difference with the classical Riemannian case is the fact that, when characterizing length-minimizing horizontal curves as minimizers of a length functional subject to the necessary constraints, the problem is no longer in the domain of the classical Calculus of Variations but in the more general setting of Control Theory. As is well-known (see [**6**] as a reference text), Euler-Lagrange equations are replaced by the Pontryagin Maximum Principle, which, as the Euler-Lagrange equations do, define the class of *extremal curves* (shortened *extremals*) that is, a set of horizontal curves among which it is possible find length-minimizers. Among these, in many cases it is possible to find a subset of critical extremals, named *abnormal extremals*, that can contain curves which are nonsmooth in some cases (non C^∞, non C^k and so on) and length-minimizing in others. We will see examples of both cases. Abnormal extremals never occur when the sub-Riemannian manifold is effectively a Riemmanian manifold.

The problem of characterizing the regularity of length-minimizing curves has led to many studies and research for minimality criteria, we cite again [**4**] and also [**7**] for a survey on regularity results. In this thesis we deal with necessary conditions that length-minimizing curves must satisfy based on the analysis of the End-Point Mapping. Loosely speaking, given a reference horizontal curve γ, this map measures "variations" both in length and in end-point of a curve obtained by perturbating the tangent vector of γ. Exploiting the fact that, if γ is length-minimizing, it cannot be open at $\dot{\gamma}$, one deduces several necessary conditions for abnormal length-minimizing curves to satisfy. After examining known results (the so-called first and second-order conditions) we derive new necessary conditions for minimality. Our contribution can be considered an introduction to third-order conditions and it is inspired by the approach to second-order conditions presented in [**8**].

In **Chapter 1** we give the notion of sub-Riemanninan manifold $(M, \mathcal{D}, G)$, a triple consisting of a smooth manifold M, a m-dimensional smooth distribution $\mathcal{D}$ of TM and of a metric G on $\mathcal{D}$. We strongly emphasize the case when the manifold is the Euclidean space $\mathbb{R}^n$. In this case, it is possible to think of $\mathcal{D}$ as spanned by m linearly independent vector fields on $\mathbb{R}^n$. Then, we give the definition of $\mathcal{D}$-horizontal curve. If $\mathcal{D}$ is bracket-generating, $(M, \mathcal{D}, G)$ is naturally a length space and we can therefore talk about *length-minimizing horizontal curves*. We show that it is possible to characterize such class of curves in terms of minimizers of a control problem and invoke the Pontryagin Maximum Principle, that leads to the definition of extremal curve.

Extremals divide between *normal* and *abnormal*. We prove that normal extremals are essentially "Riemanninan geodesics", in the sense that, due to the fact that they are characteristic curves of a classical Hamiltonian function, they are all C^∞-smooth and locally length-minimizing.

In **Chapter 2** we deal with rank-two distribution on the Eucidean space. We review some criteria concerning minimality and regularity, we provide examples of nonsmooth abnormal extremals and we present a class of distributions that admit a strictly abnormal length-minimizing curve.

In **Chapter 3** we introduce the End-Point Mapping $\mathcal{F}$, and we prove two groups of necessary conditions for length-minimizing curves. The first one is precisely the thesis of the Maximum Principle, and since it is related with first-order derivatives of $\mathcal{F}$ it assumes the name of first-order conditions. The fundamental tool to deduce these conditions is the Open Mapping Theorem. In the same spirit, since a second-order Open Mapping Theorem is available (it is stated in Appendix B and taken from [**8**]) it is possible to deduce further necessary conditions based on second-order derivatives of $\mathcal{F}$, and, from these, a very neat result, called the Goh Condition, which states that rank-two distributions do not admit strictly abnormal length-minimizing curves.

Inspired by the above approach, in **Chapter 4** we prove a third-order Open Mapping Theorem useful for our purposes. The basic step consists in dealing with the finite-dimensional, corank-one case, namely with a map $F : \mathbb{R}^N \to \mathbb{R}^n$ with $F(0) = 0$ and such that $\operatorname{corank} dF(0) = \dim \operatorname{Coker} dF(0) = 1$. If suitable conditions involving the Hessian and the third derivatives of F at 0 are satisfied, then the map will be open at 0. From this case analogous theorems valid in more general settings (arbitrary corank and infinite-dimension domain) are easily deduced. We finally present the effectiveness of this theorem to our situation by proving that a particular extremal is not length-minimizing.

More precisely, we consider the family of distributions $\mathcal{D} = \mathcal{D}(m)$ in $\mathbb{R}^3$ spanned by the vector fields X_1 and X_2 given by

$$\begin{aligned} X_1 &= \frac{\partial}{\partial x_1} \\ X_2 &= (1 - x_1)\frac{\partial}{\partial x_2} + x_1^m \frac{\partial}{\partial x_3} \end{aligned}$$

for any x in $\mathbb{R}^3$, where m is a positive integer greater or equal than 2. It will be proved that the segment γ long the x_2 axis:

$$\gamma(t) = (0, t, 0)$$

is, for any m, a strictly abnormal curve for $\mathcal{D}(m)$, but it is not length-minimizing at 0 when $m = 3$. Interestingly, the same curve is uniquely length-minimizing at 0 whenever m is even (this result is proved in Chapter 2).

According to our knowledge of the subject, we do not know any alternative way of proving this result.

Ringraziamenti. È con grande piacere che desidero brevemente ringraziare le persone che hanno reso possibile la stesura di questa tesi così come il compimento dei miei studi a Padova.

Grazie agli amici e ai colleghi padovani conosciuti in questi anni per il tempo passato assieme e per i preziosi aiuti, fuori e dentro l'ambito universitario. Grazie ad Andrea, Pietro, Luca, Nicolò, Andrea, Matteo, Francesco, Marco, Enrico, Valentina, Alessandro e a molti, molti altri. Spero di passare - non troppo presto, possibilmente - ancora del tempo assieme a voi.

Desidero esprimere la mia gratitudine al mio relatore, prof. Roberto Monti, per avermi permesso di lavorare con lui ed aver condiviso con me i suoi interessi di ricerca. È stato un piacere per me poter usufruire (speriamo non abusare) dei suoi consigli, delle sue idee e del suo supporto.

La mia più sincera riconoscenza va infine a mia madre, mio padre, mia zia Clara e Giorgio, il cui supporto affettivo, morale e materiale non è mai, e sottolineo mai, venuto meno in questi anni. Grazie a loro ho potuto dedicarmi allo studio senza ulteriori pensieri e trovare sempre, al ritorno da Padova, un'atmosfera d'affetto e amorevoli cure.

Padova, September 2014 *Francesco Palmurella*

Notation and Conventions

We establish some notation about vector spaces and differential operators that we will be using in the following chapters.

0.0.1. Duality, Coimage and Cokernel. Given a real vector space V, we denote the action of covectors over vector with $\langle \cdot, \cdot \rangle$ so that, if v is a vector in V and ξ a covector in V^*, the action of ξ over v is denoted by $\langle \xi, v \rangle$.

Given a linear mapping between real vector spaces $L : X \to Y$, we recall the definitions of Cokernel and Coimage:

$$\operatorname{Coker} L = Y / \operatorname{Im} L$$
$$\operatorname{Coim} L = X / \operatorname{Ker} L.$$

As is well-known from elementary linear algebra, there is a canonical isomorphism between $\operatorname{Coker} L$ and the complementary subspace of $\operatorname{Im} L$ in Y and between $\operatorname{Coim} L$ and the complementary subspace of $\operatorname{Ker} L$ in Y. We shall identify these isomorphic spaces, so that

$$X = \operatorname{Ker} L \oplus \operatorname{Coim} L \quad \text{and} \quad Y = \operatorname{Im} L \oplus \operatorname{Coker} L.$$

The dimension of $\operatorname{Coker} L$ is called *corank* of L.

Finally, when $Y = \mathbb{R}^n$, we have an isomorphism

$$(\operatorname{Im} L)^\perp \simeq \operatorname{Coker} L$$

given by the standard scalar product, which, in coordinates, consists in the transposition of a column vector to a row vector and vice-versa.

0.0.2. Derivatives and Differentials. If X and Y are real normed spaces, $F : X \to Y$ is a mapping, p and v are vectors of X, the derivative of F at p along v, provided it exists, is denoted as usual by $\frac{\partial F}{\partial v}(p)$.

If F is differentiable at p, its differantial or total derivative at p is denoted by $dF(p)$. When evaluating the differential at p on a vector v in X, we write $dF(p)[v]$ or sometimes $\langle dF(p), v \rangle$ in accordance with the notation introduced above. Similarly, we denote by $d^2F(p), d^3F(p), \ldots$ the differential of higher order of F at p provived they exists. When F is of class C^j, for $j \geq 2$, $d^jF(p)$ is a j-linear and symmetric operator, so we shorten expression like $d^jF(p)[v, v, \ldots, v]$ with $d^jF(p)[v^j]$ or $d^jF(p)[v, v, \ldots, v, w]$ with $d^jF(p)[v^{j-1}, w]$. We say that F is "smooth" when is of class C^∞, that is, d^jF exists for any positive integer j and are continuous mappings. When $X = \mathbb{R}^N$ and $Y = \mathbb{R}^n$ it is possible to talk about Jacobian matrix at a point p, which we denote by $DF(p)$.

We denote by $\operatorname{Hess} F(p)$ the Hessian of F at p, namely the quadratic form defined for any v in X as

$$\operatorname{Hess} F(p)[v] = d^2F(p)[v, v].$$

For more notational conventions about the Hessian see Appendix B.

When F is "time-dependent", that is, its domain is $[a,b] \times X$ for some real interval $[a,b]$, or it is a curve, we prefer to denote the derivative with repsect to $t \in [a,b]$ with $\dot{F}$, instead of $\frac{\partial F}{\partial t}$.

Finally, a point p is *critical* for F if $dF(p)$ is not surjective as a linear mapping from X to Y.

0.0.3. Commutators. Given a smooth n-dimensional manifold M and two smooth vector fields X and Y, their *Commutator*, or Lie Bracket, is the vector field defined by

$$[X,Y] = XY - YX. \tag{0.1}$$

When $M = \mathbb{R}^n$, or in local coordinates, vector fields can be regarded as functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ and we have the following formula:

$$[X,Y](p) = dY(p)[X(p)] - dX(p)[Y(p)] \tag{0.2}$$

for any p in $\mathbb{R}^n$ or the local chart. To deduce it it is sufficient to compute (0.1) in coordinates.

0.0.4. Curves on the Cotangent Bundle. Given a curve over the Euclidean space $\gamma : [a,b] \to \mathbb{R}^n$, we denote its components by $(\gamma_1(t), \ldots, \gamma_n(t))$. A curve in the cotangent space along γ, $\lambda : [a,b] \to T^*\mathbb{R}^n$ is typically identified with its fibered components, denoted with $\lambda(t) = (\lambda_1(t), \ldots, \lambda_n(t))$. If X is a vector field on $\mathbb{R}^n$, the action of λ along X is denoted by

$$\langle \lambda(t), X(\gamma(t)) \rangle = \sum_{i=1}^{n} \lambda_i(t) X_i(\gamma(t))$$

in accordance with the above notation.

CHAPTER 1

Sub-Riemannian Geometry and Extremal Curves

We introduce sub-Riemannian manifolds and the notion of extremal curve. Though most of the concepts in this chapter are suitable for a generic smooth manifold M, we are interested in a local analysis of these objects and so we will emphasize the case $M = \mathbb{R}^n$. For a more detailed introduction to sub-Riemannian geometry together with problems and applications, we refer to [**1**], [**2**] and [**3**].

Definition 1.1. Let m, n be positive integers with $m \leq n$ and let M be a smooth manifold of dimension n. A *smooth distribution of dimension* m, $\mathcal{D}$, is a m-dimensional smooth subbundle of TM, that is, a map that assigns to each p in M a dimensional subspace of T_pM of dimension m such that, for any point p, it is possible to find a *local basis for* $\mathcal{D}$ consisting of m smooth vector fields $X_1, \dots, X_m$ that span $\mathcal{D}$ in a neighborhood of p.

When $M = \mathbb{R}^n$, it is possbile to choose a global basis and so we may directly speak of distribution spanned by $X_1, \dots, X_m$.

We define by induction the following subbundles of TM: $\mathcal{D}^1 = \mathcal{D}$ and, if $i \geq 1$, $\mathcal{D}^{i+1}$ is the linear subspace spanned by the commutators (see Notation and Conventions, 0.0.3) of $\mathcal{D}$ and $\mathcal{D}^i$. For example, if $X_1 \dots X_m$ is a local basis for $\mathcal{D}$, we have the following local expression for $\mathcal{D}^2$:

$$\mathcal{D}^2 = \operatorname{span}_{\mathbb{R}}\{[X_i, X_j] : 1 \leq i, j \leq m\}$$

and for $\mathcal{D}^3$:

$$\mathcal{D}^3 = \operatorname{span}_{\mathbb{R}}\{[X_i, [X_j, X_k]] : 1 \leq i, j, k \leq m\}.$$

We set for $i \geq 1$ $\mathcal{L}^i = \mathcal{D}^1 + \dots + \mathcal{D}^i$ and for p in M, $\mathcal{D}^i(p)$ and $\mathcal{L}^i(p)$ the sets of vectors in T_pM belonging to $\mathcal{D}^i$ and $\mathcal{L}^i$ respectively.

$\mathcal{D}$ is called *bracket-generating*, or *completely non-integrable*, if for all $p \in M$ there exists a positive integer $k = k(p)$ such that $\mathcal{L}^k(p) = T_pM$. The minimum k for which $\mathcal{L}^k(p) = T_pM$ for any p in M is called *step* of the distribution $\mathcal{D}$.

Definition 1.2. Let $I = [a, b]$ be a connected, compact interval of $\mathbb{R}$. A Lipschitz-continuous curve $\gamma : I \to M$ is said to be $\mathcal{D}$*-horizontal*, or simply *horizontal* when $\mathcal{D}$ is understood, if its tangent vector $\dot{\gamma}(t)$ belongs to $\mathcal{D}$ for a.e. t in I. In local coordinates, this means that we have a m-tuple of functions $h = (h_1, \dots, h_m) \in L^\infty([a, b], \mathbb{R}^m)$, called *control of* γ, such that

$$\dot{\gamma}(t) = \sum_{j=1}^{m} h_j(t) X_j(\gamma(t)) \quad \text{for a.e. } t \in I. \tag{1.1}$$

Recall that a *metric on a bundle* $\mathcal{D} \subseteq TM$ is a 2−covariant, strictly positive and symmetric tensor on $\mathcal{D}$. In other words, it is a differentiable function $p \mapsto G_p(\cdot, \cdot)$ from M to $\mathcal{D}^* \otimes \mathcal{D}^*$ such that $G_p(v, w) = G_p(w, v)$ and $G_p(v, v) \geq 0$ for all $v, w \in T_pM$ and $G_p(v, v) = 0$ if and only if $v = 0$.

DEFINITION 1.3. A *sub-Riemannian manifold* is a triple $(M, \mathcal{D}, G)$ where M is a smooth manifold, $\mathcal{D}$ is a bracket-generating distribution of TM and G is a metric on $\mathcal{D}$.

The tensor G induces a norm on $\mathcal{D}$, called *sub-Riemannian norm induced by* G, defined for any $v \in T_pM$ by

$$||v||_G = G_p(v, v)^{1/2}.$$

Given a $\mathcal{D}$-horizontal curve $\gamma : [a, b] \to M$, we define for any $l \in [1, \infty]$ its *sub-Riemannian l-length* as

$$\text{length}_l(\gamma) = \Big\| ||\dot\gamma||_G \Big\|_{L^l[a,b]} = \begin{cases} \left(\int_a^b ||\dot\gamma(t)||_G^l dt\right)^{1/l} & \text{if } l < \infty \\ \text{esssup}_{t\in[a,b]} ||\dot\gamma(t)||_G & \text{if } l = \infty. \end{cases}$$

It is now natural to give the following definition.

DEFINITION 1.4. Given a sub-Riemannian manifold $(M, \mathcal{D}, G)$, the *sub-Riemannian l-distance* on M is defined for every x_0, x_1 in M as the nonnegative real number, or possibly $+\infty$, given by

$$\begin{aligned} d_l(x_0, x_1) = \inf\{ \text{length}_l(\gamma) :\ & \gamma : [a, b] \to M \text{ is } \mathcal{D}\text{-horizontal} \\ & \text{and such that } \gamma(a) = x_0, \gamma(b) = x_1\}. \end{aligned}$$

Similarly, one may define

$$\begin{aligned} d_{AL}(x_0, x_1) = \inf\{T :\ & \gamma : [0, T] \to M \text{ is } \mathcal{D}\text{-horizontal,} \\ & \text{parametrized by arc-length and} \\ & \text{such that } \gamma(a) = x_0, \gamma(b) = x_1\}. \end{aligned}$$

The following neat theorem holds.

THEOREM 1.5. *For any points x_0, x_1 in M and for any l in $[1, \infty]$ the equality $d_l(x_0, x_1) = d_{AL}(x_0, x_1)$ holds. When the infimum is actually a minimum, it is achieved by the same, possibly reparametrized, horizontal curve.*

PROOF. We follow [**1**]. Given a $\mathcal{D}$-horizontal curve $\gamma : [a, b] \to M$ with $\gamma(a) = x_0$ and $\gamma(b) = x_1$, Hölder inequality yields, for any l, $\text{length}_1(\gamma) \le \text{length}_l(\gamma) \le \text{length}_\infty(\gamma)$, so $d_1(x_0, x_1) \le d_l(x_0, x_1) \le d_\infty(x_0, x_1)$.

Next, if $\gamma : [0, T] \to M$ is parametrized by arc-length, define $\bar\gamma : [0, 1] \to M$ as $\bar\gamma(t) = \gamma(Tt)$ for any $t \in [0, 1]$. This curve is $\mathcal{D}$-horizontal and has the same extremal points of γ. Moreover

$$||\dot{\bar\gamma}||_G = T||\dot\gamma||_G = T$$

so $\text{length}_\infty(\bar\gamma) = T$. Since γ is arbitrary this yields $d_\infty(x_0, x_1) \le d_{AL}(x_0, x_1)$. Since any $\mathcal{D}$-horizontal curve can be parametrized by arc-length, the same procedure implies the converse inequality. Thus $d_\infty(x_0, x_1) = d_{AL}(x_0, x_1)$.

We just need to prove that $d_\infty(x_0, x_1) \le d_1(x_0, x_1)$ to conclude. To this aim, given an arbitraty $\mathcal{D}$-horizontal curve $\gamma : [a, b] \to M$ with $\gamma(a) = x_0$ and $\gamma(b) = x_1$, we want to construct another $\mathcal{D}$-horizontal curve $\bar\gamma$ defined in the same interval and with the same extremal points such that $\text{length}_\infty(\bar\gamma) = \text{length}_1(\gamma)$. Passing to the infimum over all the admissible curves will lead the desired inequality. We suppose for simplicity that $[a, b] = [0, 1]$, the general case is similar. We also assume $\text{length}_1(\gamma) > 0$. Let $\phi : [0, 1] \to [0, 1]$ be the function defined by

$$\phi(t) = \frac{1}{\text{length}_1(\gamma)} \int_0^t ||\dot\gamma(t)||_G dt.$$

This function is surjective, hence it has a right inverse, namely the function ψ given by

$$\psi(s) = \inf\{t \in [0,1] : \phi(t) = s\}$$

which is monotone and hence differentiable almost everywhere. Moreover, as ϕ is by construction absolutely continuous it maps sets of measure zero into sets of measure zero, so, differentiating the identity $\phi(\psi(s)) = s$ that holds for every s in $[0,1]$, we conclude that $1 = \dot{\phi}(\psi(s))\dot{\psi}(s)$ for almost every s. We define $\bar{\gamma} : [0,1] \to M$ as $\bar{\gamma}(s) = \gamma(\psi(s))$. This curve is a riparametrization of γ so it is $\mathcal{D}$-horizontal and has tha same extremal points of the original curve. We compute its ∞-length (keep in mind that the expressions below are defined almost everywhere):

$$\begin{aligned}
\text{length}_\infty(\bar{\gamma}) &= \operatorname*{esssup}_{s\in[0,1]} ||\dot{\bar{\gamma}}(s)||_G = \operatorname*{esssup}_{s\in[0,1]} ||\dot{\psi}(s)\dot{\gamma}(\psi(s))||_G \\
&= \operatorname*{esssup}_{s\in[0,1]} \Big(|\dot{\psi}(s)|\, ||\dot{\gamma}(\psi(s))||_G \Big) \\
&= \operatorname*{esssup}_{s\in[0,1]} \left(\frac{1}{|\dot{\phi}(\psi(s))|} ||\dot{\gamma}(\psi(s))||_G \right) \\
&= \operatorname*{esssup}_{s\in[0,1]} \left(\frac{\text{length}_1(\gamma)}{||\dot{\gamma}(\psi(s))||_G} ||\dot{\gamma}(\psi(s))||_G \right) \\
&= \text{length}_1(\gamma).
\end{aligned}$$

This proves the theorem. □

We shall henceforth indifferently denote by $d(\cdot,\cdot)$ the various distances and call it *the sub-Riemannian distance on* $(M, \mathcal{D}, G)$. We recall (see, for example, Bellaïche's article in [**3**]) that, since $\mathcal{D}$ is bracket-generating, the Chow-Rashevskii Theorem guarantees that if x_0 and x_1 belong to the same connected component of M then the distance between them is always finite, that is, there is always at least a horizontal curve connecting x_0 and x_1.

Definition 1.6. Given two points x_0, x_1 that belong to the same connected component of M, we call *length-minimizing curve from* x_0 *to* x_1 a Lipschitz curve $\gamma : [a,b] \to M$ such that $\gamma(a) = x_0$, $\gamma(b) = x_1$ and $d(x_0, x_1) = \text{length}(\gamma)$.

Viceversa, a Lipschitz curve $\gamma : [a,b] \to M$ is said to be *locally length-minimizing* if for any $c \in [a,b]$ there exists a sufficiently small $\epsilon > 0$ such that γ restricted to $[c-\epsilon, c+\epsilon]$ (or $[a, a+\epsilon]$ if $c = a$ or $[b-\epsilon, b]$ if $c = b$) is length-minimizing.

If in addition γ is the unique curve that minimizes the distance between the points, it will be called *uniquely (locally) length-minimizing.*

Locally length-minimizing curves are often referred to as "geodesics", especially in the classical Riemannian case. The Hopf-Rinow Theorem (see again Bellaïche in [**3**]) ensures that if (M, d) is a complete metric space (and this is the case when $M = \mathbb{R}^n$), then there always exists a length-minimizing curve connecting x_0 and x_1. We are interested in studying locally length-minimizing curves from a control theoretic point of view, regarding them as the minimizers of the sub-Riemannian length functional (see the next section). For a different approach involving the symplectic formalism, see [**4**].

1.1. Length-minimizing Curves and Extremals

We now make the following assumptions:

Assumptions 1.7. We suppose that

(1) $M = \mathbb{R}^n$ and $\mathcal{D}$ is spanned by the vector fields $X_1, \dots, X_m$

(2) The metric G has been chosen so that these vector fields are orthonormal.

In this way, if $\dot{\gamma}$ is in the form (1.1), $||\dot{\gamma}(t))||_G$ is simply $|h|$ and the l-length of γ is $\text{length}_l(\gamma) = ||h||_{L^l[a,b]}$.

Remark 1.8. The first assumption is not restrictive since we are interested in locally length-minimizing curves: if two points x_0, x_1 are sufficiently close they are contained in a domain of a single chart, and vice-versa any curve $\gamma : [a,b] \to M$ can be restricted to a sufficiently small compact subinterval of $[a,b]$ such that its image is in a single chart.
The second assumption is classical. However note that, given a distribution $\mathcal{D}$ in $\mathbb{R}^n$ endowed with a metric G, it is always possible to find a orthonormal frame of vector fields using the Gram-Schmidt process. It should be pointed though that choosing a priori which vector fields are orthonormal may impose some restrictions on G, resulting in a (small) loss of generality. See also Remark 1.12.

Fix two points x_0, x_1 in $\mathbb{R}^n$. Finding $\mathcal{D}$-horizontal, length-minimizing curves from x_0 to x_1 corresponds to the following minimization problem:

$$\text{Minimize } I(\gamma, h) = \int_a^b |h(t)|^l dt$$

$$\text{with } \begin{cases} \dot{\gamma}(t) & = \sum_{j=1}^m h_j(y) X_j(\gamma(t)) \\ \gamma(a) & = x_0 \\ \gamma(b) & = x_1 \\ h(\cdot) & \in L^\infty([a,b], \mathbb{R}^m). \end{cases} \tag{1.2}$$

One may note that the functional does not depend directly on the curve γ and it is possible to express the constraints exclusively by means of integral conditions over h. It is however more convenient for us to explicitly focus on the couple (γ, h), since we are in the hypothesis to apply the Pontryagin Maximum Principle (Theorem A.3 stated in Appendix A), that we restate here adapted to our setting.

Theorem 1.9 (The PMP in the sub-Riemannian case). *If (γ, h) is an optimal pair (that is, γ minimizes the distance bewteen x_0 and x_1 and has control h) there is a Lipschitz dual curve $\lambda : [a,b] \to T^*\mathbb{R}^n$ associated with γ and a constant $\lambda_0 \geq 0$ satisfying:*

adjoint equation: *for a.e. $t \in [a,b]$*

$$\dot{\lambda}(t) = -\sum_{j=1}^m h_j(t) \left\langle dX_j(\gamma(t))^T, \lambda(t) \right\rangle \tag{1.3}$$

which in components reads

$$\dot{\lambda}_k(t) = -\sum_{j=1}^m \sum_{i=1}^n h_j(t) \lambda_i(t) \frac{\partial X_{ji}}{\partial x_k}(\gamma(t)), \tag{1.4}$$

minimization: *given the Hamiltonian*

$$\mathcal{H}'(t, \lambda_0, \lambda, x, h) = \sum_{j=1}^m \langle \lambda, h_j X_j(x) \rangle + \lambda_0 |h|$$

for a.e. $t \in [a,b]$ there holds

$$\mathcal{H}'(t, \lambda_0, \lambda(t), \gamma(t), h(t)) = \min_{h \in \mathcal{U}} \mathcal{H}'(t, \lambda_0, \lambda(t), \gamma(t), h).$$

For $l = 2$, taking derivatives with respect to h and evaluating in h_* leads, after a renormalization of λ_0 so that we may suppose $\lambda_0 \in \{0, 1\}$, to

$$\langle \lambda(t), X_j(\gamma(t)) \rangle + \lambda_0 h_j(t) = 0 \quad \textit{for } j = 1 \dots m \tag{1.5}$$

for a.e. $t \in [a, b]$.

nontriviality: *for any $t \in [a, b]$*

$$(\lambda_0, \lambda(t)) \neq (0, 0). \tag{1.6}$$

The transversality condition (A.10), being $C = 0$, is trivial. In Chapter 3 we will actually prove this version of the PMP, and we will explicitly construct the dual curve λ.

DEFINITION 1.10. A (non necessarily length-minimizing) horizontal curve $\gamma : [a, b] \to M$ with control h is called *extremal* if there exists a dual curve associated with γ, $\lambda : [a, b] \to T^*\mathbb{R}^n$, satisfying the conditions of the Maximum Principle (1.3)-(1.6) stated above.
If the constant λ_0 can be choosen nonzero (and so $\lambda_0 = 1$ by the renormalization), the extremal will be called *normal*, in the opposite case *abnormal*. A *strictly abnormal extremal* is an extremal such that no associated dual curve can make it a normal extremal.

REMARK 1.11. In the classic Riemannian case, abnormal extremals do not occur. In fact, since $X_1(x), \dots, X_n(x)$ form a basis of $T_x\mathbb{R}^n$ for every $x \in \mathbb{R}^n$, condition (1.5) with $\lambda_0 = 0$ would imply $\lambda \equiv 0$, which is in contrast with the nontriviality hypotesis.

REMARK 1.12. A close inspection of Theorem A.3 reveals that when γ is an abnormal extremal, the choice of the metric G on $\mathcal{D}$ (that is, of the Lagrangian $L(h) = |h|^l$) is not relevant. Hence, *abnormal extremals are independent of the metric on $\mathcal{D}$*. They are somehow intrinsic to $\mathcal{D}$. Moreover, as we will see in a moment, normal extremals are "easy" and well understood objects, one is really interested only in studying abnormal extremals.

Given a curve $\gamma : [a, b] \to \mathbb{R}^n$, we say it is smooth, or C^k, or Lipschitz and so on if there exists a parametrization of γ such that it is smooth, or C^k, or Lipschitz and so on as a real function. In other words, we are interested in the regularity of the support of the curve regardless of its parametrization. For example, the curve defined $t \in [-1, 1]$ as $\gamma(t) = (t, |t|)$ is Lipschitz but not C^1 since it has a corner-type singularity, and no reparametrization can make it a C^1 function, while the curve defined for $t \in [0, 1]$ as $\gamma(t) = (\sqrt{t}, \sqrt{t})$ is smooth, since it is just the (badly) parametrized segment $t \mapsto (t, t)$. We are of course interested in the highest possible regular parametrizations from the point of view of differentiability; constant-speed parametrizations such as the arc-length parametrization provide such regularity.

1.1.1. Normal Extremals. The theorems presented in this subsection will tell us that, concerning regularity and minimality, normal extremals are completely characterized. The key point is the fact that they are (non-null) characteristic curves of a "classical" smooth Hamiltonian function. In fact, let us consider a normal extremal $\gamma : [a, b] \to \mathbb{R}^n$ with control h and associated dual curve λ. We define the following Hamiltonian:

$$H(x, \lambda) = -\frac{1}{2} \sum_{j=1}^{m} \langle \lambda, X_j(x) \rangle^2 \tag{1.7}$$

whose characteristic curves satisfy

$$\begin{cases} \dot{x} = \frac{\partial H}{\partial \lambda} = -\sum_{j=1}^{m} \langle \lambda, X_j(x) \rangle X_j(x) \\ \dot{\lambda} = -\frac{\partial H}{\partial x} = \sum_{j=1}^{m} \langle \lambda, X_j(x) \rangle \left\langle dX_j(x)^T, \lambda \right\rangle . \end{cases} \tag{1.8}$$

Now, being $\lambda_0 = 1$ in (1.5), $\langle \lambda(t), X_j(\gamma(t) \rangle$ is equal to $-h_j(t)$ and so it is immediate to note that the couple (γ, λ) solves a.e. the Hamilton-Jacobi equations for H.

THEOREM 1.13. *Any normal extremal is smooth.*

PROOF. Using the notation above, H is a smooth function, and so are its partial derivatives. From this we deduce that $\dot{\gamma}$ and $\dot{\lambda}$ are Lipschitz-continuous functions. But then γ and λ are C^1 functions. Taking the derivatives of (1.8) again, for the same reason we find that γ and λ are C^2. The procedure continues. □

THEOREM 1.14. *Any normal extremal is locally uniquely length-minimizing.*

PROOF. We follow the outline of [**4**]. We want to prove that, for any fixed $t_0 \in [a, b]$, there exists a sufficiently small $\epsilon > 0$ such that γ is the unique length-minimizing curve from $\gamma(t_0 - \epsilon)$ to $\gamma(t_0 + \epsilon)$, or from $\gamma(t_0)$ to $\gamma(t_0 + \epsilon)$ if $t_0 = a$, or from $\gamma(t_0 - \epsilon)$ to $\gamma(t_0)$ if $t_0 = b$. We deal with the case $t_0 \in (a, b)$, the remaining two are similar.

We suppose that γ is parametrized by arc-length so that $|h| = 1$ (if not, we just reparametrize it). Since (γ, λ) is a characteristic curve of H, $H(\gamma(t), \lambda(t))$ is constant. But, by condition (1.5) $h_j(t) = -\langle \lambda(t), X_j(\gamma(t)) \rangle$ for $j = 1, \ldots, m$ and so

$$H(\gamma(t), \lambda(t)) \equiv -1/2$$

for all t in $[a, b]$.

Let $x_0 = \gamma(t_0)$ and $\lambda(t_0) = \lambda_0$. Now comes a central claim:

Claim. There exists an open neighborhood V of x_0 in $\mathbb{R}^n$, a 1-form $\lambda = \lambda(x)$ defined on V and a small $\epsilon > 0$ such that

(1) $\lambda(\gamma(t))$ is the dual curve algong γ, $\lambda(t)$, for any t in $[t_0 - \epsilon, t_0 + \epsilon]$
(2) λ is exact, that is, there is a real valued smooth function f defined on V such that $df = \lambda$
(3) for any x in V

$$\sum_{j=1}^{m} \langle \lambda(x), X_j(x) \rangle^2 = 1.$$

In other words, the dual curve λ along γ extends in a convenient way near x_0 (whence the same name). This fact will be sufficient to prove both minimality and uniqueness.

Proof of the claim. Let S be a smooth bounded hypersuperface in $\mathbb{R}^n$ such that $x_0 \in S$ and S is orthogonal to λ_0, which means that $\operatorname{Ker} \lambda_0 = T_{x_0} S$. Using a suitable renormalization of the covector field orthogonal to S, we take a nonzero form defined on S, $\xi : S \to T^*\mathbb{R}^n$ such that it is orthogonal to all points of S ($\operatorname{Ker} \xi(p) = T_p S$ for all p in S), $\xi(x_0) = \lambda_0$ and $H(p, \xi(p)) = -1/2$ for all p in S.

Now, for a fixed $p \in S$, consider the characteristic curve (γ_p, λ_p) of H through $(p, \xi(p))$ at t_0, that is, the solution to the Cauchy problem

$$\begin{cases} \dot{\gamma}_p = \frac{\partial H}{\partial \lambda}(\gamma_p, \lambda_p) \\ \dot{\lambda}_p = -\frac{\partial H}{\partial x}(\gamma_p, \lambda_p) \\ \gamma_p(t_0) = p \\ \lambda_p(t_0) = \xi(p) \end{cases}$$

Note in particular that $(\gamma_{x_0}, \lambda_{x_0})$ is the normal extremal (γ, λ). Since $\bar{S}$ is compact, there exists a sufficiently small $\delta > 0$ such that, for any p, any (γ_p, λ_p) is defined in $(t_0-\delta, t_0+\delta)$. We then define the following flow mapping $\Phi : (t_0-\delta, t_0+\delta)\times S \to \mathbb{R}^n$

$$\Phi(t, p) = \gamma_p(t).$$

Now, fix a basis $v_1, \ldots, v_{n-1}$ of $T_{x_0}S$. Being $\Phi(t_0, p) = p$ for all p in S, it is

$$\frac{\partial \Phi}{\partial v_i}(t_0, x_0) = v_i \quad i = 1, \ldots, n-1.$$

Moreover

$$\frac{\partial \Phi}{\partial t}(t_0, x_0) = \frac{d}{dt}\gamma_{x_0}(t)\bigg|_{t=t_0} = \lambda_0$$

and λ_0 is orthogonal to $T_{x_0}S$.

It follows that the vectors $\frac{\partial\Phi}{\partial t}(t_0, x_0)$, $\frac{\partial\Phi}{\partial v_1}(t_0, x_0), \ldots, \frac{\partial\Phi}{\partial v_{n-1}}(t_0, x_0)$ are linearly independent and so $\det d\Phi(t_0, x_0)$ is not zero, which means that Φ is a local diffeomorphism at (t_0, x_0). Possibly shrinking δ and S, we deduce that Φ is a diffeomorphism from $U = (t_0-\delta, t_0+\delta)\times S$ onto its image $V = \Phi(U)$ which is an open neighborhood of x_0. The inverse function $\Phi^{-1} : V \to U$ must then be of the form $\Phi^{-1} = (f, F)$, where f is a smooth real-valued function such that, for all t in $(t_0 - \delta, t_0 + \delta)$ and p in S

$$f(\gamma_p(t)) = t. \tag{1.9}$$

Now, any x in V is of the form $x = \gamma_p(t)$ for a unique couple (t, p), so we define the 1-form λ of the claim on V by setting $\lambda(\gamma_p(t)) = \lambda_p(t)$.With this definition, condition (1) of the claim is satisfied.

Condition (2) is then equivalent to

$$\lambda_p(t) = df(\gamma_p(t)) \tag{1.10}$$

for all (t, p) in $(t_0 - \delta, t_0 + \delta) \times S$. We want to prove this. In the process of doing so also condition (3) will be proved.

Define the vector field X on V by setting

$$X(\gamma_p(t)) = \dot{\gamma}_p(t) = \frac{\partial H}{\partial \lambda}(\gamma_p, \lambda_p) = \sum_{j=1}^{m} \langle \lambda_p(t), X_j(\gamma_p(t))\rangle\, X_j(\gamma_p(t)). \tag{1.11}$$

The functions defined on V (that extend the control of γ on V) by

$$h_j(\gamma_p(t)) = \langle \lambda_p(t), X_j(\gamma_p(t))\rangle$$

do satisfy

$$\sum_{j=1}^{m} h_j(\gamma_p(t))^2 = \sum_{j=1}^{m} \langle \lambda_p, X_j(\gamma_p)\rangle^2 = 1 \tag{1.12}$$

for any $(t, p) \in (t_0 - \delta, t_0 + \delta) \times S$, since H is constant along its Hamiltonian flow and the equality is true when $p = x_0$. This in particular proves condition (3).

We set $S_t = \{x \in V : f(x) = t\}$. Note in particular that $S_{t_0} = S$ and that, Φ being a flow, for any t in $(t_0 - \delta, t_0 + \delta)$ $\Phi(t, \cdot) = \Phi_t$ is a diffeomorphism from S onto S_t. Fix an arbitrary v_0 in T_pS, and define $v(t) = d\Phi_t(p)[v_0] \in T_{\Phi_t(p)}S_t$. The function ϕ defined for $t \in (t_0 - \delta, t_0 + \delta)$ by

$$\phi(t) = \langle \lambda_p(t), v(t) \rangle \tag{1.13}$$

satisfies $\phi(t_0) = \langle \lambda_0(p), v_0 \rangle = \langle \xi(p), v_0 \rangle = 0$. Its derivative is

$$\dot{\phi} = \langle \dot{\lambda}_p, v \rangle + \langle \lambda_p, \dot{v} \rangle . \tag{1.14}$$

We focus on the first term. Differentiating with respect to x_i for $i = 1, \ldots, n$ (1.12) we deduce that, on V

$$\sum_{j=1}^{m} h_j \frac{\partial h_j}{\partial x_i} = 0. \tag{1.15}$$

We define the Hamiltonian K on V (that is, on T^*V) by $K(x, \lambda) = \langle \lambda, X(x) \rangle$ (recall that X is defined in 1.11) and we note that characteristic curves of H with range in V are also characteristic curves of K, that is the satisfy the ODE

$$\begin{cases} \dot{\gamma} = \frac{\partial K}{\partial \lambda}(\gamma, \lambda) \\ \dot{\lambda} = -\frac{\partial K}{\partial x}(\gamma, \lambda). \end{cases}$$

In fact, the first equation follows immediately from the definition of the vector field X; as for the second one, we differentiate and use (1.15):

$$\begin{aligned} \frac{\partial K}{\partial x_i} &= \sum_{j=1}^{m} \left(\frac{\partial h_j}{\partial x_i} h_j + h_j \left\langle \lambda, \frac{\partial X_j}{\partial x_i} \right\rangle \right) \\ &= \sum_{j=1}^{m} h_j \frac{\partial X_j}{\partial x_i} = \frac{\partial H}{\partial x_i}. \end{aligned}$$

We conclude that $\dot{\lambda}_p = -\frac{\partial K}{\partial x}(\gamma_p, \lambda_p) = -\langle \lambda_p, D_x X(\gamma_p) \rangle$ and so the first term of equation (1.14) becomes

$$\langle \dot{\lambda}_p, v \rangle = -\langle \lambda_p, d_x X(\gamma_p)[v] \rangle . \tag{1.16}$$

We study on the second term of (1.14). We compute $\dot{v}$:

$$\begin{aligned} \dot{v} &= \frac{d}{dt} d_p \Phi_t(p)[v_0] = d_p \left(\frac{d}{dt} \Phi_t(p)[v_0] \right) = d_p(X(\Phi_t(p))[v_0] \\ &= d_x X(\Phi_t(p))[d_p \Phi_t(p)[v_0]] = d_x X(\Phi_t(p))[v]. \end{aligned}$$

Thus, $\langle \lambda_p, \dot{v} \rangle = \langle \lambda_p, d_x X(\Phi_t(p))[v] \rangle$. Together with (1.16) it leads $\dot{\phi} = 0$ and since $\phi(t_0) = 0$, ϕ identically vanishes. By looking at the definition of ϕ in (1.13), λ_p must be orthogonal to $v(t)$ in $T_{\Phi_t(p)} S_t$ for any choiche of the tangent vector $v_0 \in T_p S$. This means that λ_p is orthogonal to $T_{\gamma_p(t)} S_t$, and since df is orthogonal to the same space, we conclude that there exists a real function β on V such that

$$df(\gamma_p) = \beta(\gamma_p) \lambda_p.$$

To prove that β is identically 1, differentiate with respect to t equation (1.9) and use relation (1.12) to find

$$\begin{aligned} 1 &= df(\gamma_p)[\dot{\gamma_p}] = \langle df(\gamma_p), X(\gamma_p) \rangle \\ &= \left\langle \beta(\gamma_p) \lambda_p, \sum_{j=1}^{m} \langle \lambda_p, X_j(\gamma_p) \rangle X_j(\gamma_p) \right\rangle = \beta(\gamma_p). \end{aligned}$$

This concludes the proof of condition (3) and of the claim.

Proof of the local minimality. Fix any $\epsilon \in (0, \delta)$, consider tha restriction of γ to $[t_0 - \epsilon, t_0 + \epsilon]$ and let $\bar{\gamma} : [t_0 - \epsilon, t_0 + \epsilon] \to \mathbb{R}^n$ be a competitor for γ, that is, a $\mathcal{D}$-horizontal curve such that $\bar{\gamma}(t_0) = x_0$, $\bar{\gamma}(t_0 - \epsilon) = \gamma(t_0 - \epsilon)$ and $\bar{\gamma}(t_0 + \epsilon) = \gamma(t_0 + \epsilon)$. We may suppose that the support of $\bar{\gamma}$ is in V. If $\bar{h} = (\bar{h}_1, \ldots, \bar{h}_m)$ is the control of such curve, its length is

$$\text{length}(\bar{\gamma}) = \int_{t_0-\epsilon}^{t_0+\epsilon} |\bar{h}(t)| dt.$$

By definition of f (see equation (1.9)), and since γ is parametrized by arc-length, we have that $\text{length}(\gamma) = 2\epsilon = f(\gamma(t_0 + \epsilon)) - f(\gamma(t_0 - \epsilon)) = f(\bar{\gamma}(t_0 + \epsilon)) - f(\bar{\gamma}(t_0 - \epsilon))$. Denoting by $X_j(\bar{\gamma}(t))(f)$ the action of X_j on f at the point $\bar{\gamma}(t)$ as a differential operator, by the Cauchy-Schwartz inequality it follows that

$$\begin{aligned} \text{length}(\gamma) &= \int_{t_0-\epsilon}^{t_0+\epsilon} \frac{d}{dt} f(\bar{\gamma}(t)) dt \\ &= \int_{t_0-\epsilon}^{t_0+\epsilon} df(\bar{\gamma}(t))[\dot{\bar{\gamma}}(t)] dt = \int_{t_0-\epsilon}^{t_0+\epsilon} \sum_{j=1}^{m} \bar{h}_j X_j(\bar{\gamma}(t))(f) dt \\ &\leq \int_{t_0-\epsilon}^{t_0+\epsilon} |\bar{h}| \left(\sum_{j=1}^{m} (X_j(\bar{\gamma}(t))(f))^2 \right)^{1/2} dt = \text{length}(\bar{\gamma}) \end{aligned} \tag{1.17}$$

Where the last equality is just the application of condition (2) of the claim. This proves the minimality of γ.

Proof of the uniqueness. We finally want to prove that, if $\text{length}(\gamma) = \text{length}(\bar{\gamma})$ then the two curves coincide over $[t_0 - \epsilon, t_0 + \epsilon]$. Define the 1-form along $\bar{\gamma}$ as $\bar{\lambda}(t) = \lambda(\bar{\gamma}(t))$. Since both (γ, λ) and $(\bar{\gamma}, \bar{\lambda})$ take the same value on $t_0 - \epsilon$, it is sufficient to prove that $(\bar{\gamma}, \bar{\lambda})$ is a characteristic curve of the Hamiltonian H defined in (1.7).

Starting from (1.17), since we assume $\text{length}(\gamma) = \text{length}(\bar{\gamma})$ the Cauchy-Schwartz inequality is actually an equality, so there must be a real-valued function $s = s(t)$ such that

$$\bar{h}_j = s X_j(\bar{\gamma})(f) = s \langle \lambda(\bar{\gamma}), X_j(\bar{\gamma}) \rangle \quad j = 1, \ldots, m$$

where the last equality follows from (1.10). Since $\text{length}(\bar{\gamma}) = \text{length}(\gamma) = 2\epsilon$, it is possible to assume that $\bar{\gamma}$ has unitary speed, i.e. $|\bar{h}|$=1, and so the above equation together with condition (3) of the claim implies $s = 1$. This leads to $\dot{\bar{\gamma}} = \frac{\partial H}{\partial \lambda}(\bar{\gamma}, \bar{\lambda})$.

As for the dual curve, we compute:

$$
\begin{aligned}
\dot{\bar\lambda}_i &= \frac{d}{dt}\frac{\partial f}{\partial x_i}(\bar\gamma) = \sum_{k=1}^{n} \frac{\partial^2 f}{x_k x_i}(\bar\gamma)\dot{\bar\gamma}_k \\
&= \sum_{j=1}^{m}\sum_{k=1}^{n} \frac{\partial^2 f}{x_i x_k}\left\langle \bar\lambda, X_j \right\rangle X_{jk}(\bar\gamma) \\
&= \sum_{j=1}^{m}\sum_{k=1}^{n} \frac{\partial \bar\lambda_k}{\partial x_i}\left\langle \bar\lambda, X_j \right\rangle X_{jk}(\bar\gamma) \\
&= \sum_{j=1}^{m} \left\langle \bar\lambda, X_j \right\rangle \left\langle \frac{\partial \bar\lambda}{\partial x_i}, X_j \right\rangle (\bar\gamma).
\end{aligned} \tag{1.18}
$$

If we differentiate with respect to x_i condition (3) of the claim we obtain

$$
\sum_{j=1}^{m} \left\langle \bar\lambda, X_j \right\rangle \left(\left\langle \frac{\partial \bar\lambda}{\partial x_i}, X_j \right\rangle + \left\langle \bar\lambda, \frac{\partial X_j}{\partial x_i} \right\rangle \right) = 0
$$

which, evaluated in $\bar\gamma$ and combined with (1.18) lead eventually to

$$
\bar\lambda_i = -\sum_{j=1}^{m} \left\langle \bar\lambda, X_j(\bar\gamma) \right\rangle \left\langle \bar\lambda, \frac{\partial X_j}{\partial x_i}(\bar\gamma) \right\rangle = -\frac{\partial H}{\partial x_i}(\bar\gamma, \bar\lambda)
$$

which proves that $(\bar\gamma, \bar\lambda)$ is a characteristic curve of H, and hence completes the proof of the theorem. □

1.1.2. Abnormal Extremals. For abnormal extremals the situation is quite different. First of all, a general regularity result analogous to Theorem 1.13 cannot hold.

Example 1.15. Let $M = \mathbb{R}^5$ and $\mathcal{D}$ be the rank-two distribution spanned by

$$
\begin{aligned}
X_1(x) &= \frac{\partial}{\partial x_1} \\
X_2(x) &= \frac{\partial}{\partial x_2} + x_1 \frac{\partial}{\partial x_3} + x_1^5 \frac{\partial}{\partial x_4} + x_1 x_2^5 \frac{\partial}{\partial x_5}
\end{aligned}
$$

A straightforward computation proves that it is bracket generating of step 5. If $\gamma : I \to \mathbb{R}^n$ is an $\mathcal{D}$-horizontal curve with control $h = (h_1, h_2)$, i.e.

$$
\gamma(t) = h_1(t)X_1(\gamma(t)) + h_2(t)X_2(\gamma(t)) \quad \text{for a.e } t \in I
$$

clearly then $h_1 = \dot\gamma_1$ and $h_2 = \dot\gamma_2$. The remaining components γ_3, γ_4 and γ_5 are determined by γ_1 and γ_2. The curve $\gamma : [0,1] \to M$ defined by

$$
\begin{cases} \gamma_1(t) = t \\ \gamma_2(t) = t^{4/3} \end{cases}
$$

is an abnormal extremal with dual curve $\lambda(t) = (0, \frac{4}{5}t^5, 0, \frac{1}{5}, -1)$. Since it is nonsmooth, by Theorem 1.13 γ must be strictly abnormal. On [**7**], through lengthy and difficult calculations it is proved that the curve is not length-minimizing at the singular point $t = 0$.

Remark 1.16. Many examples of this kind can be obtained using the techniques presented in section 2.2 of the following chapter.

However, strictly abnormal extremals can be length-minimizing. In Section 2.4 of Chapter 2 we will discuss a detailed example of a strictly abnormal, length-minimizing smooth curve. The question on whether nonsmooth length-minimizing curves (thus necessarily abnormal extremals) exist is currently open: at the present time, no examples of this kind are known but, on the opposite side, authors still have dot developed a general regularity theory for abnormal length-minimizing curves.

CHAPTER 2

Rank-Two Distributions

In this chapter we deal with rank-two distributions. Our aim is to produce concrete examples of extremal curves and present some regularity results. Recall Assumptions 1.7 along with Remarks 1.8 and 1.12.

2.1. Regular Abnormal Extremals

We summarize here the main result obtained by Liu and Sussmann in [**4**]. Recall formula (0.1) stated in Notation and Conventions and let us consider the distribution $\mathcal{D}$ in $\mathbb{R}^n$ spanned by two vector fields X_1 and X_2 and a Lipschitz $\mathcal{D}$-horizontal curve $\gamma : [a,b] \to \mathbb{R}^n$ with control $h = (h_1, h_2)$ as in Definition 1.2. If γ is an abnormal extremal, then there exists an associated dual curve $\lambda : [a,b] \to T^*\mathbb{R}^n$ satisfying, for any t in $[a,b]$

$$\langle \lambda(t), X_1(\gamma(t)) \rangle = 0 \qquad j = 1, 2.$$

Let us differentiate the expression corresponding to $j = 1$ with respect to t:

$$\begin{aligned} 0 &= \left\langle \dot\lambda(t), X_1(\gamma(t)) \right\rangle + \langle \lambda(t), dX_1(\gamma(t))[\dot\gamma(t)] \rangle \\ &= \left\langle \dot\lambda(t), X_1(\gamma(t)) \right\rangle + \langle \lambda(t), dX_1(\gamma(t))[h_1(t)X_1(\gamma(t)) + h_2(t)X_2(\gamma(t))] \rangle \\ &= \left\langle \dot\lambda(t), X_1(\gamma(t)) \right\rangle + \\ &\quad h_1(t) \langle \lambda(t), dX_1(\gamma(t))[X_1(\gamma(t))] \rangle + h_2(t) \langle \lambda(t), dX_1(\gamma(t))[X_2(\gamma(t))] \rangle . \end{aligned} \tag{2.1}$$

On the other hand, from the adjoint equation (1.3), we deduce

$$\begin{aligned} \left\langle \dot\lambda(t), X_1(\gamma(t)) \right\rangle = -\Big(& h_1(t) \langle \lambda(t), dX_1[X_1(\gamma(t))] \rangle + \\ & h_2(t) \langle \lambda(t), dX_2(\gamma(t))[X_1(\gamma(t))] \rangle \Big). \end{aligned} \tag{2.2}$$

Inserting expression (2.2) in equation (2.1), we get

$$\begin{aligned} 0 &= -h_2 \langle \lambda(t), dX_2(\gamma(t))[X_1(\gamma(t))] \rangle + h_2(t) \langle \lambda(t), dX_1(\gamma(t))[X_2(\gamma(t))] \rangle \\ &= h_2(t) \langle \lambda(t), -dX_2(\gamma(t))[X_1(\gamma(t))] + dX_1(\gamma(t))[X_2(\gamma(t))] \rangle \\ &= -h_2(t) \langle \lambda(t), [X_1, X_2](\gamma(t)) \rangle . \end{aligned}$$

Exactly in the same way, if we differentiate $\langle \lambda(t), X_2(\gamma(t)) \rangle = 0$ we deduce that $h_1(t) \langle \lambda(t), [X_1, X_2](\gamma(t)) \rangle = 0$. Since the control h can be choosen to be always nonzero (for example, parametrizing γ by arc-length) it is necessary that

$$\langle \lambda(t), [X_1, X_2](\gamma(t)) \rangle = 0 \tag{2.3}$$

for any t in $[a,b]$. This condition is called *Goh condition* and can be extended to distributions of any rank, see Chapter 3. We can rephrase the above result as follows: given an abnormal extremal $\gamma : [a,b] \to \mathbb{R}^n$ and an associated dual curve $\lambda : [a,b] \to T^*\mathbb{R}^n$, we *always* have

- $\lambda(t)$ is orthogonal to $\mathcal{D}^1(\gamma(t)) = \mathcal{D}(\gamma(t))$ for any t by definition of abnormal extremal
- $\lambda(t)$ is orthogonal to $\mathcal{D}^2(\gamma(t))$ for any t by the Goh condition (2.3).

We hence give the following definition.

DEFINITION 2.1. Let $\gamma : [a,b] \to \mathbb{R}^n$ be an abnormal extremal curve with control h and dual curve $\lambda : [a,b] \to T^*\mathbb{R}^n$. We say that γ is a *regular abnormal extremal* if it is at least C^1 and $\lambda(t)$ is not orthogonal to $\mathcal{D}^3(\gamma(t))$ for any t in $[a,b]$. This means that we must have

$$\begin{aligned} &\langle \lambda(t), [X_1,[X_1,X_2]](\gamma(t))\rangle \neq 0, \quad \text{or} \\ &\langle \lambda(t), [X_2,[X_1,X_2]](\gamma(t))\rangle \neq 0 \end{aligned}$$

for any t in $[a,b]$.

THEOREM 2.2 (Liu and Sussmann). *Any regular abnormal extremal is smooth and locally uniquely length-minimizing.*

Since its hypotesis are particularly easy to verify, Theorem 2.2 is very useful. We will make use of it in the following sections.

2.2. An "aut-aut" Theorem

In this Section we prove a Theorem stated in [**7**] (Theorem 9 in Section 5) that characterizes extremals of a rank-two distributions $\mathcal{D}$ generated by the vector fields X_1 and X_2 given, for any $x \in \mathbb{R}^n$, by

$$\begin{aligned} X_1(x) &= \frac{\partial}{\partial x_1} \\ X_2(x) &= \frac{\partial}{\partial x_2} + \sum_{i=3}^n f_i(x)\frac{\partial}{\partial x_i} \end{aligned} \tag{2.4}$$

where the functions $f_i(x) = f_i(x_1,x_2)$ for $i = 3,\dots,n$ are smooth functions depending only on the first two variables. These distributions arise as a result of a limiting process (or "blow-up") that is employed, in proving non-minimality of extremal curves with corners, see [**9**].

If $\gamma : I \to \mathbb{R}^n$ is a horizontal curve parametrized by arc-length, then there exist measurable functions u, v such that $u^2 + v^2 = 1$ and

$$\gamma(t) = u(t)X_1(\gamma(t)) + v(t)X_2(\gamma(t)) \quad \text{for a.e } t \in I.$$

Clearly then $u = \dot\gamma_1$ and $v = \dot\gamma_2$. From now on when omit the explicitation of the t dependence.
If we want the curve to be an extremal, Theorem 1.9 states that we must find a dual curve $\lambda : I \to T^*\mathbb{R}^n$ such that $\lambda \neq 0$ and

$$\begin{cases} \langle \lambda, X_1(\gamma)\rangle = -\lambda_0 u \\ \langle \lambda, X_2(\gamma)\rangle = -\lambda_0 v \end{cases} \quad \text{and } \dot\lambda_k = -\sum_{i=3}^n \lambda_i \frac{\partial X_{1i}}{\partial x_k}(\gamma)u + \lambda_i\frac{\partial X_{2i}}{\partial x_k}(\gamma)v$$

for $k = 1,\dots,n$ and $\lambda_0 \in \{0,1\}$. In our specific case, we must than have

$$\begin{cases} \lambda_1 = -\lambda_0 u \\ \lambda_2 + \sum_{i=3}^n \lambda_i f_i(\gamma) = -\lambda_0 v \end{cases} \tag{2.5}$$

and

$$\begin{cases} \dot\lambda_1 = -v\left(\sum_{i=3}^n \lambda_i \frac{\partial f_i}{\partial x_1}(\gamma)\right) \\ \dot\lambda_2 = -v\left(\sum_{i=3}^n \lambda_i \frac{\partial f_i}{\partial x_2}(\gamma)\right) \\ \dot\lambda_3 = 0 \\ \vdots \\ \dot\lambda_n = 0 \end{cases} \tag{2.6}$$

These equations clearly imply that $\lambda_3, \dots \lambda_n$ are constant numbers and that λ_1 and λ_2 are entirely determined by these constants and the controls.

Abnomal extremals. We suppose $\lambda_0 = 0$. Differentiating the second equation in (2.5) with respect to t we obtain

$$\begin{aligned} \dot\lambda_2 &= -\sum_{i=3}^n \lambda_i \left(\frac{\partial f_i}{\partial x_1}(\gamma)\dot\gamma_1 + \frac{\partial f_i}{\partial x_2}(\gamma)\dot\gamma_2\right) \\ &= -\sum_{i=3}^n \lambda_i \left(\frac{\partial f_i}{\partial x_1}(\gamma)u + \frac{\partial f_i}{\partial x_2}(\gamma)v\right). \end{aligned}$$

comparing this equation with (2.6), it is necessary that

$$-u\sum_{i=3}^n \lambda_i \frac{\partial f_i}{\partial x_1}(\gamma) = 0.$$

Now, from (2.5) we also get that $\lambda_1 = 0$, so

$$0 = \dot\lambda_1 = -v\sum_{i=3}^n \lambda_i \frac{\partial f_i}{\partial x_1}(\gamma).$$

Since $(u,v) \neq (0,0)$ we conclude that

$$\sum_{i=3}^n \lambda_i \frac{\partial f_i}{\partial x_1}(\gamma) = 0 \quad \text{for a.e. } t \in I. \tag{2.7}$$

This is a necessary condition for γ to be an abnormal extremal for $\mathcal{D}$. Viceversa, choosing constants $\lambda_3, \dots \lambda_n \in \mathbb{R}$ such that (2.7) holds allows us to define a dual curve λ by choosing $\lambda_1 = 0$ and λ_2 defined by (2.5).

Normal extremals. We now suppose $\lambda_0 = 1$, that is γ is a normal extremal. We know from Theorem 1.13 that both γ and λ are smooth. Comparing the equations for λ_1 in (2.5) and (2.6) we get

$$\dot u = v\sum_{i=3}^n \lambda_i \frac{\partial f_i}{\partial x_1}(\gamma) \tag{2.8}$$

while for λ_2 we get

$$v\sum_{i=3}^n \lambda_i \frac{\partial f_i}{\partial x_1}(\gamma) = \dot v + \sum_{i=3}^n \lambda_i \left(\frac{\partial f_i}{\partial x_1}(\gamma)u + \frac{\partial f_i}{\partial x_2}(\gamma)v\right)$$

that is

$$\dot{v} = -u \sum_{i=3}^{n} \lambda_i \frac{\partial f_i}{\partial x_1}(\gamma). \tag{2.9}$$

Viceversa, if equations (2.8), (2.9) hold, we deduce that λ_0 cannot be zero, hence the extremal must be normal. We have proved:

THEOREM 2.3. *For a rank-two distribution of the form* (2.4)*, a $\mathcal{D}$-horizontal curve $\gamma : I \to \mathbb{R}^n$ with controls $(u, v) = (\dot{\gamma}_1, \dot{\gamma}_2)$, the following facts hold:*

(1) *γ is a normal extremal if and only if there exist constants $\lambda_3, \ldots, \lambda_n \in \mathbb{R}$ such that*
$$\begin{cases} \dot{u} = v \sum_{i=3}^{n} \lambda_i \frac{\partial f}{\partial x_1}(\gamma) \\ \dot{v} = -u \sum_{i=3}^{n} \lambda_i \frac{\partial f}{\partial x_1}(\gamma) \end{cases}$$
in particular this means that $(\ddot{\gamma}_1, \ddot{\gamma}_2) \perp (\dot{\gamma}_1, \dot{\gamma}_2)$;

(2) *γ is an abnormal extremal if and only if there exist constants $\lambda_3, \ldots, \lambda_n \in \mathbb{R}$ such that*
$$\sum_{i=3}^{n} \lambda_i \frac{\partial f_i}{\partial x_1}(\gamma) = 0 \quad \textit{for a.e } t \in I.$$

2.3. An extremal of regularity $C^{1,1/m}$

We shall now study a specific case of (2.4) that was suggested in [**7**], namely, the distribution in $\mathbb{R}^5$ given by

$$X_1(x) = \frac{\partial}{\partial x_1}$$
$$X_2(x) = \frac{\partial}{\partial x_2} + x_1 \frac{\partial}{\partial x_3} + x_1^{2m} \frac{\partial}{\partial x_4} + x_1 x_2^m \frac{\partial}{\partial x_5}$$

where m is an integer greater or equal than 2. We are interested in studying its abnormal nonsmooth minimizers. As we have seen, all the components of the curve are determined by γ_1 and γ_2. Due to Theorem 2.3, all abnormal minimizers satisfy the algebraic condition

$$\lambda_3 + \lambda_4 \gamma_1^{2m-1} + \lambda_5 \gamma_2^m = 0$$

for some real constants λ_3, λ_4 and λ_5 that completely determine the associated dual curve λ. We distinguish cases:

If $\lambda_4 = 0$: we get $\lambda_3 + \lambda_5 \gamma_2^m = 0$, which implies that γ_2 is constant and so $\dot{\gamma}_2 = 0$ a.e. This means that γ is an integral curve of X_1 and so a smooth curve

If $\lambda_5 = 0$: we get similarly that γ is an integral curve of X_2

If λ_3, λ_4 and λ_5 are not zero: then for any t $(\gamma_1(t), \gamma_2(t)) \neq (0, 0)$. This is enough to conclude since the curve in the plane $t \mapsto (\gamma_1(t), \gamma_2(t))$ is implicitly defined by the analytic function $f(x, y) = \lambda_3 + \lambda_4 x^{2m-1} + \lambda_5 y^m$, whose differential has matrix

$$Df(x, y) = \begin{pmatrix} (2m-1)\lambda_4 x^{2m-2} \\ m\lambda_5 y^{m-1} \end{pmatrix}.$$

Since it is $Df(x, y) \neq 0$ for $(x, y) \neq (0, 0)$, the Implicit Function Theorem allows us to conclude that $(\gamma_1(t), \gamma_2(t))$ is as analytic, hence smooth.

The only remaining case is when $\lambda_3 = 0, \lambda_4 \neq 0 \neq \lambda_5$. Supposing for simplicity $-\lambda_4/\lambda_5 = 1$, we get $\gamma_1^{2m-1} = \gamma_2^m$ which yields to $\gamma_2(t) = \gamma_1^{2-1/m}(t)$. If we want the extremal to have a chance of being nonsmooth we must have that, for some $\bar{t}$, $\gamma_1(\bar{t}) = 0$. The possibly singular point of the curve is then the origin. Next, we note that

$$[X_1, [X_1, X_2]](x) = (2m-1)x_1^{2m-1}\frac{\partial}{\partial x_4}$$
$$[X_2, [X_1, X_2]](x) = m x_2^{m-1}\frac{\partial}{\partial x_5}$$

so

$$\langle \lambda(t), [X_1, [X_1, X_2]](\gamma(t))\rangle = \lambda_4(2m-1)\gamma_1(t)^{2m-1}$$
$$\langle \lambda(t), [X_2, [X_1, X_2]](\gamma(t))\rangle = \lambda_5 m \gamma_2(t)^{m-1}.$$

We can apply Theorem 2.2: away from 0, γ is a regular abnormal extremal, and so uniquely locally length minimizing.

A parametrization of the curve near 0 can be obtained in the following way. The domain of the curve is $I = [0, 1]$, $\gamma_1(t) = t$ and so $\gamma_2(t) = t^{1-1/m}$. This curve has regularity exactly $C^{1,1/m}$, and is uniquely locally length-minimizing far from 0. At the present time we do not know if this curve is also minimizing at the singuar point. For more on this extremal, see also Subsection 3.2.1 of Chapter 3.

REMARK 2.4. Elementary modifications of the techniques presented in this and in the previous sections lead to the construction of Example 1.15.

2.4. Strictly Abnormal, Non Regular Length-Minimizers

In this section we present a family of distributions that admit a strictly anbormal extremals that are not "regular" according to Definition 2.1 at any point and for which, consequently, Theorem 2.2 does not apply. We also prove that for some of these distributions they actually are length-minimizing. This is substantially a generalization of an example presented in Section 2.3 of [**4**].

Let $\mathcal{D} = \mathcal{D}(m)$ be the distribution in $\mathbb{R}^3$ defined by

$$X_1 = \frac{\partial}{\partial x_1}$$
$$X_2 = (1 - x_1)\frac{\partial}{\partial x_2} + x_1^m\frac{\partial}{\partial x_3} \tag{2.10}$$

where m is a positive integer. For $m = 1$, the distribution has step two, so Goh condition (2.3) implies that there are no abnormal extremals: the (immediate) proof will be given for general-rank distributions, see Corollary 3.10.

For $m \geq 2$, the ditribution has step $m + 1$ at the origin, that is, the "critical" point in $\mathbb{R}^3$ is 0, far from which the distribution has step 2. Consequently, abnormal extremals necessarily pass through the origin and we may look for $\mathcal{D}$-horizontal curves $\gamma : [0, 1] \to \mathbb{R}^3$ such that $\gamma(0) = 0$. If $h = (h_1, h_2)$ is the control, and $\lambda : [0, 1] \to \mathbb{R}^3$ an associated dual curve, Theorem 1.9 leads to the following equations:

$$\begin{cases} \dot\gamma_1(t) = h_1(t) \\ \dot\gamma_2(t) = h_2(t)(1-\gamma_1(t)) \\ \dot\gamma_3(t) = h_2(t)\gamma_1(t)^m \\ (\gamma_1(0), \gamma_2(0), \gamma_3(0)) = (0,0,0) \end{cases} \quad \text{and} \quad \begin{cases} \lambda_1(t) = 0 \\ \lambda_2(t)(1-\gamma_1(t)) + \lambda_3(t)\gamma_1^m(t) = 0 \\ \dot\lambda_1(t) = -h_2(t)(-\lambda_2(t) + m\lambda_3(t)\gamma_1(t)^{m-1}) \\ \dot\lambda_2(t) = 0 \\ \dot\lambda_3(t) = 0. \end{cases}$$

Clearly then $\lambda = (0, \lambda_2, \lambda_3)$ for some real constants λ_2, λ_3 non both zero (due the nontriviality condition). Evaluating in 0 the second equation of the group on the right we get

$$\lambda_2(0)(1-\gamma_1(0)) + \lambda_3(0)\gamma_1^m(0) = 0$$

and taking into account that $\gamma_1(0) = 0$, we conclude that $\lambda_2 = 0$. But then, the same equation leads to $\lambda_3\gamma_1(t)^m = 0$, and so being λ_3 nonzero, we conclude $\gamma_1 = 0$. Necessarily, then, the only possible abnormal extremal is the segment along the x_2 axis:

$$\gamma(t) = (0, t, 0)$$

and the control is then $h = (0, 1)$. Abnormality is strict: if γ were normal, again Theorem 1.9 would lead to the incompatible conditions

$$\begin{cases} \lambda_1(t) = 0 \\ \lambda_2(t) = -1 \\ \dot\lambda_1(t) = -\lambda_2(t) \end{cases}$$

so that no dual curve can be associated to γ to make it a normal extremal.

Since

$$[X_j, [X_1, X_2]](\gamma) \equiv 0 \quad \text{for } j = 1, 2 \text{ and } m > 2$$

the curve is not a regular abnormal extremal (according to Definition 2.1) when m is strictly greater than 2.

Now, we prove that *when m is even*, for $0 < b - a \le \bar\tau = \frac{2}{m+1}$ the curve

$$\gamma : [a, b] \to \mathbb{R}^3$$
$$t \mapsto (0, t, 0)$$

is uniquely locally length-minimizing. Let us suppose that $\delta : [0, \tau] \to \mathbb{R}^3$ is another $\mathcal{D}$-horizontal curve in $\mathbb{R}^3$ parametrized by arc-length such that $\delta(0) = (0, a, 0)$ and $\delta(\tau) = (0, b, 0)$. There exist unique measurable functions u, v with $u^2 + v^2 = 1$ a.e. such that

$$\dot\delta(t) = u(t)X_1(\delta(t)) + v(t)X_2(\delta(t))$$

for a.e. $t \in [0, \tau]$. In particular, if $\delta(t) = (x(t), y(t), z(t))$, it must be

$$\dot x(t) = u(t)$$
$$\dot y(t) = (1 - x(t))v(t)$$
$$\dot z(t) = x(t)^m v(t)$$

Since

$$\text{length}_G(\gamma) = b - a = y(\tau) - y(0) = \int_0^\tau (1 - x(s))v(s)ds$$

our claim is proved through the following lemma.

LEMMA 2.5. *Let $0 < \tau \le \bar{\tau} = \frac{2}{m+1}$, and let $u, v : [0,\tau] \to \mathbb{R}$ be measurable functions with $|u(t)|, |v(t)| \le 1$ for a.e. $t \in [0,\tau]$. Define $x : [0,\tau] \to \mathbb{R}$ by*

$$x(t) = \int_0^t u(s)ds.$$

If $x(\tau) = 0$ and $\int_0^\tau x(s)^m v(s)ds = 0$ then

$$\int_0^\tau (1 - x(s))v(s)ds \le \tau \tag{2.11}$$

and the equality holds if and only of $u = 0, v = 1$ a.e.

PROOF. Set

$$\begin{aligned} A &= \int_0^\tau (1 - x(s))v(s)ds, \\ V(t) &= \int_0^t v(s)ds, \\ \alpha &= \tau - V(\tau), \\ \beta &= ||x||_\infty. \end{aligned}$$

Since $|v| \le 1$,

$$\begin{aligned} A &\le \int_0^\tau v(s)ds + \left| \int_0^\tau x(s)v(s)ds \right| \le V(\tau) + \beta\tau \\ &\le V(\tau) - \tau + \tau + \beta\tau = \tau - \alpha + \beta\tau. \end{aligned}$$

It is then enough to prove that $-\alpha + \beta\tau \le 0$, and this is true if $\beta \le \frac{\alpha}{\tau}$. We prove this last inequality by showing that

$$\bar{\tau}\beta^{m+1} \le \int_0^\tau x(s)^m ds \le \beta^m \alpha.$$

Inequality on the right. Since $\int_0^\tau x(s)^m v(s)ds = 0$ by hypotesis,

$$\begin{aligned} \int_0^\tau x(s)^m ds &\le \int_0^\tau x(s)^m (1 - v(s))ds \\ &\le \beta^m \int_0^\tau (1 - v(s))ds = \beta^m(\tau - V(\tau)) = \beta^m \alpha. \end{aligned}$$

Inequality on the left. Let $a \in [0,\tau]$ such that $|x(a)| = \beta$. Since $x(0) = 0 = x(\tau)$ and $|\dot{x}(t)| = |u(t)| \le 1$ it is $a \ge \beta$ and $\tau - a \ge \beta$. So the intervals $I_1 = [a - \beta, a]$ and $I_2 = [a, a + \beta]$ are contained in $[0,\tau]$. Moreover these conditions also say that $|x|$ is bounded below by the linear functions ϕ_1 in I_1 and ϕ_2 in I_2 such that $\phi_1(a-\beta) = 0 = \phi_2(a+\beta)$ and $\phi_1(a) = \beta = \phi_2(a)$. Since for $j = 1, 2$ $\int_{I_j} \phi_j^m = \frac{\beta^{m+1}}{m+1}$ it follows that

$$\int_0^\tau x(s)^m ds \ge \int_{I_1 \cup I_2} x(s)^m ds = \frac{2}{m+1}\beta^{m+1}$$

which concludes the proof of the minimality.

We have thus proved (2.11). When the equality holds, $A = \tau$ in (2.4), and this implies $-\alpha + \beta\tau \ge 0$. Since we proved that $-\alpha + \beta\tau \le 0$, we have the equality. In particular, it is $\beta \ge \frac{\alpha}{\tau}$ which, together with the converse inequality just proved above, leads $\beta = \frac{\alpha}{\tau}$. Matching this equality with (2.4), we find that

$$\int_0^\tau x^m(s)ds = \beta^m \alpha.$$

This is only possible if and olny if $x \equiv \beta = ||x||_\infty$. Being $x(0) = 0$, we must have $x \equiv 0$ and consequently $u = 0$ and $v = 1$ a.e. □

REMARK 2.6. When m is odd (and greater or equal than 3), the argument used in the proof to obtain the inequality on the left is no longer valid. In fact, we will use the techniques developed in Chapter 4 to prove that γ is actually not a length-minimizing curve at least when $m = 3$.

CHAPTER 3

The End-Point Mapping

In this Chapter we introduce the fundamental object of our analysis: the End-point Mapping, which is intended to somehow measure "variations" along a reference curve, as we will explain below. We deduce necessary conditions for horizontal curves to be length-minimizing, namely Theorem 1.9 of Chapter 1 and the so-called second-order conditions. We always work under Assumptions 1.7.

Let $X_1, \dots, X_m$ be a set of vector fields spanning $\mathcal{D}$. Let γ be a reference $\mathcal{D}$-horizontal curve with control h and extremal points x_0 and x_1 (for example, an extremal cruve). For the sake of simplicity we assume that its domain is the interval $[0, 1]$, so that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. We shorten $\dot{\gamma}(t) = \sum_{j=1}^{m} h_j(t) X_j(\gamma(t))$ by $h(t) \cdot X(\gamma(t))$ and similar expressions in the same way.

3.1. First-order Conditions

The underlying idea of the below definitions consists in considering perturbations of the reference control h while leaving the initial point x_0 fixed.

Definition 3.1. Let $V = L^2([0, 1], \mathbb{R}^m)$. For any fixed t in $[0, 1]$ the *End-Point Mapping at the time t* is the map $\mathcal{E}_t : V \to \mathbb{R}^n$ defined by the following Cauchy problem

$$\begin{cases} \dot{\mathcal{E}}_t(v) = v(t) \cdot X(\mathcal{E}_t(v)) \\ \mathcal{E}_0(v) = x_0. \end{cases} \tag{3.1}$$

The End-Point Mapping is defined by $\mathcal{E} = \mathcal{E}_1$.

The *Extended End-Point Mapping* $\mathcal{F} : V \to \mathbb{R} \times \mathbb{R}^n$ is defined by

$$\mathcal{F}(v) = (\mathcal{L}(v), \mathcal{E}(v)) \tag{3.2}$$

where $\mathcal{L}$ is the normalized L^2-squared norm, $\mathcal{L}(v) = \frac{1}{2}||v||^2_{L^2}$.

Remark 3.2. As we assume a generic $\mathcal{D}$-horizontal curve to be Lipschitz-continuous (see Definition 1.2), the domain of $\mathcal{F}$ should be $V = L^\infty([0, 1], \mathbb{R}^m)$ while we took $V = L^2([0, 1], \mathbb{R}^m)$ which is a bigger space. The reason is that this is the "natural" domain for $\mathcal{F}$ because of the definition of $\mathcal{L}$, but of course we should include also absolutely continuous $\mathcal{D}$-horizontal curves in our analysis. This would not be a great effort; however, we point out that, regarding the definitions and the results presented in the following chapters, either this option or restricting $\mathcal{F}$ to $L^\infty([0, 1], \mathbb{R}^m)$ would be perfectly equivalent alternatives.

The crucial observation is then the following: from the definition of length-minimizing curve and of $\mathcal{E}$, $\mathcal{F}$ cannot be open at h, because otherwise we would find another control h' such that $\mathcal{E}(h') = \mathcal{E}(h)$ but $||h'||_{L^2} < ||h||_{L^2}$ and this would ultimately violate the minimality of γ. We want to investigate this necessary condition for $\mathcal{F}$. For computational reasons, though, it is more convenient to work with a modified version of the end-point mapping. We begin with the following:

DEFINITION 3.3. For any fixed t in $[0,1]$, *the flow relative to the control h at time t* is the map $\Phi_t : \mathbb{R}^n \to \mathbb{R}^n$ defined by the following Cauchy problem

$$\begin{cases} \dot{\Phi}_t(x) = h(t) \cdot X(\Phi_t(x)) \\ \Phi_0(x) = x. \end{cases} \tag{3.3}$$

Note that $\Phi_t(x_0) = \gamma(t)$ for all t. In practice, the construction of Φ_t is dual to the one of $\mathcal{E}_t$: we are varying the initial point of γ leaving fixed the control h.

DEFINITION 3.4. The *Modified End-Point Mapping at the time t* is defined by

$$\hat{\mathcal{E}}_t = \Phi_t^{-1} \circ \mathcal{E}_t. \tag{3.4}$$

The *Modified End-Point Mapping* is then $\hat{\mathcal{E}} = \hat{\mathcal{E}}_1 = \Phi_1^{-1} \circ \mathcal{E}$ and the *Modified Extended End-Point Mapping* is the function $\hat{\mathcal{F}} : V \to \mathbb{R} \times \mathbb{R}^n$ defined by

$$\hat{\mathcal{F}}(v) = \left(\mathcal{L}(v), \hat{\mathcal{E}}(v)\right). \tag{3.5}$$

Note that, being Φ_t a diffeomorphism for any t, the critical points (see Notation and Conventions, 0.0.2) of $\mathcal{E}$ (resp. $\mathcal{F}$) are exactly those of $\hat{\mathcal{E}}$ (resp. $\hat{\mathcal{F}}$).

REMARK 3.5. Note that, for any t in $[0,1]$, $\mathcal{E}_t(h) \equiv \gamma(t)$ and $\hat{\mathcal{E}}_t(h) \equiv x_0$ by definition of $\mathcal{E}_t$ and $\hat{\mathcal{E}}_t$.

We want to further investigate the non-openness of $\mathcal{F}$ (equivalently, of $\hat{\mathcal{F}}$) at h. The Open Mapping Theorem (see Appendix B, Theorem B.2) implies that $d\hat{\mathcal{F}}(h)$ cannot be surjective from $V = L^2([0,1], \mathbb{R}^m)$ to $\mathbb{R}^{n+1}$. We want to compute this differential. The formula for $d\mathcal{L}(h)$ is elementary:

$$d\mathcal{L}(h)[v] = \int_0^1 (h \cdot v) dt. \tag{3.6}$$

We turn to $d\hat{\mathcal{E}}_t$. Differentiating the identity $\mathcal{E}_t(v) = \Phi_t(\hat{\mathcal{E}}_t(v))$ with respect to t we get ($d_x\Phi_t$ denotes the differential of Φ_t with respect to the space variables only)

$$\begin{aligned} \dot{\mathcal{E}}_t(v) &= \dot{\Phi}_t(\hat{\mathcal{E}}_t(v)) + d_x\Phi_t(\hat{\mathcal{E}}_t(v)) \left[\dot{\hat{\mathcal{E}}}_t(v)\right] \\ &= h \cdot X(\Phi_t(\hat{\mathcal{E}}_t(v))) + d_x\Phi_t(\hat{\mathcal{E}}_t(v)) \left[\dot{\hat{\mathcal{E}}}_t(v)\right] \\ &= h \cdot X(\mathcal{E}_t(v)) + d_x\Phi_t(\hat{\mathcal{E}}_t(v)) \left[\dot{\hat{\mathcal{E}}}_t(v)\right]. \end{aligned}$$

On the other hand, by definition of $\mathcal{E}_t$

$$\dot{\mathcal{E}}_t(v) = v \cdot X(\mathcal{E}_t(v))$$

so we get the formula

$$\dot{\hat{\mathcal{E}}}_t(v) = d_x\Phi_t(\hat{\mathcal{E}}_t(v))^{-1} \left[(v-h) \cdot X(\mathcal{E}_t(v))\right].$$

By integrating with respect to the time variable from 0 to t and recalling once more the identity $\mathcal{E}_t(v) = \Phi_t(\hat{\mathcal{E}}_t(v))$, we obtain the following integral formula for $\hat{\mathcal{E}}_t(v)$:

$$\hat{\mathcal{E}}_t(v) = x_0 + \int_0^t d_x\Phi_\tau(\hat{\mathcal{E}}_\tau(v))^{-1} \left[(v-h) \cdot X(\Phi_\tau(\hat{\mathcal{E}}_\tau(v)))\right] d\tau. \tag{3.7}$$

Consequently,

(3.8)
$$\begin{aligned}\frac{\partial \hat{\mathcal{E}}_t(h+sv)}{\partial s} =& \frac{\partial}{\partial s}\int_0^t d_x\Phi_\tau(\hat{\mathcal{E}}_\tau(h+sv))^{-1}\Big[(sv)\cdot X(\Phi_\tau(\hat{\mathcal{E}}_\tau(h+sv)))\Big]\,d\tau \\ =& \frac{\partial}{\partial s}\left(s\int_0^t d_x\Phi_\tau(\hat{\mathcal{E}}_\tau(h+sv))^{-1}\Big[v\cdot X(\Phi_\tau(\hat{\mathcal{E}}_\tau(h+sv)))\Big]\,d\tau\right) \\ =& \int_0^t d_x\Phi_\tau(\hat{\mathcal{E}}_\tau(h+sv))^{-1}\Big[v\cdot X(\Phi_\tau(\hat{\mathcal{E}}_\tau(h+sv)))\Big]\,d\tau \\ &+ s\int_0^t \frac{\partial}{\partial s}\Big(d_x\Phi_\tau(\hat{\mathcal{E}}_\tau(h+sv))^{-1}\Big[v\cdot X(\Phi_\tau(\hat{\mathcal{E}}_\tau(h+sv)))\Big]\Big)\,d\tau.\end{aligned}$$

Since we are interested in evaluating the expression in $s=0$, we can (for now) save the computation of the second term. Now (see Remark 3.5) $\hat{\mathcal{E}}_\tau(h+sv)\big|_{s=0} = \hat{\mathcal{E}}_\tau(h) = x_0$ and $\Phi_\tau(\hat{\mathcal{E}}_\tau(h+sv))\big|_{s=0} = \mathcal{E}_\tau(h) = \gamma(\tau)$. We have found a formula for the differential of $\hat{\mathcal{E}}$ at h:

$$d\hat{\mathcal{E}}(h)[v] = \left.\frac{\partial \hat{\mathcal{E}}(h+sv)}{\partial s}\right|_{s=0} = \int_0^1 d_x\Phi_\tau(x_0))^{-1}\,[v\cdot X(\gamma(\tau))]\,d\tau. \tag{3.9}$$

Since $\hat{\mathcal{F}}$ is not surjective at h, we can take a nonzero covector in $(\operatorname{Im} d\hat{\mathcal{F}}(h))^\perp$, whose coordinates are, say, $(\lambda_0, \lambda(0))$ with $\lambda_0 \in \{0,1\}$. Then for any $v \in V = L^2([0,1],\mathbb{R}^m)$

$$0 = \langle(\lambda_0,\lambda(0)), d\hat{\mathcal{F}}(h)[v]\rangle = \lambda_0 d\mathcal{L}(h)[v] + \langle\lambda(0), d\hat{\mathcal{E}}(h)[v]\rangle. \tag{3.10}$$

We have that $d\mathcal{L}(h)[v]$ is as in (3.6) and $d\hat{\mathcal{E}}(h)[v]$ is as in (3.9), so equation (3.10) becomes

$$\begin{aligned}0 &= \int_0^1 \lambda_0 h\cdot v + \big\langle\lambda(0), d_x\Phi_\tau(x_0)^{-1}[v\cdot X(\gamma(\tau))]\big\rangle\,d\tau \\ &= \int_0^1 \sum_{j=1}^r v_j\left(\lambda_0 h_j + \big\langle\lambda(0), d_x\Phi_\tau(x_0)^{-1}[X_j(\gamma(\tau))]\big\rangle\right)d\tau.\end{aligned}$$

Since v is arbitrary, the Fundamental Lemma of Calculus of Variations allows us to conclude that $\lambda_0 h_j + \big\langle\lambda(0), d_x\Phi_t(x_0)^{-1}X_j(\gamma(t))\big\rangle = 0$ a.e. in $[0,1]$. Setting

$$\lambda(t) = d_x\Phi_t(x_0)^{-T}[\lambda(0)] \tag{3.11}$$

(where $(\cdot)^{-T}$ means inverse of the transpose), we conclude that for any $j = 1,\ldots,m$ and for almost every t in $[0,1]$

$$\lambda_0 h_j + \langle\lambda(t), X_j(\gamma(t))\rangle = 0. \tag{3.12}$$

Note that these equations coincide with equations (1.5) of Theorem 1.9. Note also that the nontriviality condition 1.6 is satisfied since $(\lambda_0,\lambda(0))$ is nonzero and $d_x\Phi_t(x_0)$ is a diffeomorphism. We now prove that λ is by all means a dual curve associated with γ by proving that λ satisfies the adjoint equation (1.3).

Starting from equation (3.11), we differentiate with respect to t the identity

$$d_x\Phi_t(x_0)^T[\lambda(t)] = \lambda(0)$$

obtaining

$$0=\frac{\partial}{\partial t}\left(d_x\Phi_t(x_0)^T[\lambda(t)]\right)=\frac{\partial}{\partial t}\Big(d_x\Phi_t(x_0)^T\Big)[\lambda(t)]+d_x\Phi_t(x_0)^T[\dot\lambda(t)]. \tag{3.13}$$

Now,

$$\begin{aligned}\frac{\partial}{\partial t}d_x\Phi_t(x_0)&=d_x\left(\frac{\partial}{\partial t}\Phi_t(x_0)\right)=d_x(h(t)\cdot X(\Phi_t(x_0)))\\&=h(t)\cdot d_x(X(\Phi_t(x_0)))=h(t)\cdot d_x(X(y))\Big|_{y=\Phi_t(x_0)}\circ d_x(\Phi_t(x_0))\\&=h(t)\cdot dX(\gamma(t))\circ d_x\Phi_t(x_0)\end{aligned}$$

Inserting this result in (3.13), and recalling that $d_x\Phi_t(x_0)$ is invertible since the flow relative to h is a diffeomorphism, we obtain

$$\begin{aligned}0&=h(t)\cdot(dX(\gamma(t))\circ d_x\Phi_t(x_0))^T[\lambda(t)]+d_x\Phi_t(x_0)^T[\dot\lambda(t)]\\&=h(t)\cdot d_x\Phi_t(x_0)^T[dX(\gamma(t))^T[\lambda(t)]]+d_x\Phi_t(x_0)^T[\dot\lambda(t)]\\&=h(t)\cdot dX(\gamma(t))^T[\lambda(t)]+\dot\lambda(t)\\&=\sum_{j=1}^m h_j(t)dX_j(\gamma(t))^T[\lambda(t)]+\dot\lambda(t)\\&=\sum_{j=1}^m h_j(t)\left\langle dX_j(\gamma(t))^T,\lambda(t)\right\rangle+\dot\lambda(t)\end{aligned}$$

that is exactly the adjoint equation (1.3).

Recall that (Definition 1.10) extremals were defined as $\mathcal{D}$-horizontal curves such that there exists an associated dual curve so that the thesis of the Maximum Principle holds. The above computations allow us to characterize them in terms of end-point mappings:

Theorem 3.6. *Given a distribution $\mathcal{D}$ on $\mathbb{R}^n$ generated by the vector fields $X_1,\dots X_m$, fix two point x_0 and x_1 in $\mathbb{R}^n$ and consider the horizontal curves $\gamma:[0,1]\to\mathbb{R}^n$ such that $\gamma(0)=x_0$ and $\gamma(1)=x_1$. Then, among these curves*

(1) *extremal curves are* exactly *the critical points of the extended end-pont mapping $\mathcal{F}$ (equivalently, of $\hat{\mathcal{F}}$)*
(2) *abnormal extremals are* exactly *the critical points of the end-point mapping $\mathcal{E}$ (equivalently, of $\hat{\mathcal{E}}$).*

Proof. If $d\mathcal{F}$ is not surjective at the control h of γ, the above discussion leads through the construction of a dual curve λ associated with γ. Viceversa, if $\lambda(t)$ is a dual curve, it is possible to proceed backward from equation (3.12) to equation (3.10) and, due to the nontriviality condition (1.6), deduce that $d\mathcal{F}$ is not surjective at h. This proves (1).

As for (2), being $\lambda_0=0$ in (3.12) and (3.10), the non-surjectivity of $\mathcal{F}$ at h is equivalent to the non-sujectivity of $\mathcal{E}$ at h. □

Remark 3.7. Note that for both $\mathcal{E}$ and $\mathcal{F}$ (equivalently, $\hat{\mathcal{E}}$ and $\hat{\mathcal{F}}$), their range is a finite-dimensional vector space, while their domain is $L^2([0,1],\mathbb{R}^m)$ which is infinite-dimensional. This implies that, independently on h, their coimages (see Notation and Conventions, 0.0.1) $\operatorname{Coim} d\mathcal{E}(h)$ and $\operatorname{Coim} d\mathcal{F}(h)$ are finite-dimensional subspaces of V. We will make use of this fact later on.

The results obtained in this subsections involve first derivatives of the end-point mapping, and are consequently called *first-order conditions*. In the following section, we deal with *second-order conditions* by working on second order derivatives of $\mathcal{E}$.

3.2. Second-Order Conditions

Since we know a second-order Open Mapping Theorem (see Appendix B, the result is proved in [**8**]) it is possible deduce further necessary conditions (or, equivalently sufficient conditions for nonminimiality) for a length-minimizing curve with control h based on second-order derivatives of of $\hat{\mathcal{E}}$. The immediate adaptation of Theorem B.4 to our context is then:

. *Let $\gamma : [0,1] \to \mathbb{R}^n$ be a $\mathcal{D}$-horizontal curve with extremal points x_0 and x_1 and control h. If $d\hat{\mathcal{E}}(h)$ has corank r and the restricted Hessian $H_{\hat{\mathcal{E}}}(h)$ has index greater or equal than r at h:*

$$\operatorname{ind} H_{\hat{\mathcal{E}}}(h) \geq r$$

then $\hat{\mathcal{E}}$ is open at h, and thus γ is not a length-minimizing curve.

We want to perform some computations of the second derivative of $\hat{\mathcal{E}}$ in order to get some more handful formulas for the Hessian of $\hat{\mathcal{E}}$ at h. First of all, we compute second-order directional derivatives of $\hat{\mathcal{E}}$ at h along a control v resuming from formula (3.8).

In what follows, v is a control in $V = L^2([0,1], \mathbb{R}^m)$. Define the following time-dependent vector field $\Psi_{t,v} : \mathbb{R}^n \to \mathbb{R}^n$ by setting, for any t in $[0,1]$

$$\Psi_{t,v}(x) = d_x\Phi_t(x)^{-1}[v(t) \cdot X(\Phi_t(x))]. \tag{3.14}$$

In this way, (3.8) becomes easier to write:

$$\frac{\partial \hat{\mathcal{E}}_t(h+sv)}{\partial s} = \int_0^t \Psi_{\tau,v}(\hat{\mathcal{E}}_\tau(h+sv))d\tau + s\int_0^t \frac{\partial}{\partial s}(\Psi_{\tau,v}(\hat{\mathcal{E}}_\tau(h+sv)))d\tau$$

and then

$$\frac{\partial^2 \hat{\mathcal{E}}_t(h+sv)}{\partial s^2} = 2\int_0^t \frac{\partial}{\partial s}\Psi_{\tau,v}(\hat{\mathcal{E}}_\tau(h+sv))d\tau + s\int_0^t \frac{\partial^2}{\partial s^2}(\Psi_{\tau,v}(\hat{\mathcal{E}}_\tau(h+sv)))d\tau. \tag{3.15}$$

Again, the fact that we will evaluate the expression in $s = 0$ allows us to focus only on the first term. Now, the chain rule yields

$$\frac{\partial}{\partial s}\Psi_{t,v}(\hat{\mathcal{E}}_t(h+sv)) = d_x\Psi_{t,v}(\hat{\mathcal{E}}_t(h+sv))\left[\frac{\partial}{\partial s}\hat{\mathcal{E}}_t(h+sv)\right]$$

so

$$\begin{aligned}\left.\frac{\partial}{\partial s}\Psi_{t,v}(\hat{\mathcal{E}}_t(h+sv))\right|_{s=0} &= d_x\Psi_{t,v}(x_0)\left[\left.\frac{\partial}{\partial s}\hat{\mathcal{E}}_t(h+sv)\right|_{s=0}\right] \\ &= d_x\Psi_{t,v}(x_0)\left[\int_0^t \Psi_{\tau,v}(x_0)d\tau\right].\end{aligned}$$

Consequently, (3.15) evalutaed in $s = 0$ becomes

$$\begin{aligned}\left.\frac{\partial^2 \hat{\mathcal{E}}_t(h+sv)}{\partial s^2}\right|_{s=0} &= 2\int_0^t \frac{\partial}{\partial s}\Psi_{\tau,v}(\hat{\mathcal{E}}_\tau(h+sv))\Big|_{s=0} d\tau \\ &= 2\int_0^t \int_0^{\tau_1} d_x\Psi_{\tau_1,v}(x_0)[\Psi_{\tau_2,v}(x_0)]d\tau_2 d\tau_1.\end{aligned} \tag{3.16}$$

This formula for the second directional derivative of $\hat{\mathcal{E}}_t$ at h along v holds de facto for general controls h and v. Theorem B.4 tells us to look at the *restricted* Hessian oh $\hat{\mathcal{E}}$: we consequently assume that $v \in \operatorname{Ker} d\hat{\mathcal{E}}(h) = \operatorname{Ker} d\mathcal{E}(h)$, which means by equation (3.9) that

$$\int_0^1 \Psi_{\tau,v}(x_0)d\tau = 0.$$

This assumption allows us to make a very clever computation over equation (3.16) when $t = 1$:

$$\begin{aligned}\left.\frac{\partial^2 \hat{\mathcal{E}}(h+sv)}{\partial s^2}\right|_{s=0} &= 2\int_0^1 \int_0^{\tau_1} d_x\Psi_{\tau_1,v}(x_0)[\Psi_{\tau_2,v}(x_0)]d\tau_2 d\tau_1 \\ &= -2\int_0^1 \int_{\tau_1}^{1} d_x\Psi_{\tau_1,v}(x_0)[\Psi_{\tau_2,v}(x_0)]d\tau_2 d\tau_1 \\ &= -2\int_0^1 \int_0^{\tau_2} d_x\Psi_{\tau_1,v}(x_0)[\Psi_{\tau_2,v}(x_0)]d\tau_1 d\tau_2\end{aligned} \tag{3.17}$$

where in the last step we changed the order of integration. Now, renaming the variables, we deduce that (recall the formula for the commutator (0.2) in Notation and Conventions)

$$\begin{aligned}\left.\frac{\partial^2 \hat{\mathcal{E}}(h+sv)}{\partial s^2}\right|_{s=0} &= \int_0^1 \int_0^{\tau_1} d_x\Psi_{\tau_1,v}(x_0)[\Psi_{\tau_2,v}(x_0)] - d_x\Psi_{\tau_2,v}(x_0)[\Psi_{\tau_1,v}(x_0)]d\tau_2 d\tau_1 \\ &= \int_0^1 \int_0^{\tau_1} [\Psi_{\tau_2,v}, \Psi_{\tau_1,v}](x_0)d\tau_2 d\tau_1.\end{aligned}$$

We have proved the following formula for the Hessian of $\hat{\mathcal{E}}$ at h:

$$\operatorname{Hess}\hat{\mathcal{E}}(h)[v] = \int_0^1 \int_0^{\tau_1} [\Psi_{\tau_2,v}, \Psi_{\tau_1,v}](x_0)d\tau_2 d\tau_1. \tag{3.18}$$

for any v in $\operatorname{Ker} d\hat{\mathcal{E}}(h)$.

REMARK 3.8. Note that, aside from the computation of the integral, formula (3.18) can be computed explicitly if we are able to compute the flow relative to h, Φ_t (see Definition 3.3), because in that case $\Psi_{t,v}$, defined through (3.14), can be explicitely written. However, even for corank-one extremals, it may be difficult to determine the index of $H_{\hat{\mathcal{E}}}(h)$ as the case presented in the following subsection demonstrates.

3.2.1. Hessian for a $C^{1,1/2}$ abnormal extremal. We compute $\operatorname{Hess}\hat{\mathcal{E}}(h)$ and $H_{\hat{\mathcal{E}}}(h)$ for the extremal introduced in Section 2.3 when $m = 2$ dividing our computations in steps. Recall that the curve is defined by its two first components $(\gamma_1(t), \gamma_2(t)) = (t, t^{3/2})$, has control $h = (1, \frac{3}{2}t^{1/2})$ and passes through 0 at $t = 0$.

Step 1: computation of the flow $\Phi_{t,v}$. We compute the flow relative to h at the time t, $\Phi_t = (\Phi_t^1, \dots, \Phi_t^5)$ which is the solution to the Cauchy problem

$$\begin{cases} \dot{\Phi}_t(x) = h_1 X_1(\Phi_t(x)) + h_2 X_2(\Phi_t(x)) \\ \Phi_0(x) = x \end{cases}$$

that is

$$\begin{cases} \dot{\Phi}_t^1(x) = 1 \\ \dot{\Phi}_t^2(x) = \frac{3}{2} t^{1/2} \\ \dot{\Phi}_t^3(x) = \frac{3}{2} t^{1/2} \Phi_t^1(x) \\ \dot{\Phi}_t^4(x) = \frac{3}{2} t^{1/2} (\Phi_t^1(x))^4 \\ \dot{\Phi}_t^5(x) = \frac{3}{2} t^{1/2} \Phi_t^1(x)(\Phi_t^1(x))^2 \\ (\Phi_0^1(x), \dots, \Phi_0^5(x)) = (x_1, \dots, x_5). \end{cases}$$

The integration of this system is elementary and leads to

$$\Phi_t(x) = \begin{pmatrix} t + x_1 \\ t^{3/2} + x_2 \\ \frac{3}{5} t^{5/2} + t^{3/2} x_1 + x_3 \\ \frac{3}{11} t^{11/2} + \frac{4}{3} t^{9/2} x_1 + \frac{18}{7} t^{7/2} x_1^2 + \frac{12}{5} t^{5/2} x_1^3 + t^{3/2} x_1^4 + x_4 \\ \frac{3}{11} t^{11/2} + \frac{1}{3} t^{9/2} x_1 + \frac{3}{4} t^4 x_2 + + t^3 x_1 x_2 + \frac{3}{5} t^{5/2} x_2^2 + t^{3/2} x_1 x_2^2 + x_5 \end{pmatrix}.$$

Step 2: computation of $\Psi_{t,v}$ and of $d\hat{\mathcal{E}}(h)$. We compute the vector field $\Psi_{t,v}$ defined through formula (3.14). The Jacobian matrix with respect to the space variable of Φ_t is

$$D_x \Phi_t(x) =$$
$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ t^{3/2} & 0 & 1 & 0 & 0 \\ \frac{4}{3} t^{9/2} + \frac{36}{7} x_1 t^{7/2} + \frac{36}{5} x_1^2 t^{5/2} + 4 x_1^3 t^{3/2} & 0 & 0 & 1 & 0 \\ \frac{1}{3} t^{9/2} + x_2 t^3 + x_2^2 t^{3/2} & \frac{3}{4} t^4 + x_1 t^3 + \frac{6}{5} x_2 t^{5/2} + 2 x_1 x_2 t^{3/2} & 0 & 0 & 1 \end{pmatrix}$$

and so the vector field is

$$\Psi_{t,v}(x) = D_x \Phi_t(x)^{-1} [v(t) \cdot X(\Phi_t(x))] =$$
$$\begin{pmatrix} v_1(t) \\ v_2(t) \\ v_2(t)(t + x_1) - t^{3/2} v_1(t) \\ v_2(t)(t + x_1)^4 - \frac{4}{105} t^{3/2} v_1(t) \left(35 t^3 + 135 t^2 x_1 + 189 t x_1^2 + 105 x_1^3\right) \\ -t^{3/2} v_1(t) x_2^2 - \frac{1}{3} t^{9/2} v_1(t) + \frac{4}{5} t^{5/2} v_2(t) x_2 + \frac{1}{4} t^4 v_2(t) - t^3 v_1(t) x_2 + t v_2(t) x_2^2 + v_2(t) x_1 x_2^2 \end{pmatrix}.$$

The initial point is $x_0 = 0$, so

$$\Psi_{t,v}(0) = \begin{pmatrix} v_1(t) \\ v_2(t) \\ t v_2(t) - t^{3/2} v_1(t) \\ t^4 v_2(t) - \frac{4}{3} t^{9/2} v_1(t) \\ \frac{1}{4} t^4 v_2(t) - \frac{1}{3} t^{9/2} v_1(t) \end{pmatrix}. \tag{3.19}$$

Being $d\hat{\mathcal{E}}(h)[v] = \int_0^1 \Psi_{\tau,v}(0) d\tau$, it is clear that $\operatorname{Coker} d\hat{\mathcal{E}}(h) \simeq (\operatorname{Im} d\hat{\mathcal{E}}(h))^\perp = \operatorname{span}_{\mathbb{R}}\{\lambda\}$ with $\lambda = (0, 0, 0, 1, -4)$, in particular the corank of this extremal is one.

Moreover, v is in $\operatorname{Ker} d\hat{\mathcal{E}}(h)$ if and only if it satisfies the following integral conditions:

$$\begin{cases} \int_0^1 v_1(\tau)d\tau = 0 \\ \int_0^1 v_2(\tau)d\tau = 0 \\ \int_0^1 (\tau v_2(\tau) - \tau^{3/2}v_1(\tau)d\tau) = 0 \\ \int_0^1 (\tau^4 v_2(\tau) - \frac{4}{3}\tau^{9/2}v_1(\tau))d\tau = 0 \\ \int_0^1 (\frac{1}{4}\tau^4 v_2(\tau) - \frac{1}{3}\tau^{9/2}v_1(\tau))d\tau = 0. \end{cases} \tag{3.20}$$

Step 3: Computation of the Hessian. The formula for $\operatorname{Hess}\hat{\mathcal{E}}(h)$ is (3.18) for any v in $\operatorname{Ker} d\hat{\mathcal{E}}(h)$. Recalling formula (0.2) for the commutator the term inside the integral is:

$$[\Psi_{\tau_2,v}, \Psi_{\tau_1,v}](0) = D_x\Psi_{\tau_1,v}(0)[\Psi_{\tau_2,v}(0)] - D_x\Psi_{\tau_2,v}(0)[\Psi_{\tau_1,v}(0)] =$$
$$\begin{pmatrix} 0 \\ 0 \\ v_1(\tau_2)v_2(\tau_1) - v_1(\tau_1)v_2(\tau_2) \\ -\frac{4}{7}(-7\tau_1^3 v_1(\tau_2)v_2(\tau_1) + v_1(\tau_1)(9(\tau_1^{7/2} - \tau_2^{7/2})v_1(\tau_2) + 7\tau_2^3 v_2(\tau_2))) \\ \tau_2^3 v_1(\tau_2)v_2(\tau_1) - \frac{1}{5}(5\tau_1 v_1(\tau_1) + 4(-\tau_1^{5/2} + \tau_2^{5/2})v_2(\tau_1))v_2(\tau_2) \end{pmatrix}.$$

$\operatorname{Hess}\hat{\mathcal{E}}(h)$ is then just the double integral in τ_1 and τ_2 of the above vector-valued function.

We turn to the restricted Hessian $H_{\hat{\mathcal{E}}}(h)$.

$$H_{\hat{\mathcal{E}}}(h)[v] = \left\langle \lambda, \operatorname{Hess}\hat{\mathcal{E}}(h)[v] \right\rangle =$$
$$\int_0^1 \int_0^{\tau_1} \frac{4}{5}v_2(\tau_1)(5(\tau_1^3 - \tau_2^3)v_1(\tau_2) + 4(-\tau_1^{5/2} + \tau_2^{5/2})v_2(\tau_2)$$
$$+ \frac{4}{7}v_1(\tau_1)(9(-\tau_1^{7/2} + \tau_2^{7/2})v_1(\tau_2) + 7(\tau_1^3 - \tau_2^3)v_2(\tau_2))d\tau_2 d\tau_1$$
$$= \int_0^1 \int_0^{\tau_1} \begin{pmatrix} v_1(\tau_2) \\ v_2(\tau_2) \end{pmatrix} \cdot \begin{pmatrix} 36(-\tau_1^{7/2} + \tau_2^{7/2}) & 4(\tau_1^3 - \tau_2^3) \\ 4(\tau_1^3 - \tau_2^3) & +16(-\tau_1^{5/2} + \tau_2^{5/2}) \end{pmatrix} \begin{pmatrix} v_1(\tau_1) \\ v_2(\tau_1) \end{pmatrix} d\tau_2 d\tau_1.$$

The last line is intended to to emphasize the fact that the map is indeed quadratic in the argument v. However, since v must satisfy also conditions (3.20), it is difficult to establish if it is definite, semidefinite or indefinite.

3.2.2. Goh Condition. Although determining directly the index of $H_{\mathcal{E}}(h)$ can be in general a difficult task, thanks to formula (3.18) it is possible to deduce a very clever consequence of Theorem B.4, known as the Goh condition.

THEOREM 3.9 (The Goh Condition). *Let $\gamma : [0,1] \to \mathbb{R}^n$ be an abnormal length-minimizing curve with control h for the distribution $\mathcal{D}$ in $\mathbb{R}^n$ spanned by the vector fields $X_1, \ldots X_m$ and let r be the corank of $d\hat{\mathcal{E}}(h)$. Then there exists a dual curve $\lambda : [0,1] \to T^*\mathbb{R}^n$ associated with γ such that for all i,j in $\{1,\ldots,m\}$*

$$\langle \lambda(t), [X_i, X_j](\gamma(t)) \rangle = 0 \tag{3.21}$$

for any $t \in [0,1]$.

PROOF. Theorem B.4 implies that, as $\hat{\mathcal{E}}$ is not open at h, the index of $H_{\hat{\mathcal{E}}}$ must be less than r, so there must be a nonzero covector $\lambda(0)$ in $(\operatorname{Im}\hat{\mathcal{E}}(h))^\perp$ such that

$$\operatorname{ind}\left(\left\langle \lambda(0), \operatorname{Hess}\hat{\mathcal{E}}(h)\big|_{\operatorname{Ker} d\hat{\mathcal{E}}(h)} \right\rangle\right) < r.$$

Let λ be the dual curve defined through (3.11). Assume by contradiction that (3.21) is false: we find a t_0 in $[0,1]$ and i,j in $\{1,\ldots,m\}$ such that

$$\langle \lambda(t_0), [X_i, X_j](\gamma(t_0))\rangle \neq 0. \tag{3.22}$$

We prove that if this condition holds then $\mathrm{ind}(\langle \lambda(0), \mathrm{Hess}\,\hat{\mathcal{E}}(h)\big|_{\mathrm{Ker}\,d\hat{\mathcal{E}}(h)}\rangle)$ is greater or equal than r. We pick a nonconstant control v in $L^2([0,1],\mathbb{R}^m)$ of the following form:

$$v(t) = v_i(t)\mathbf{e_i} + v_j(t)\mathbf{e_j} \tag{3.23}$$

where $\mathbf{e_i}, \mathbf{e_j}$ are the i-th and j-th vector of the canonical basis of $\mathbb{R}^m$ and v_i, v_j are real functions in $L^2([0,1])$ to be specified later. We localize it in t_0 by setting, for small $\epsilon > 0$

$$v^\epsilon(t) = v\left(\frac{t-t_0}{\epsilon}\right) \tag{3.24}$$

extending it by 0 when the member on the right of (3.24) is not defined. Note that v^ϵ is supported in $[t_0, t_0+\epsilon]$ and since v is nonconstant, the infinite family $\{v^\epsilon\}_{\epsilon>0}$ is linearly independent. For this reason, as $\mathrm{Ker}\,d\hat{\mathcal{E}}(h)$ has finite complementary subspace in $L^2([0,1],\mathbb{R}^m)$ (see Remark 3.7) there are infinite values of ϵ arbitrarily close to 0 such that v^ϵ is in $\mathrm{Ker}\,d\hat{\mathcal{E}}(h)$. We compute $\mathrm{Hess}\,\hat{\mathcal{E}}(h)$ along one of these controls. By a change of variables we obtain

$$\begin{aligned}
\mathrm{Hess}\,\hat{\mathcal{E}}(h)[v^\epsilon] &= \int_0^1 \int_0^{\tau_1} [\Psi_{\tau_2,v^\epsilon}, \Psi_{\tau_1,v^\epsilon}](x_0)d\tau_2 d\tau_1 \\
&= \int_{t_0}^{t_0+\epsilon} \int_{t_0}^{\tau_1} [\Psi_{\tau_2,v^\epsilon}, \Psi_{\tau_1,v^\epsilon}](x_0)d\tau_2 d\tau_1 \\
&= \epsilon^2 \int_0^1 \int_0^{\tau_1} [\Psi_{t_0+\epsilon\tau_2,v(\tau_2)}, \Psi_{t_0+\epsilon\tau_1,v(\tau_1)}](x_0)d\tau_2 d\tau_1
\end{aligned}$$

where $\Psi_{t_0+\epsilon\tau,v(t)} = d_x\Phi_{t_0+\epsilon t}(x_0)^{-1}[v(t)\cdot X(\Psi_{t_0+\epsilon t}(x_0))]$ - note the different dependence on the time variable comparing it with the definition of $\Psi_{t,v}$ given in equation (3.14).

Now, if $\{\epsilon_k\}_k$ is any sequence converging to zero such that v^{ϵ_k} is in $\mathrm{Ker}\,d\hat{\mathcal{E}}(h)$ for any k, we compute the limit:

$$\begin{aligned}
\lim_{k\to\infty} \Psi_{t_0+\epsilon_k t, v(t)}(x_0) &= \lim_{k\to\infty} d_x\Phi_{t_0+\epsilon_k t}(x_0)^{-1}[v(t)\cdot X(\Psi_{t_0+\epsilon_k t}(x_0))] \\
&= d_x\Phi_{t_0}(x_0)^{-1}[v(t)\cdot X(\Phi_{t_0}(x_0))]
\end{aligned}$$

so, passing to the limit under the integral sign

$$\lim_{k\to\infty}\frac{\operatorname{Hess}\hat{\mathcal{E}}(h)[v^{\epsilon_k}]}{{\epsilon_k}^2}=$$
$$=\int_0^1\int_0^{\tau_1}[d_x\Phi_{t_0}^{-1}[v(\tau_2)\cdot X(\Phi_{t_0})],d_x\Phi_{t_0}^{-1}[v(\tau_1)\cdot X(\Phi_{t_0})]](x_0)d\tau_2 d\tau_1$$
$$=\int_0^1\int_0^{\tau_1}d_x\Phi_{t_0}^{-1}(x_0)[[v(\tau_2)\cdot X,v(\tau_1)\cdot X]](\Phi_{t_0}(x_0))d\tau_2 d\tau_1$$
$$=d_x\Phi_{t_0}^{-1}(x_0)\left[\int_0^1\int_0^{\tau_1}[v(\tau_2)\cdot X,v(\tau_1)\cdot X](\Phi_{t_0}(x_0))d\tau_2 d\tau_1\right].$$

Recalling the form of v in (3.23),

$$[v(\tau_2)\cdot X,v(\tau_1)\cdot X](\Phi_{t_0}(x_0))=$$
$$=[(v_i(\tau_2)\mathbf{e_i}+v_j(\tau_2)\mathbf{e_j})\cdot X,(v_i(\tau_1)\mathbf{e_i}+v_j(\tau_1)\mathbf{e_j})\cdot X](\Phi_{t_0}(x_0))$$
$$=v_i(\tau_2)v_j(\tau_1)[X_i,X_j](\Phi_{t_0}(x_0))+v_j(\tau_2)v_i(\tau_1)[X_j,X_i](\Phi_{t_0}(x_0))$$
$$=(v_i(\tau_2)v_j(\tau_1)-v_j(\tau_2)v_i(\tau_1))\,[X_i,X_j](\Phi_{t_0}(x_0))$$

so

$$\lim_{k\to\infty}\frac{\operatorname{Hess}\hat{\mathcal{E}}(h)[v^{\epsilon_k}]}{{\epsilon_k}^2}=$$
$$=d_x\Phi_{t_0}^{-1}(x_0)\left[\int_0^1\int_0^{\tau_1}(v_i(\tau_2)v_j(\tau_1)-v_j(\tau_2)v_i(\tau_1))\,[X_i,X_j](\Phi_{t_0}(x_0))d\tau_2 d\tau_1\right]$$
$$=d_x\Phi_{t_0}^{-1}(x_0)[[X_i,X_j](\Phi_{t_0}(x_0))]\int_0^1\int_0^{\tau_1}(v_i(\tau_2)v_j(\tau_1)-v_j(\tau_2)v_i(\tau_1))\,d\tau_2 d\tau_1.$$

By the definition of the dual curve λ in equation (3.11) we have

$$\left\langle\lambda(0),d_x\Phi_{t_0}^{-1}(x_0)[[X_i,X_j](\Phi_{t_0}(x_0))]\right\rangle=\langle\lambda(t_0),[X_i,X_j](\gamma(t_0))\rangle$$

we infer from the above computation that

$$\lim_{k\to\infty}\frac{1}{{\epsilon_k}^2}\left\langle\lambda,\operatorname{Hess}\hat{\mathcal{E}}(h)[v^{\epsilon_k}]\right\rangle=$$
$$\langle\lambda(t_0),[X_i,X_j](\gamma(t_0))\rangle\int_0^1\int_0^{\tau_1}(v_i(\tau_2)v_j(\tau_1)-v_j(\tau_2)v_i(\tau_1))\,d\tau_2 d\tau_1.$$

Now comes the clever choice of the control v. First of all, recall that we are assuming that (3.22) holds true - for example, $\langle\lambda(t_0),[X_i,X_j](\gamma(t_0))\rangle>0$. Take any nonzero sequence $\{a_h\}_h$ in $l^2(\mathbb{R})$ and set, for t in $[0,1]$

$$v_i(t)=\sum_{h=1}^{\infty}a_h\sin(2\pi ht),\qquad v_j(t)=\sum_{h=1}^{\infty}a_h\cos(2\pi ht). \tag{3.25}$$

The control $v=v_i\mathbf{e_i}+v_j\mathbf{e_j}$ is in $L^2([0,1],\mathbb{R}^m)$ and since

$$\int_0^1\int_0^{\tau_1}(v_i(\tau_2)v_j(\tau_1)-v_j(\tau_2)v_i(\tau_1))\,d\tau_2 d\tau_1=\frac{1}{2\pi}\sum_{h=1}^{\infty}\frac{a_h^2}{h}$$

we conclude that

$$\lim_{k\to\infty} \frac{1}{{\epsilon_k}^2} \langle \lambda, \operatorname{Hess} \hat{\mathcal{E}}(h)[v^{\epsilon_k}] \rangle = \langle \lambda(t_0), [X_i, X_j](\gamma(t_0)) \rangle \frac{1}{2\pi} \sum_{h=1}^{\infty} \frac{a_h^2}{h} > 0.$$

Due to the continuous dependence on ϵ of $\operatorname{Hess} \hat{\mathcal{E}}(h)[v^\epsilon]$ there must be a sufficiently large k (possibly depending on $\{a_h\}_h$) for which $\langle \lambda, \operatorname{Hess} \hat{\mathcal{E}}(h)[v^{\epsilon_k}] \rangle > 0$. Take now r linearly independent sequences $\{a_h^1\}_h, \dots, \{a_h^r\}_h$ in $l^2(\mathbb{R})$ and construct r controls $v_1, \dots, v_r$ through (3.25): for a sufficiently small k such that

- $\langle \lambda, \operatorname{Hess} \hat{\mathcal{E}}(h)[v_p^{\epsilon_k}] \rangle > 0$ for $p = 1, \dots, r$
- $v_p^{\epsilon_k}$ all belong to $\operatorname{Ker} d\hat{\mathcal{E}}(h)$ for $p = 1, \dots r$

we deduce that there exists a r-dimensional subspace of $\operatorname{Ker} d\hat{\mathcal{E}}(h)$ in which $\operatorname{Hess} \hat{\mathcal{E}}(h)$ is positive definite. Exactly by the same argument applied to the control v defined through

$$v_i(t) = \sum_{h=1}^{\infty} a_h \sin(2\pi h t), \qquad v_j(t) = -\sum_{h=1}^{\infty} a_h \cos(2\pi h t).$$

we deduce that there exists a r-dimensional subspace of $\operatorname{Ker} d\hat{\mathcal{E}}(h)$ in which $\operatorname{Hess} \hat{\mathcal{E}}(h)$ is negative definite. But then $\operatorname{ind}\left(\langle \lambda, \operatorname{Hess} \hat{\mathcal{E}}(h)\big|_{\operatorname{Ker} d\hat{\mathcal{E}}(h)} \rangle\right)$ is greater or equal than r. We have reached the claimed consequence of hypotesis (3.22). This gives rise to a contradition, since Theorem B.4 implies that $\hat{\mathcal{E}}$ is open and γ is length-minimizing.

□

As an immediate consequence we have:

Corollary 3.10. *Any distribution $\mathcal{D}$ in $\mathbb{R}^n$ of step 2 has no abnormal length-minimizing curves.*

Proof. In fact, if $\mathcal{D}$ is spanned by the vector fields $X_1, \dots, X_m$ and $\gamma : [0,1] \to \mathbb{R}^n$ were an abnormal extremal with associated dual curve $\lambda : [0,1] \to T^*\mathbb{R}^n$, the minimization condition (1.5) together with the above Goh condition (3.21) would imply

$$\begin{cases} \langle \lambda(t), X_j(\gamma(t)) \rangle = 0 & \text{for } j = 1, \dots, m \\ \langle \lambda(t), [X_i, X_j](\gamma(t)) \rangle = 0 & \text{for } i, j = 1, \dots, m \end{cases}$$

but since $\mathcal{D}$ is of rank 2, the spanning vector fields and their commutator of degree 1 form a basis of $T_x\mathbb{R}^n$ for any x in $\mathbb{R}^n$, so this would imply $\lambda(t) = 0$ for any t in $[0,1]$. This would contradict the nontriviality condition (1.6).

□

CHAPTER 4

A Third-Order Open Mapping Theorem

In the previous Chapter, we have seen that necessary conditions involving the End-Point Mapping are based on open mapping theorems. Guided by this outline, we now prove a third-order Open Mapping Theorem, that has anyway its own interest. We finally prove its relevance to our situation by proving the nonminimality of an abnormal extremal presented in Section 2.4 of Chapter 2.

4.1. A Third-Order Open Mapping Theorem

The fundamental step in proving the third-order Open Mapping Theorem we are interested in is the following Euclidean, corank-one case. The notion of restricted Hessian and of index of a quadratic form are specified on Appendix B.

THEOREM 4.1. *Let n, N be positive integers with $N \geq n$, and let $F : \mathbb{R}^N \to \mathbb{R}^n$ be a smooth map with $F(0) = 0$. Suppose that $dF(0)$ has corank 1, that the restricted Hessian of F at 0*

$$H_F(0) : \operatorname{Ker} dF(0) \to \operatorname{Coker} dF(0)$$

is semidefinite and that the isotropic space of H_F at 0

$$\operatorname{Iso} H_F(0) = \{w \in \operatorname{Ker} dF(0) : H_F(0)[w] = 0\}$$

is nontrivial. If there exists a nonzero vector v in $\operatorname{Iso} H_F(0)$ such that the projection over $\operatorname{Coker} dF(0)$ of the third-order derivative of F along v:

$$\pi_{\operatorname{Coker} dF(0)}\left(\frac{\partial^3 F}{\partial v^3}(0)\right) = \lim_{t\to 0} \pi_{\operatorname{Coker} dF(0)}\left(\frac{F(tv) - F(0)}{t^3}\right)$$

is nonzero, then F is open at 0.

PROOF. The underlying idea of this proof consists in composing F with a suitable "perturbation map" and prove the openness of this composition. This approach was inspired by the proof of Theorem 20.3 in [**8**].

Let $\epsilon > 0$ be a real parameter, x_0 and x_1 two vectors in $\mathbb{R}^N$ to be chosen later. We define the map ϕ_ϵ from $\operatorname{Coim} dF(0) \oplus \operatorname{span}_{\mathbb{R}}\{v\}$ to $\mathbb{R}^N$ by setting

$$\phi_\epsilon(x, y) = \frac{\epsilon^3 y^3}{3!} v + \frac{\epsilon^6 y^6}{6!} x_0 + \frac{\epsilon^9 y^9}{9!} x_1 + \frac{\epsilon^9}{9!} x \tag{4.1}$$

for any $(x, y) \in \operatorname{Coim} dF(0) \oplus \operatorname{span}_{\mathbb{R}}\{v\}$. Note that ϕ is smooth in x, y and ϵ and $\phi_\epsilon(0) = 0$. Consider now the composition of F with ϕ_ϵ:

$$\Phi_\epsilon(x, y) = F(\phi_\epsilon(x, y)) \tag{4.2}$$

proving that Φ_ϵ is open at 0 (for some nonzero fixed $\epsilon > 0$) would guarantee that F is a fortiori open at 0. We will achieve this goal basically by computing enough derivatives of Φ_ϵ with respect to ϵ and studying its Taylor expansion. In the following computations,we omit the ϵ-dependence of ϕ_ϵ and we simply write

ϕ. We denote the derivatives with respect to ϵ of ϕ and Φ by $\phi', \phi'', \dots, \phi^{(j)}$ and $\Phi'_\epsilon, \Phi''_\epsilon, \dots, \Phi^{(j)}_\epsilon$ respectively.

We are going to differentiate Φ_ϵ nine times. We first group the derivatives of ϕ in a table for future use.

derivative of ϕ w.r.t. ϵ	eval. in $\epsilon = 0$	behavior as $\epsilon \to 0$
$\phi' = \frac{\epsilon^2}{2!} y^3 v + \frac{\epsilon^5}{5!} y^6 x_0 + \frac{\epsilon^8}{8!} y^9 x_1 + \frac{\epsilon^8}{8!} x$	0	$O(\epsilon^2)$
$\phi'' = \epsilon y^3 v + \frac{\epsilon^4}{4!} y^6 x_0 + \frac{\epsilon^7}{7!} y^9 x_1 + \frac{\epsilon^7}{7!} x$	0	$O(\epsilon^1)$
$\phi''' = y^3 v + \frac{\epsilon^3}{3!} y^6 x_0 + \frac{\epsilon^6}{6!} y^9 x_1 + \frac{\epsilon^6}{6!} x$	$y^3 v$	nonvanishing
$\phi^{(4)} = \frac{\epsilon^2}{2!} y^6 x_0 + \frac{\epsilon^5}{5!} y^9 x_1 + \frac{\epsilon^5}{5!} x$	0	$O(\epsilon^2)$
$\phi^{(5)} = \epsilon y^6 x_0 + \frac{\epsilon^4}{4!} y^9 x_1 + \frac{\epsilon^4}{4!} x$	0	$O(\epsilon^1)$
$\phi^{(6)} = y^6 x_0 + \frac{\epsilon^3}{3!} y^9 x_1 + \frac{\epsilon^3}{3!} x$	$y^6 x_0$	nonvanishing
$\phi^{(7)} = \frac{\epsilon^2}{2!} y^9 x_1 + \frac{\epsilon^2}{2!} x$	0	$O(\epsilon^2)$
$\phi^{(8)} = \epsilon y^9 x_1 + \epsilon x$	0	$O(\epsilon^1)$
$\phi^{(9)} = y^9 x_1 + x$	$y^9 x_1 + x$	nonvanishing

We start differentiating Φ_ϵ with respect to ϵ. The first-order derivative is:

$$\Phi'_\epsilon = \frac{d}{d\epsilon} F(\phi(x,y)) = dF(\phi)[\phi']$$

The second-order derivative is:

$$\begin{aligned}\Phi''_\epsilon &= d^2F(\phi)[\phi', \phi'] + dF(\phi)[\phi''] \\ &= d^2F[\phi'^2] + dF[\phi''].\end{aligned}$$

The third-order derivative is (for brevity, we omit the point at which the d^jF are evaluated):

$$\begin{aligned}\Phi'''_\epsilon &= d^3F[\phi'^3] + d^2F[2\phi', \phi''] + d^2F''[\phi', \phi''] + dF[\phi'''] \\ &= d^3F[\phi'^3] + 3d^2F[\phi', \phi''] + dF[\phi'''].\end{aligned}$$

Looking at the table, Φ'_ϵ is $O(\epsilon^2)$ and Φ''_ϵ is $O(\epsilon)$ so

$$\Phi'_\epsilon\Big|_{\epsilon=0} = \Phi''_\epsilon\Big|_{\epsilon=0} = 0.$$

Moreover, evaluated in $\epsilon = 0$ (4.1) reduces to

$$\Phi'''_\epsilon\Big|_{\epsilon=0} = dF(0)[y^3 v] = y^3 dF(0)[v] = 0$$

since $v \in \operatorname{Ker} dF(0)$ by hypothesis.

Let us proceed computing higher-order derivatives. The fourth-order derivative with respect to ϵ is

$$\begin{aligned}\Phi^{(4)}_\epsilon =& d^4F[\phi'^4] + d^3F[3\phi'^2, \phi''] + 3d^3F[\phi'^2, \phi''] + 3d^2F[\phi''^2] \\ &+ 3d^2F[\phi', \phi'''] + d^2F[\phi', \phi'''] + dF[\phi^{(4)}] \\ =& d^4F[\phi'^4] + 6d^3F[\phi'^2, \phi''] + 4d^2F[\phi', \phi'''] \\ &+ 3d^2F[\phi''^2] + dF[\phi^{(4)}].\end{aligned}$$

The fifth-order derivative is

$$\begin{aligned}\Phi_\epsilon^{(5)} =& d^5F[\phi'^5] + 10d^4F[\phi'^3, \phi''] + 15d^3F[\phi', \phi''^2] \\ &+ 10d^3F[\phi'^2, \phi'''] + 5d^2F[\phi', \phi^{(4)}] + 10d^2F[\phi'', \phi'''] \\ &+ dF[\phi^{(5)}] \\ =& 15d^3F[\phi', \phi''^2] + 10d^3F[\phi'^2, \phi'''] + 5d^2F[\phi', \phi^{(4)}] \\ &+ 10d^2F[\phi'', \phi'''] + dF[\phi^{(5)}] + O(\epsilon^5).\end{aligned}$$

The sixth-order derivative is

$$\begin{aligned}\Phi_\epsilon^{(6)} =& 15d^4F[\phi'^2, \phi''^2] + 15d^3F[\phi''^3] + 15d^3F[\phi', 2\phi'', \phi'''] \\ &+ 10d^4F[\phi'^3, \phi'''] + 10d^3F[2\phi', \phi'', \phi'''] + 10d^3F[2\phi'^2, \phi^{(4)}] \\ &+ 5d^3F[\phi'^2, \phi^{(4)}] + 5d^2F[\phi'', \phi^{(4)}] + 5d^2F[\phi', \phi^{(5)}] \\ &+ 10d^3F[\phi', \phi'', \phi'''] + 10d^2F[\phi'''^2] + 10d^2F[\phi'', \phi^{(4)}] \\ &+ d^2F[\phi', \phi^{(5)}] + dF[\phi^{(6)}] + O(\epsilon^4) \\ =& 15d^3F[\phi'''^3] + 60d^3F[\phi', \phi'', \phi'''] + 6d^2F[\phi', \phi^{(5)}] \\ &+ 15d^2F[\phi'', \phi^{(4)}] + 10d^2F[\phi'''^2] + dF[\phi^{(6)}] + O(\epsilon^4).\end{aligned}$$

As before, $\Phi_\epsilon^{(4)}$ is $O(\epsilon^2)$ and $\Phi_\epsilon^{(5)}$ is $O(\epsilon)$ so they vanish when evaluated at $\epsilon = 0$, while when evaluating $\Phi_\epsilon^{(6)}$ at $\epsilon = 0$ we are left with

$$\begin{aligned}\Phi_\epsilon^{(6)}\Big|_{\epsilon=0} =& 10d^2F(0)[y^3v, y^3v] + dF(0)[y^6x_0] \\ =& 10\,\mathrm{Hess}\,F[y^3v] + dF(0)[y^6x_0] \\ =& y^6\Big(10\,\mathrm{Hess}\,F(0)[v] + dF(0)[x_0]\Big).\end{aligned}$$

since by hypotesis $v \in \mathrm{Iso}\,H_F(0)$, $\mathrm{Hess}\,F(0)[v]$ is in $\mathrm{Im}\,dF(0)$, we choose x_0 in the definition of ϕ_ϵ (4.1) so that $10\,\mathrm{Hess}\,F(0)[v] + dF(0)[x_0] = 0$. In this way $\Phi_\epsilon^{(6)}\Big|_{\epsilon=0} = 0$.

The seventh-order derivative of Φ_ϵ with respect to ϵ is

$$\begin{aligned}\Phi_\epsilon^{(7)} =& 15d^4F[\phi', \phi''^3] + 15d^3F[3\phi''^2, \phi'''] + 60d^4F[\phi'^2, \phi'', \phi'''] \\ &+ 60d^3F[\phi''^2, \phi'''] + 60d^3F[\phi', \phi'''^2] + 60d^3F[\phi', \phi'', \phi^{(4)}] \\ &+ 6d^3F[\phi'^2, \phi^{(5)}] + 6d^2F[\phi'', \phi^{(5)}] + 6d^2F[\phi', \phi^{(6)}] \\ &+ 15d^3F[\phi', \phi'', \phi^{(4)}] + 15d^2F[\phi''', \phi^{(4)}] + 15d^2F[\phi'', \phi^{(5)}] \\ &+ 10d^3F[\phi', \phi'''^2] + 10d^2F[2\phi''', \phi^{(4)}] + d^2F[\phi', \phi^{(6)}] \\ &+ dF[\phi^{(7)}] + O(\epsilon^3) \\ =& 105d^3F[\phi''^2, \phi'''] + 70d^3[\phi', \phi'''^2] + 21d^2F[\phi'', \phi^{(5)}] \\ &+ 7d^2F[\phi', \phi^{(6)}] + 35d^2F[\phi''', \phi^{(4)}] + dF[\phi^{(7)}] + O(\epsilon^3).\end{aligned} \tag{4.3}$$

At this point it is clear that $\Phi_\epsilon^{(7)}$ is $O(\epsilon^2)$ and $\Phi_\epsilon^{(8)}$ is $O(\epsilon)$, so it will be $\Phi_\epsilon^{(7)}\Big|_{\epsilon=0} = \Phi_\epsilon^{(8)}\Big|_{\epsilon=0} = 0$. In order to compute $\Phi_\epsilon^{(9)}$, we differentiate twice the right-hand side of (4.3) keeping in mind that we are interested in an expression modulo $O(\epsilon)$. In particular, we not interested in terms whose arguments involve different elements than $\phi''', \phi^{(6)}$ and $\phi^{(9)}$. We get

$$\begin{aligned}\Phi_\epsilon^{(9)} =&105 d^3F[2\phi'''^3] + 70 d^3F[\phi'''^3] + 35 d^2F[\phi''', \phi^{(6)}] \\ &+ 42 d^2F[\phi''', \phi^{(6)}] + 7 d^2F[\phi''', \phi^{(6)}] + dF[\phi^{(9)}] + O(\epsilon) \\ =&280 d^3F[\phi'''^3] + 84 d^2F[\phi''', \phi^{(6)}] + dF[\phi^{(9)}] + O(\epsilon)\end{aligned}$$

and so

$$\begin{aligned}\Phi_\epsilon^{(9)}\Big|_{\epsilon=0} =&280 d^3F[(y^3v)] + 84 d^2F[y^3v, y^6x_0] + dF[y^9x_1 + x] \\ =&y^9\Big(280 d^3F(0)[v^3] + 84 d^2F(0)[v, x_0] + dF(0)[x_1]\Big) + dF(0)[x].\end{aligned}$$

We decompose as a direct sum

$$d^3F(0)[v^3] = \pi_{\mathrm{Coker}\, dF(0)}(d^3F(0)[v^3]) + \pi_{\mathrm{Im}\, dF(0)}(d^3F(0)[v^3])$$

and claim that it is possible to find a vector x_1 in $\mathbb{R}^N$ such that

$$\pi_{\mathrm{Im}\, dF(0)}(d^3F(0)[v^3]) + 84 d^2F(0)[v, x_0] + dF(0)[x_1] = 0.$$

In fact, this is equivalent to prove that $d^2F(0)[v, x_0]$ is in $\mathrm{Im}\, dF(0)$.

We point out the following fact: given a real-valued, positive-semidefinite bilinear operator $B[\cdot,\cdot]$, and a vector u such that $B[u,u] = 0$, it follows that $B[u,\cdot]$ is zero as a linear operator. This can be seen, for example, by diagonalizing the matrix asssociated with B and writing u as a linear combination of the associated eigenvectors.

Since by hypotesis $H_F(0)$ is positive-semidefinite and $v \in \mathrm{Iso}\, H_F(0)$ we can apply this fact to $\pi_{\mathrm{Coker}\, dF(0)} d^2F(0)[\cdot,\cdot]$: we conclude that $\pi_{\mathrm{Coker}\, dF(0)} d^2F(0)[v,\cdot]$ is zero as a linear operator, so, in particular $\pi_{\mathrm{Coker}\, dF(0)} d^2F(0)[v, x_0]$ is zero, and this means exactly that $d^2F(0)[v, x_0]$ is $\mathrm{Im}\, dF(0)$.

Now, x_1 is fixed and ϕ_ϵ completely determined. The computations made above and the choices of x_0 and x_1 in the definition of ϕ (4.1) allowed us to conclude that the derivatives with respect to ϵ of Φ_ϵ from first to eigth order vanish. As a consequence of the Taylor formula, we can write

$$\Phi_\epsilon(x,y) = \frac{\epsilon^9}{9!}\Phi_\epsilon^{(9)}(x,y)\Big|_{\epsilon=0} + R_\epsilon(x,y). \tag{4.4}$$

where the reminder $R_\epsilon(x,y)$ is $O(\epsilon^{10})$ as ϵ goes to 0.

So

$$\frac{1}{\epsilon^9}\Phi_\epsilon(x,y) = \frac{1}{9!}\left(280 y^9 \pi_{\mathrm{Coker}\, dF(0)}\left(\frac{\partial^3 F}{\partial v^3}(0)\right) + dF(0)[x]\right) + \frac{1}{\epsilon^9}R_\epsilon(x,y).$$

We now set

$$\Psi_\epsilon(x,y) = \frac{1}{\epsilon^9}\Phi_\epsilon(x, y^{1/9}).$$

Since Ψ_ϵ is obtained composing $\frac{1}{\epsilon^9}\Phi_\epsilon$ with a homeorphism, proving that Ψ_ϵ is open at 0 will also prove that Φ_ϵ is open at 0.

Claim. $R_\epsilon(x, y^{1/9})$ is $o(|x| + |y|)$ as $(x,y) \to 0$, equivalently

$$\lim_{(x,y)\to 0} \frac{R_\epsilon(x,y)}{|x| + |y|^9} = 0. \tag{4.5}$$

If this claim holds true, we would have, as $(x,y)\to 0$,

$$\Psi_\epsilon(x,y) = AdF(0)[x] + By\pi_{\mathrm{Coker}\,dF(0)}\left(\frac{\partial^3 F}{\partial v^3}(0)\right) + R_\epsilon(x,y^{1/9}).$$

with $A = 1/9!$, $B = 280/9!$ and $R_\epsilon(x,y^{1/9})$ that is $o(|x|+|y|)$. In particular, $d\Psi_\epsilon$ exists and has full rank. Thus, after a linear change of coordinates we are reduced to prove that

$$\Psi_\epsilon(x,y) = (x,y) + R_\epsilon(x,y^{1/9})$$

is open at 0. Now, the openness at 0 can be characterized by the fact that that we can find a sufficiently small $\delta_0 > 0$ such that, for any $0 < \delta < \delta_0$, $\Psi_\epsilon(B_{\mathbb{R}^n}(0,\delta])$ contains an ball centered at 0 of some radius, say $\eta = \eta(\delta)$. This can be seen using the classical Brouwer's Fixed Point Theorem (see e.g. Theorem 6.1 in Chapter Two of [**10**]) in the following way:

- choose δ_0 such that $|R_\epsilon(x,y^{1/9})| \le \frac{\delta}{2}(|x|+|y|)$ for $|x|+|y| \le \delta_0$, which is possible due to the Claim
- set $\eta = \delta/2$
- fix an arbitrary ξ in $B_{\mathbb{R}^n}(0,\delta]$ and define the map

$$\chi(x,y) = -\Psi_\epsilon(x,y) + \xi + (x,y).$$

The problem becomes equivalent to verify that χ has a fixed point in $B_{\mathbb{R}^n}(0,\delta]$, a condition that is ensured by Browuer's Theorem if prove that χ maps $B_{\mathbb{R}^n}(0,\delta]$ into itself. Our choices of δ_0 and η allow us to do so since

$$\begin{aligned}|\chi(x,y)| =& |-\Psi_\epsilon(x,y) + \xi + (x,y)|\\ =& |R_\epsilon(x,y^{1/9}) + \xi|\\ \le& |R_\epsilon(x,y^{1/9})| + |\xi| \le \delta/2 + \delta/2 = \delta.\end{aligned}$$

We are just left to prove the Claim (4.5). By definition of $R_\epsilon(x,y)$ it is

$$R_\epsilon(x,y) = \Phi_\epsilon(x,y) - \frac{280\epsilon^9}{9!}\pi_{\mathrm{Coker}\,dF(0)}\left(\frac{\partial^3 F}{\partial v^3}(0)\right) - \frac{1}{9!}dF(0)[x]. \tag{4.6}$$

Exactly as before we have

$$\begin{aligned}
&\left.\frac{\partial\phi}{\partial\partial y}\right|_{y=0} = 0, && \left.\frac{\partial^2\phi}{\partial y^2}\right|_{y=0} = 0, && \left.\frac{\partial^3\phi}{\partial y^3}\right|_{y=0} = \epsilon^3 v,\\
&\left.\frac{\partial^4\phi}{\partial y^4}\right|_{y=0} = 0, && \left.\frac{\partial^5\phi}{\partial y^5}\right|_{y=0} = 0, && \left.\frac{\partial^6\phi}{\partial y^6}\right|_{y=0} = \epsilon^6 x_0,\\
&\left.\frac{\partial^7\phi}{\partial y^7}\right|_{y=0} = 0, && \left.\frac{\partial^8\phi}{\partial y^8}\right|_{y=0} = 0, && \left.\frac{\partial^9\phi}{\partial y^9}\right|_{y=0} = \epsilon^9 x_1.
\end{aligned}$$

Again, the same computations made above lead to

$$\frac{\partial^k\Phi_\epsilon}{\partial y^k}(0) = 0 \quad \text{for } j = 1,...,8 \tag{4.7}$$

and

$$\frac{\partial^9\Phi_\epsilon}{\partial y^9}(0) = 280\epsilon^9\pi_{\mathrm{Coker}\,dF(0)}\left(\frac{\partial^3 F}{\partial v^3}(0)\right). \tag{4.8}$$

On the other hand,by the chain rule

$$D_x \Phi_\epsilon(x,y)\Big|_{(x,y)=0} = \frac{\epsilon^9}{9!} D_x F(0). \tag{4.9}$$

Relations (4.7)-(4.9) force the Taylor formula for Φ_ϵ at 0 to be

$$\Phi_\epsilon(x,y) = d_x F(0)[x] + O(|x||y|) + \frac{y^9}{9!}\frac{\partial^9 \Phi_\epsilon}{\partial y^9}(0) + o(|x| + |y|^9) \tag{4.10}$$

where the big-oh $O(|x||y|)$ sums up the mixed terms in x and y until the ninth-order. But as $O(|x||y|)$ is $o(|x|+|y|^9)$, Formula (4.10) compared with (4.6) allows us to conclude that $R_\epsilon(x,y)$ is $o(|x| + |y^9)$ as $(x,y) \to 0$, which is exactly the limiting relation (4.5). The claim is proved, and so is the Theorem. □

Remark 4.2. Let us point out a fact that we will use in a moment. What we have proved in Theorem 4.1 is that, having set $V = \mathrm{span}_{\mathbb{R}}\{v\}$, the (restriction of the) map F:

$$\mathrm{Coim}\, dF(0) \oplus V \xrightarrow{F} \mathbb{R}^n = \mathrm{Im}\, dF(0) \oplus \mathrm{span}_{\mathbb{R}}\{\pi\left(\frac{\partial^3 F}{v^3}(0)\right)\}$$

is open in a neighborhood of $(0,0)$. As V is contained in $\mathrm{Ker}\, dF(0)$ (since v is an isotropic vector for the restricted Hessian), it is necessarily mapped into $\mathrm{Coker}\, dF(0)$, which is spanned by $\pi_{\mathrm{Coker}\, dF(0)}\left(\frac{\partial^3 F}{v^3}(0)\right)$.

By the first-order Open Mapping Theorem, we know that the restriction of F to $\mathrm{Coim}\, dF(0)$

$$\mathrm{Coim}\, dF(0) \xrightarrow{F} \mathrm{Im}\, dF(0)$$

is open at 0. The key point was to prove that also the restriction of F to V

$$V \xrightarrow{F} \mathrm{span}_{\mathbb{R}}\{\pi\left(\frac{\partial^3 F}{v^3}(0)\right)\}$$

is open at 0.

The generalization to the arbitrary corank case is now straightforward.

Theorem 4.3. *Let n, N be positive integers with $N \geq n$, and let $F : \mathbb{R}^N \to \mathbb{R}^n$ be a smooth map with $F(0) = 0$ and let r be the corank of $dF(0)$. Suppose that the restricted Hessian of F at 0*

$$H_F(0) : \mathrm{Ker}\, dF(0) \to \mathrm{Coker}\, dF(0)$$

is semidefinite and that the isotropic space of H_F at 0

$$\mathrm{Iso}\, H_F(0) = \{w \in \mathrm{Ker}\, dF(0) : H_F(0)[w] = 0\}$$

is nontrivial. If there exist r vectors $v_1, \dots, v_r$ in $\mathrm{Iso}\, H_F(0)$ *such that the projected third-order derivatives of F over* $\mathrm{Coker}\, dF(0)$ *along each v_j*

$$\pi_{\mathrm{Coker}\, dF(0)}\left(\frac{\partial^3 F}{v_1^3}(0)\right), \pi_{\mathrm{Coker}\, dF(0)}\left(\frac{\partial^3 F}{v_2^3}(0)\right), \dots, \pi_{\mathrm{Coker}\, dF(0)}\left(\frac{\partial^3 F}{v_r^3}(0)\right)$$

are linearly independent vectors, then F is open at 0.

Proof. We proceed by finite induction over r.

The case $r = 1$ is Theorem 4.1. If the thesis holds true for $r-1$, we set, for $j = 1, \dots, r$, $V_j = \operatorname{span}_{\mathbb{R}}\{v_j\}$, $W_j = \operatorname{span}_{\mathbb{R}}\{\pi_{\operatorname{Coker} dF(0)}\left(\frac{\partial^3 F}{v_j^3}(0)\right)\}$ and we consider

$$\operatorname{Coim} dF(0) \oplus V_1 \oplus \cdots \oplus V_r \xrightarrow{F} \operatorname{Im} dF(0) \oplus W_1 \oplus \cdots \oplus W_r.$$

By inductive hypotesis F is open at 0 from $\operatorname{Coim} dF(0) \oplus V_1 \oplus \cdots \oplus V_{r-1}$ to (necessarily) $\operatorname{Im} dF(0) \oplus W_1 \oplus \cdots \oplus W_{r-1}$, since each V_j is mapped into W_j, see Remark 4.2. But then we are as in the case of corank 1 and, applying Theorem 4.1 again we conclude that F is open at 0. □

Remark 4.4. If r vectors $v_1, \dots, v_m$ as in the above Theorem exist, they need to be linearly independent since the map $v \mapsto \pi_{\operatorname{Coker} dF(0)}\left(\frac{\partial^3 F}{v^3}(0)\right)$ is smooth. This implies that a necessary condition for the openness of F at 0 is that $\operatorname{Iso} H_F(0)$ must have dimension at least r as a vector space.

The Theorem we are looking for now descends easily from Theorem 4.3.

Theorem 4.5 (A Third-Order Open Mapping Theorem). *Let X be a Banach space, x_0 a point in X, $Y \approx \mathbb{R}^n$ be a finite dimensional normed space of dimension n and let $F : X \to Y$ be a smooth map. Suppose that $dF(x_0)$ has corank r, that the restricted Hessian of F at x_0*

$$H_F(0) : \operatorname{Ker} dF(x_0) \to \operatorname{Coker} dF(x_0)$$

is semidefinite and that the isotropic space of H_F at x_0

$$\operatorname{Iso} H_F(x_0) = \{w \in \operatorname{Ker} dF(x_0) : H_F(x_0)[w] = 0\}$$

is nontrivial. If there exist r vectors $v_1, \dots, v_r$ in $\operatorname{Iso} H_F(x_0)$ such that the projected third-order derivatives of F over $\operatorname{Coker} dF(x_0)$ along each v_j

$$\pi_{\operatorname{Coker} dF(x_0)}\left(\frac{\partial^3 F}{v_1^3}(x_0)\right), \pi_{\operatorname{Coker} dF(x_0)}\left(\frac{\partial^3 F}{v_2^3}(x_0)\right), \dots, \pi_{\operatorname{Coker} dF(x_0)}\left(\frac{\partial^3 F}{v_r^3}(x_0)\right)$$

are linearly independent vectors, then F is open at x_0.

Proof. We suppose $x_0 = 0$. If this is not the case, we just compose F with a translation that, being a homeomorphism, does not affect the openness of F at x_0.

Set, for $j = 1, \dots, r$, $V_j = \operatorname{span}_{\mathbb{R}}\{v_j\}$ and consider the restriction of F

$$\operatorname{Coim} dF(0) \oplus V_1 \oplus \cdots \oplus V_r \xrightarrow{F} \operatorname{Im} dF(0) \oplus W_1 \oplus \cdots \oplus W_r.$$

It is sufficient to prove that this restriction is open at 0 (the zero vector in the finite dimensional space) because then a fortiori F will be open at 0 in X. Since the range space is finite dimensional, $\operatorname{Coim} dF(0)$ is finite dimensional, and so one may suppose that X itself is finite dimensional. But then the result is proved by Theorem 4.3. □

4.2. Third-Order Derivatives of the End-Point Mapping

Theorem 4.5 fully applies to our situation and provides necessary conditions for length-minimizing curves to satisfy, or equivalently sufficient conditions for a curve to be nonminimizing. Namely (we use the notation and definitions introduced in Sections 3.1 and 3.2 of Chapter 3):

. Let $\gamma : [0,1] \to \mathbb{R}^n$ be a $\mathcal{D}$-horizontal curve with extremal points x_0 and x_1 and control h. If $d\hat{\mathcal{E}}(h)$ has corank r and the restricted Hessian $H_{\hat{\mathcal{E}}}(h)$ is semidefinite, if we can find r vectors $v_1, \ldots, v_r$ in $\operatorname{Iso} H_{\hat{\mathcal{E}}}(h)$ such that

$$\pi_{\operatorname{Coker} d\hat{\mathcal{E}}(h)}\left(\frac{\partial^3 \hat{\mathcal{E}}}{v_1^3}(h)\right), \pi_{\operatorname{Coker} d\hat{\mathcal{E}}(h)}\left(\frac{\partial^3 \hat{\mathcal{E}}}{v_2^3}(h)\right), \ldots, \pi_{\operatorname{Coker} d\hat{\mathcal{E}}(h)}\left(\frac{\partial^3 \hat{\mathcal{E}}}{v_r^3}(h)\right)$$

are linearly independent, then $\hat{\mathcal{E}}$ is open at h, and thus γ is not a length-minimizing curve.

We want to demonstrate the applicability of our result by proving the nonminimality of the abnormal extremal γ, discussed in Section 2.4, Chapter 2 when m is 3. Recall that since γ is not a regular abnormal extremal, according to Definition 2.1.

First of all, we deduce handful formula for the third-order derivative of $\hat{\mathcal{E}}$ at h. We resume from (3.15). Differentiating with respect to s once more, we obtain

$$\frac{\partial^3}{\partial s^3}\hat{\mathcal{E}}_t(h+sv) = 3\int_0^t \frac{\partial^2}{\partial s^2}\Psi_{\tau,v}(\hat{\mathcal{E}}_\tau(h+sv))d\tau + s\int_0^t \frac{\partial^3}{\partial s^3}\Psi_{\tau,v}(\hat{\mathcal{E}}_\tau(h+sv))d\tau$$

so

$$\frac{\partial^3}{\partial s^3}\hat{\mathcal{E}}(h+sv)\bigg|_{s=0} = 3\int_0^1 \frac{\partial^2}{\partial s^2}\Psi_{\tau,v}(\hat{\mathcal{E}}_\tau(h+sv))\bigg|_{s=0} d\tau =$$

$$3\int_0^1 d_x^2\Psi_{\tau,v}(x_0)\left[\left(\frac{\partial}{\partial s}\hat{\mathcal{E}}_\tau(h+sv)\bigg|_{s=0}\right)^2\right] + d_x\Psi_{\tau,v}(x_0)\left[\frac{\partial^2}{\partial s^2}\hat{\mathcal{E}}_\tau(h+sv)\bigg|_{s=0}\right]d\tau.$$

We are interested in the case $v \in \operatorname{Iso} H_{\hat{\mathcal{E}}}(h)$, in particular v is in $\operatorname{Ker} d\hat{\mathcal{E}}(h)$, so condition (3.17) holds. As a consequence, we can perform the following computation on the first term:

$$\int_0^1 d_x^2\Psi_{\tau_1,v}(x_0)\left[\frac{\partial}{\partial s}\hat{\mathcal{E}}_{\tau_1}(h+sv)\bigg|_{s=0}, \frac{\partial}{\partial s}\hat{\mathcal{E}}_{\tau_1}(h+sv)\bigg|_{s=0}\right]d\tau_1 =$$

$$\int_0^1 d_x^2\Psi_{\tau_1,v}(x_0)\left[\int_0^{\tau_1}\Psi_{\tau_2,v}(x_0)d\tau_2, \int_0^{\tau_1}\Psi_{\tau_3,v}(x_0)d\tau_3\right]d\tau_1 =$$

$$-\int_0^1 d_x^2\Psi_{\tau_1,v}(x_0)\left[\int_{\tau_1}^1\Psi_{\tau_2,v}(x_0)d\tau_2, \int_0^{\tau_1}\Psi_{\tau_3,v}(x_0)d\tau_3\right]d\tau_1 =$$

$$-\int_0^1\int_{\tau_1}^1 d_x^2\Psi_{\tau_1,v}(x_0)\left[\Psi_{\tau_2,v}(x_0), \int_0^{\tau_1}\Psi_{\tau_3,v}(x_0)d\tau_3\right]d\tau_2 d\tau_1 =$$

$$-\int_0^1\int_0^{\tau} d_x^2\Psi_{\tau_1,v}(x_0)\left[\Psi_{\tau_2,v}(x_0), \int_0^{\tau_1}\Psi_{\tau_3,v}(x_0)d\tau_3\right]d\tau_1 d\tau_2 =$$

rename the variables t and τ to obtain

$$-\int_0^1\int_0^{\tau_1}\int_0^{\tau_2} d_x^2\Psi_{\tau_2,v}(x_0)\left[\Psi_{\tau_1,v}(x_0), \Psi_{\tau_3,v}(x_0)\right]d\tau_3 d\tau_2 d\tau_1.$$

As for the second term, we just use (3.16):

$$\int_0^1 d_x\Psi_{t,v}(x_0)\left[\frac{\partial^2}{\partial s^2}\hat{\mathcal{E}}_t(h+sv)\Big|_{s=0}\right]dt =$$
$$\int_0^1 d_x\Psi_{t,v}(x_0)\left[2\int_0^t\int_0^\tau d_x\Psi_{\tau,v}(x_0)[\Psi_{\sigma,v}(x_0)]d\sigma d\tau\right]dt$$
$$2\int_0^1\int_0^{\tau_1}\int_0^{\tau_2} d_x\Psi_{\tau_1,v}(x_0)[d_x\Psi_{\tau_2,v}(x_0)[\Psi_{\tau_3,v}(x_0)]]d\tau_3 d\tau_2 d\tau_1.$$

So, for any v in $\operatorname{Ker} d\hat{\mathcal{E}}(h)$ we have

$$\frac{\partial^3\hat{\mathcal{E}}}{\partial v^3}(h) = 3\int_0^1\int_0^{\tau_1}\int_0^{\tau_2} - d_x^2\Psi_{\tau_2,v}(x_0)\left[\Psi_{\tau_1,v}(x_0),\Psi_{\tau_3,v}(x_0)\right] + 2d_x\Psi_{\tau_1,v}(x_0)[d_x\Psi_{\tau_2,v}(x_0)[\Psi_{\tau_3,v}(x_0)]]d\tau_3 d\tau_2 d\tau_1.$$

We turn to our specific extremal. Recall that the curve is $\gamma(t) = (0,t,0)$ with control $h = (0,1)$ for the distribution spanned in $\mathbb{R}^3$ by the vector fields defined by (2.10). We follow the scheme outlined in Subsection 3.2.1 of Chapter 3.

Step 1: computation of the flow $\Phi_{t,v}$. We compute the vector field $\Phi_t = (\Phi_t^1,\Phi_t^2,\Phi_t^3)$ solution of the Cauchy problem

$$\begin{cases}\dot{\Phi}_t(x) = h_1X_1(\Phi_t(x)) + h_2X_2(\Phi_t(x))\\ \Phi_0(x) = x\end{cases}$$

that is

$$\begin{cases}\dot{\Phi}_t^1(x) = 0\\ \dot{\Phi}_t^2(x) = 1-\Phi_t^1\\ \dot{\Phi}_t^3(x) = (\Phi_t^1(x))^3\\ (\Phi_0^1(x),\Phi_0^2(x),\Phi_0^3(x)) = (x_1,x_2,x_3).\end{cases}$$

The integration of this system is elementary and the solution is

$$\Phi_t(x) = (x_1, x_2 + t(1-x_1), x_3 + tx_1^3).$$

Step 2: computation of $\Psi_{t,v}$ and $d\hat{\mathcal{E}}(h)$. We compute $\Psi_{t,v}$ defined in formula (3.14). The Jacobian matrix with respect to the space variable of Φ_t is

$$D_x\Phi_t(x) = \begin{pmatrix}1 & 0 & 0\\ -t & 1 & 0\\ 3tx_1^2 & 0 & 1\end{pmatrix} \tag{4.11}$$

whose inverse is

$$D_x\Phi_t(x)^{-1} = \begin{pmatrix}1 & 0 & 0\\ t & 1 & 0\\ -3tx_1^2 & 0 & 1\end{pmatrix}$$

so

$$\Psi_{t,v}(x) = D_x\Phi_t(x)^{-1}[v(t)\cdot X(\Phi_t(x))]$$
$$= \begin{pmatrix}1 & 0 & 0\\ t & 1 & 0\\ -3tx_1^2 & 0 & 1\end{pmatrix}\begin{pmatrix}v_1(t)\\ v_2(t)(1-x_1)\\ v_2(t)(x_1^3)\end{pmatrix} = \begin{pmatrix}v_1(t)\\ tv_1(t)+v_2(t)(1-x_1)\\ -3tv_1(t)x_1^2 + v_2(t)x_1^3\end{pmatrix}.$$

The initial point is $x_0 = 0$, so

$$d\hat{\mathcal{E}}(h)[v] = \int_0^1 \Psi_{\tau,v}(0)d\tau = \left(\int_0^1 v_1(\tau)d\tau, \int_0^1 \tau v_1(\tau) + v_2(\tau)d\tau, 0\right). \tag{4.12}$$

From this expression it is clear that $\operatorname{Coker} d\hat{\mathcal{E}}(h) \simeq (\operatorname{Im} d\hat{\mathcal{E}}(h))^\perp = \operatorname{span}_{\mathbb{R}}\{\lambda\}$ with $\lambda = (0, 0, 1)$, in particular the corank of this extremal is one.

Step 3: Computation of the restricted Hessian $H_{\hat{\mathcal{E}}}(h)$. We want to determine the restricted hessian of $\hat{\mathcal{E}}$ at h given by

$$H_{\hat{\mathcal{E}}}(h)[v] = \left\langle \lambda, \operatorname{Hess} \hat{\mathcal{E}}(h)[v] \right\rangle$$

for v in $\operatorname{Ker} d\hat{\mathcal{E}}(h)$. The formula for $\operatorname{Hess} \hat{\mathcal{E}}(h)$ is (3.18). Recalling formula (0.2) for the commutator, we compute

$$D_x\Psi_{\tau_1,v}(0)[\Psi_{\tau_2,v}(0)] = \begin{pmatrix} 0 & 0 & 0 \\ -v_2(\tau_1) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_1(\tau_2) \\ \tau_2 v_1(\tau_2) + v_2(\tau_2) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -v_2(\tau_1)v_1(\tau_2) \\ 0 \end{pmatrix}.$$

A similar result is obtained for $D_x\Psi_{\tau_2,v}(0)[\Psi_{\tau_1,v}(0)]$ exchanging the roles of τ_2 and τ_1. In any case, it is clear that the x_3 component of the Hessian is zero, and thus $H_{\hat{\mathcal{E}}}(h)[v] = \left\langle \lambda, \operatorname{Hess} \hat{\mathcal{E}}(h)[v] \right\rangle = 0$. In particular this trivial restricted Hessian is semidefinite.

Step 4: third-order derivatives of $\hat{\mathcal{E}}$ at h. We compute now the third-order derivatives of $\hat{\mathcal{E}}$ at h using (4.2). Looking at the Jacobian matrix of $\Psi_{t,v}$ (4.11), it is clear that

$$\frac{\partial^2 \Psi^k_{t,v}}{\partial x_i x_j}(0) = -6tv_1(t) \quad \text{when } k = 3 \text{ and } i, j = 1$$

and zero for any other i, j, k. Moreover, the second term in the integral (4.2) has, as before, no x_3 component, so we conclude that

$$\begin{aligned} \pi_{\operatorname{Coker} d\hat{\mathcal{E}}(h)}\left(d^3\hat{\mathcal{E}}(h)[v^3]\right) &= \int_0^1 \int_0^{\tau_1} \int_0^{\tau_2} \frac{\partial^2 \Psi^3_{\tau_2,v}}{\partial x_1^2}(0)\Psi^1_{\tau_1,v}(0)\Psi^1_{\tau_3,v}(0)d\tau_3 d\tau_2 d\tau_1 \\ &= \int_0^1 \int_0^{\tau_1} \int_0^{\tau_2} \left(-6\tau_2 v_1(\tau_2)v_1(\tau_1)v_1(\tau_3)\right) d\tau_3 d\tau_2 d\tau_1. \end{aligned}$$

To conclude, we just need to find a control $v = (v_1, v_2)$ in $\operatorname{Ker} d\hat{\mathcal{E}}(h)$ (i.e. such that (4.12) vanishes) such that this last expression is not zero. For example, we may take $v_1(t) = \sin(2\pi t)$ or $v_2(t) = \operatorname{sgn}(t)$ and $v_2(t) = -tv_1(t)$.

We have proven through Theorem 4.5 that $\hat{\mathcal{E}}$ is open at h, and consequently that the abnormal extremal γ found in Section 2.4 is not length-minimizing for the distribution when $m = 3$.

Remark 4.6. According to our knowledge on the subject, we do not know any alternative way of proving this result. We also point out that, when the integer m in the definition of $\mathcal{D}$ is greater than 3 and odd, the third-order derivatives of $\hat{\mathcal{E}}$ at h vanish, as can be easily deduced, and we cannot draw any conclusion using Theorem 4.5.

APPENDIX A

Statement of the Pontryagin Maximum Principle

In this appendix we state precisely the Pontryagin Maximum Principle (PMP) along with the necessary definitions. A reference book is, for example, [**6**].

A.0.1. The minimization problem. Approximating cones. We start with the following minimization problem:

$$\text{Minimize } I(x,h) = g(x(b)) + \int_a^b L(t,x(t),u(t))dt \tag{A.1}$$

$$\text{with } \begin{cases} \dot{x}(t) & = f(t,x(t),u(t)) \\ x(a) & = \bar{x} \\ x(b) & \in \mathcal{S} \\ h(\cdot) & \in \mathcal{U} \end{cases}$$

where:

- $\mathcal{S}$ is a subset of $\mathbb{R}^n$, called *target*
- $\mathcal{U}$ is a family of curves, called *controls* from $[a,b]$ to a set U, the *set of control values*, that can be of many kinds, e.g. a subset of $\mathbb{R}^n$, $\{-1,1\}$ or a subset of some infinite dimensional vector space
- $f : [a,b] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^n$ and $L : [a,b] \times \mathbb{R}^n \times U \to \mathbb{R}^n$ are maps called, respectively, the *dynamics* and the *Lagrangian* of the system

The *associated Hamiltonian* is $\mathcal{H} : [a,b] \times (\mathbb{R}^{n+1})^* \times \mathbb{R}^n \times U \to \mathbb{R}$ defined by

$$\mathcal{H}(t,\lambda_0,\lambda,x,u) = \langle \lambda, f(t,x,u) \rangle - \lambda_0 L(t,x,u). \tag{A.2}$$

We are then given a *reference control* u_* and the following further conditions:

- $\mathcal{U}$ is a *needle variational neighborhood* of u_*, in the following sense: given any elements $u_1,\ldots,u_k$ of U, there exists a sufficiently small $\alpha > 0$ such that, if $I_1,\ldots,I_k$ are pairwise disjoint closed subintervals of $[a,b]$ with $\sum_{i=1}^k |I_k| < \alpha$, then

$$u_* + \sum_{i=1}^k \chi_{I_i}(u_i - u_*) \in \mathcal{U} \tag{A.3}$$

- for any control $u(\cdot)$ which is constant or equal to $u_*(\cdot)$ the vector field $G : A = [a,b] \times \mathbb{R}^n \to \mathbb{R}^{n+1}$ given by

$$G(t,x) = (L(t,x,u(t)), f(t,x,u(t))) \tag{A.4}$$

  satisfies the C^1-Caratheodory conditions:
  - **(A1):** for any $x \in \mathbb{R}^n$ such that the x-section $A_x = \{t \in [a,b] : (t,x) \in A\}$ is not empty, the x-section of G, $t \mapsto G(t,x)$ is measurable; the same, symmetric condition must hold also when taking the t-sections for any t in $[a,b]$

(A2): for any compact set Q in A, there exists L^1 maps $\phi_Q, \psi_Q : \mathbb{R} \to [0, +\infty[$ such that for any $(t,x), (t,y)$ in Q

$$\begin{cases} |G(t,x)| \leq \phi_Q(t) \\ |G(t,x) - G(t,y)| \leq \psi_Q(t)|x-y| \end{cases} \tag{A.5}$$

Definition A.1. An *optimal pair* is a couple (x_*, u_*) consisting of an admissible curve x_* and a control $u_* \in \mathcal{U}$ that attains the minimum in (A.1), that is $I(x_*, u_*) \leq I(x, u)$ for any other admissible curve x and control $u \in \mathcal{U}$.

Let X be a real vector space. A subset $C \in X$ is called *cone* if it is closed under multiplication by nonnegative scalars, that is, if $v \in C$ and $\alpha \in [0, +\infty[$ then $\alpha x \in C$.
Given a cone C its *polar cone* $C^\perp \subseteq X^*$ is defined by

$$C^\perp = \{p \in X^* : \langle p, v \rangle \geq 0 \quad \forall v \in C\} \tag{A.6}$$

We will now deal with convex cones.

Definition A.2. Let S be any subset of $\mathbb{R}^n$ and $s \in S$. A convex cone $K \in \mathbb{R}^n$ is called a *(Boltyanskii) approximating cone to S at s* if for some integer ν there exists a convex come $C \subseteq \mathbb{R}^\nu$, an open neighborhood of 0 in $\mathbb{R}^\nu$ and a continuous map $F : V \cap C \to S$ such that $F(0) = s$, F admits directional derivatives in 0 along the vectors of C and $\{\partial_v F(0) : v \in C\} = K$.

For example, if S is a smooth manifold in $\mathbb{R}^n$, any convex cone contained in $T_s S$ is an approximating cone for S at s. We also note that in the trivial case $S = \{s\}$ the only possible approximating cone is the null space $\{0\}$. We will be dealing with this situation in the future.

A.0.2. Statement of the theorem. We are now ready to state the Maximum Principle.

Theorem A.3 (The Pontriagyn Maximum Principle). *Let (x_*, u_*) be an optimal pair and let C be and approximating cone to $\mathcal{S}$ in $x_*(b)$. There exists an absolutely continuous dual curve $\lambda : [a,b] \to T^*\mathbb{R}^n$ along x_* and a constant $\lambda_0 \geq 0$ such that*

adjoint equation: *for all $t \in [a,b]$*

$$\dot{\lambda}(t) = -\Big(\big\langle d_x f(t, x_*(t), u_*(t))^T, \lambda(t) \big\rangle - \lambda_0 d_x L(t, x_*(t), u_*(t)) \Big) \tag{A.7}$$

maximization: *for a.e. $t \in [a,b]$*

$$\mathcal{H}(t, \lambda_0, \lambda(t), x_*(t), u_*(t)) = \max_{u \in \mathcal{U}} \mathcal{H}(t, \lambda_0, \lambda(t), x_*(t), u) \tag{A.8}$$

nontriviality: *for all $t \in [a,b]$*

$$(\lambda_0, \lambda(t)) \neq (0,0) \tag{A.9}$$

transversality: *if $C^\perp$ denotes the polar cone of C (see eq.* (A.6)*)*

$$\lambda(b) + \lambda_0 dg(x(b)) \in C^\perp. \tag{A.10}$$

The curve λ is called *dual curve associated with γ*. If we replace $\mathcal{H}$ with $\mathcal{H}' : [a,b] \times (\mathbb{R}^{n+1})^* \times \mathbb{R}^n \times U \to \mathbb{R}$ defined by

$$\mathcal{H}('t, \lambda_0, \lambda, x, u) = \langle \lambda, f(t,x,u) \rangle + \lambda_0 L(t,x,u), \tag{A.11}$$

the maximization condition (A.8) can be proven to be equivalent to the following:

minimization: for a.e. $t \in [a.b]$

$$\mathcal{H}'(t, \lambda_0, \lambda(t), x_*(t), u_*(t)) = \min_{u \in \mathcal{U}} \mathcal{H}'(t, \lambda_0, \lambda(t), x_*(t), u). \tag{A.12}$$

APPENDIX B

Open Mapping Theorems

In this Appendix we state two Theorem, along with the necesasry notation and references, that present sufficient conditions for a smooth mapping to be open. We do not state them in a full generality, rather, in a formulation adapted to our contest.

Recall that a map between topological spaces $F : X \to Y$ is called *open* if it maps open sets into open sets. If $x \in X$, F is said to be *open at x* is there exist an open neighborhood U of x in X such that $F\big|_U$ is open. The classical theorem is the following.

THEOREM B.1 (The Open Mapping Theorem). *Let X and Y be Banach spaces and let $L : X \to Y$ be a linear and continuous map. Then, if it is surjective, it is an open mapping.*

A proof can be found in many textbooks of Functional Analysis, see for example Theorem 5.10, Chapter 5 of [**11**]. We are interested in the case when Y is finite dimensional, in what follows we state theorems under this assumption.

THEOREM B.2. *Let X and Y be normed spaces, Y be finite dimensional, and let Ω be an open subset of X. Let x be in Ω and $F : \Omega \to Y$ be smooth map. If its differential at x, $dF(x) : X \to Y$, is surjective, than F is open at x.*

The proof is a direct consequence of Theorem B.1 and of the Inverse Function Theorem.

B.1. Second-order conditions

We first recall some definitions about quadratic form over vector spaces.

Given two real vector spaces V and W, a *symmetric bilinear map* from V to W is a map $B : V \times V \to W$ such that it is linear in each argument and $B(v, v') = B(v', v)$ for any v, v' in V. The *quadratic form associated with B* is the map $Q_B : V \to W$ defined by $Q_B(v) = B(v, v)$ for any $v \in V$. Viceversa a quadratic form Q completely determines the bilinear symmetric form to which it is associated by the formula $B(v, v') = \frac{1}{2}(Q(v + v') - Q(v) - Q(v'))$.

A real quadtratic form $Q : V \to \mathbb{R}$ is said to be *positive semidefinite* if $Q(v, v) \geq 0$ for any $v \in V$, and *positive definite* if $Q(v, v) > 0$ for any v in $V \setminus \{0\}$. When a quadratic form is positive semidefinite (resp. positive definite) one often writes $Q \geq 0$ (resp. $Q > 0$).

DEFINITION B.3. Given a quadratic form $Q : V \to W$ and a covector $\lambda \in W^*$, *the λ-index of* Q is the (possibly infinite) nonnegative integer $\operatorname{ind}_\lambda Q$ defined by

$$\operatorname{ind}_\lambda Q = \max\{\dim U : U \text{ is a subspace of } V \text{ so that } \lambda \circ Q > 0 \text{ on } U\}. \tag{B.1}$$

The index of Q is the (possibly infinite) nonnegative integer $\operatorname{ind} Q$ defined by

$$\operatorname{ind} Q = \min\{\operatorname{ind}_\lambda Q : \lambda \in W^* \setminus \{0\}\}. \tag{B.2}$$

The fundamental example is when $V = \mathbb{R}^n$ and $W = \mathbb{R}$. In this situation any nonzero covector λ consists of a multiplication by a nonzero real number. So, if Q is a quadratic form with eigenvalues $\mu_1, \dots \mu_n$, its index is the maximum $k \in \mathbb{N}$ such that there exist k positive eigenvalues and k negative eigenvalues between $\mu_1, \dots \mu_n$.

Now recall also the definitions of Cokernel and Hessian matrix given in Notation and Conventions 0.0.1 and 0.0.2. Given two normed spaces X and Y, an open set Ω in X, a point x in Ω and a C^2 map $F : \Omega \to Y$, the *restricted Hessian* of F at the point x is the quadratic form obtained first by restricting it to $\operatorname{Ker} dF(x)$, then projecting it to $\operatorname{Coker} dF(x)$. In other words, if

- $\cdot\big|_{\operatorname{Ker} dF(x)}$ is the restriction from X to $\operatorname{Ker} dF(x)$
- $\pi_{\operatorname{Coker} dF(x)}$ is the canonical projection from Y to $\operatorname{Coker} dF(x)$

$$H_F(x) : \operatorname{Ker} dF(x) \to \operatorname{Coker} dF(x) \tag{B.3}$$

is defined to be $H_F(x) = \pi_{\operatorname{Coker} dF(x)} \left(\operatorname{Hess} F(x)\big|_{\operatorname{Ker} dF(x)} \right)$.

THEOREM B.4 (A Second-Order Open Mapping Theorem). *Let X and Y be Banach spaces, Y be finite dimensional, and let Ω be an open subset of X. Let x be in Ω and $F : \Omega \to Y$ be a smooth map. Let $r = \operatorname{corank} dF(x)$. If $\operatorname{ind} H_F(x) \geq r$, then F is open at x.*

This theorem is a particular case of Theorem 20.6 in Chapter 20 of [**8**], and a proof can be found therein.

Bibliography

[1] R. Monti, "Distances, Boundaries and Surface Measures in Carnot-Carathéodory Spaces," doctoral thesis, University of Trento, 2001.

[2] R. Montgomery, *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, vol. 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2002.

[3] A. Bellaïche and J.-J. Rislers, eds., *Sub-Riemannian Geometry*, vol. 144 of *Progress in Mathematics*. Birkhäuser, 1996.

[4] W. Liu and H. J. Sussmann, *Shortest Paths for Sub-Riemannian Metrics on Rank-Two Distributions*, vol. 118 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 1995.

[5] B. A. Dubrovin, S. P. Novikov, and A. T. Fomenko, *Contemporary Geometry (Part 1)*, vol. 93 of *Graduated Text in Matematics*. Springer-Verlag, 1991.

[6] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, vol. 2 of *Applied Mathematics*. American Institute of Mathematical Sciences, 2007.

[7] R. Monti, "The regularity problem for sub-Riemannian geodesics," *Geometric Control Theory and Sub-Riemannian Geometry*, 2014.

[8] A. Agracev and Y. Sachov, *Control Theory from the Geometric Viewpoint*, vol. 87 of *Encyclopaedia of Mathematical Sciences*. Springer, 2004.

[9] G. P. Leonardi and R. Monti, "End-point equations and regularity of sub-riemannian geodesics," *GAFA Geometric and Functional Analysis*, 2008.

[10] W. S. Massey, *Algebraic Topology: An Introduction*, vol. 56 of *Graduated Text in Matematics*. Springer-Verlag, 1989.

[11] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, vol. 87 of *Pure and Applied Mathematics*. Wiley, second ed., 1999.