1

Introduction to Nonlinear Optimal Control

Bernard Bonnard[1] and Jean-Baptiste Caillau[2]

[1] Université de Bourgogne, Bâtiment Sciences Mirande, BP 47870, F-21078 Dijon. E-mail: Bernard.Bonnard@u-bourgogne.fr

[2] ENSEEIHT-IRIT (UMR CNRS 5505), 2 Rue Camichel, F-31071 Toulouse. E-mail: caillau@n7.fr.

The maximum principle is presented in the weak and general forms. The standard proofs are detailed, and the connection with the shooting method for numerical resolution is made. A brief introduction to the micro-local analysis of extremals is also provided. Regarding second-order conditions, small time-optimality is addressed by means of high order generalized variations. As for local optimality of extremals, the conjugate point theory is introduced both for regular problems and for minimum time singular single input affine control systems. The analysis is applied to the minimum time control of the Kepler equation, and the numerical simulations for the corresponding orbit transfer problems are given. In the case of state constrained optimal control problems, necessary conditions are stated for boundary arcs. The junction and reflection conditions are derived in the Riemannian case.

1.1 Introduction

The objective of this article is to present available techniques to analyze optimal control problems of systems governed by ordinary differential equations. Coupled with numerical methods, they provide tools to solve practical problems. This will be illustrated by the minimum time transfer between Keplerian orbits.

The material is organized as follows. Section 1.2 is devoted to the standard maximum principle who was formulated and proved by Pontryagin and his collaborators in 1956. We follow in the presentation the line of the discovery, see [9]. First of all, we give the weak version, assuming the control domain open. Then we formulate and prove the general theorem along the lines of [15].

Work supported in part by the French Space Agency (contract 02/CNES/0257/00) and done in the framework of the HYCON Network of Excellence, contract number FP6-IST-511368.

A. Loría et al. (Eds.): Advanced Topics in Control Systems Theory, LNCIS 328, pp. 1–60, 2006.

The maximum principle is a necessary optimality result and further conditions are usually required to select minimizers. The aim of Section 1.3 is to present the recent techniques developed to achieve this task. They use the second-order variation along a reference extremal solution of the maximum principle and are directly applicable when the control domain is open. The problem is to test the sign of this second variation. This is done in two steps. First, we must check optimality for small time. To this end, we use special variations and make direct evaluations of the accessibility set, especially using the Baker-Campbell-Hausdorff formula. This approach has provided a generalization of the maximum principle called the high order maximum principle, first obtained by Krener [13]. This result can be applied in the so-called singular case where the standard maximum principle is not able to distinguish minima from maxima. A consequence of this generalization is to get second-order computable conditions in the singular case: generalized Legendre and Goh conditions. The second step, which does not concern small time, is the concept of conjugate point: the problem is to compute in the $\mathcal{C}^1$ topology the first time when a reference trajectory loses local optimality. We present an algorithm to compute this time in the smooth case. This computation is based on the concept of Lagrangian singularity related to the second-order derivative. We give the elements of symplectic geometry necessary to the understanding. One practical motivation for the discovery of the maximum principle was coming from the space engineering. In Section 1.4 we present applications of the afore-mentioned techniques to investigate the minimum time transfer of a spaceship between Keplerian orbits. They are combined with geometrical analysis and numerical simulations so as to compute the optimal solution. The final section deals with the necessary conditions for state constrained problems. The presentation is geometric, is the spirit of Gamkrelidze approach [18]. The conditions, due to Weierstraß, are proved in the planar case.

1.2 Optimal Control and Maximum Principle

1.2.1 Preliminaries

In this section, we consider a system written in local coordinates as

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

where, for each time t, $x(t)$ is in $\mathbf{R}^n$, $u(t)$ in $U \subset \mathbf{R}^m$, and where (x, u) represents a trajectory-control pair defined on an interval $[0, T]$. We denote by $\mathcal{U}$ the class of admissible controls. To each trajectory we assign a cost of the form

$$c(x, u) = \int_0^T f^0(x, u)dt$$

where T can be fixed or not. The optimal control problem is to minimize this cost functional among all trajectories of the system satisfying prescribed boundary conditions of the form

$$x(0) \in M_0,\ x(T) \in M_1.$$

Our system can be extended to a state-cost system according to

$$\dot{x}^0 = f^0(x, u) \tag{1.1}$$

$$\dot{x} = f(x, u) \tag{1.2}$$

which we will also write, with $\widetilde{x} = (x^0, x)$ and $x^0(0) = 0$,

$$\dot{\widetilde{x}} = \widetilde{f}(\widetilde{x}, u).$$

In order to define the necessary optimality conditions, our problem has to be tamed in the following way. For each admissible control u, the corresponding solution $\widetilde{x}(t, \widetilde{x}_0, u)$ starting at time $t = 0$ from $\widetilde{x}_0 = (0, x_0)$ has to be uniquely defined on a maximal interval and has to be an absolutely continuous solution of the system (1.1)-(1.2) almost everywhere. Moreover, the differential of this solution with respect to the initial condition has to be defined, absolutely continuous and solution of the linear differential system

$$\frac{d}{dt}\frac{\partial \widetilde{x}}{\partial \widetilde{x}_0} = \frac{\partial \widetilde{f}}{\partial \widetilde{x}}(\widetilde{x}(t, \widetilde{x}_0, u))\frac{\partial \widetilde{x}}{\partial \widetilde{x}_0}$$

called the *variational system*. Those basic existence, uniqueness and regularity results are standard under the following assumptions.

(i) The set of admissible controls if the set of locally bounded mappings defined on the real line.
(ii) The function $\widetilde{f}$ and its partial derivative $\partial \widetilde{f}/\partial \widetilde{x}$ are continuous.
(iii) The prescribed boundary manifolds are regular submanifolds of $\mathbf{R}^n$.

The approach of our work is geometric and the important concept is the *accessibility set* attached to the system $\dot{x} = f(x, u)$ defined by

$$\mathcal{A}_{x_0,T} = \{x(T, x_0, u),\ u \in \mathcal{U}\}$$

when the initial condition is x_0 and the final time T. Observe that if (x, u) is optimal, the extremity $\widetilde{x}(T, \widetilde{x}_0, u)$ of the extended trajectory must clearly belong to the boundary of the accessibility set of the extended system. The maximum principle is a necessary condition for $\widetilde{x}(T, \widetilde{x}_0, u)$ to belong to $\partial \mathcal{A}_{\widetilde{x}_0,T}$.

1.2.2 The Weak Maximum Principle

We assume that f is smooth and that the set of admissible controls is the set of locally bounded mappings taking values in U, an *open* subset of $\mathbf{R}^m$. If we introduce the *endpoint mapping*, x_0 and T being fixed,

$$E_{x_0,T} \;:\; u \in \mathcal{U} \mapsto x(T, x_0, u)$$

then the accessibility set is the image of the mapping. Since the final time is fixed, the set $\mathcal{U}$ is endowed with the $\mathrm{L}^\infty([0,T])$-norm topology:

$$\|u\| = \text{Ess Sup}_{t\in[0,T]}|u(t)|$$

where $|.|$ is any equivalent norm on $\mathbf{R}^n$.

First and Second Variation

It can be easily proved that the endpoint mapping is $\mathcal{C}^\infty$ for the L^∞ topology and that the first and second variations are computed in the following way. Fix $x(0) = x_0$ and denote by (x, u) the reference solution defined on $[0, T]$. Let $x + \delta x$ be the solution starting from x_0 and generated by $u + \delta u$ where δu is an L^∞ variation. Since f is smooth we can write:

$$\begin{aligned} f(x+\delta x, u+\delta u) &= f(x,u) + \frac{\partial f}{\partial x}(x,u)\delta x + \frac{\partial f}{\partial u}(x,u)\delta u \\ &+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x,u)(\delta x, \delta x) + \frac{\partial^2 f}{\partial x \partial u}(x,u)(\delta x, \delta u) + \frac{1}{2}\frac{\partial^2 f}{\partial u^2}(x,u)(\delta u, \delta u) + \cdots \end{aligned}$$

Writing that $x + \delta x$ is solution, we have

$$\dot{x} + \delta\dot{x} = f(x+\delta x, u+\delta u)$$

and we can decompose δx as $\delta_1 x + \delta_2 x + \cdots$ where $\delta_1 x$ is linear in u and $\delta_2 x$ quadratic. By identification,

$$\delta_1\dot{x} = \frac{\partial f}{\partial x}(x(t),u(t))\delta_1 x + \frac{\partial f}{\partial u}(x(t),u(t))\delta u(t)$$

that is $\delta_1 x$ is solution of the system linearized along the reference trajectory and

$$\begin{aligned} \delta_2\dot{x} &= \frac{\partial f}{\partial x}(x(t),u(t))\delta_2 x + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x(t),u(t))(\delta_1 x(t), \delta_1 x(t)) \\ &+ \frac{\partial^2 f}{\partial x \partial u}(x(t),u(t))(\delta_1 x(t), \delta u(t)) + \frac{1}{2}\frac{\partial^2 f}{\partial u^2}(x(t),u(t))(\delta u(t), \delta u(t)) \end{aligned} \tag{1.3}$$

with $\delta_1 x(0) = \delta_2 x(0) = 0$ since $\delta x(0) = 0$. Let $A(t)$ be the matrix $\partial f/\partial x$ along $(x(t), u(t))$ and let Φ be the matrix valued fundamental solution of

$$\dot{\Phi} = A(t)\Phi$$

with $\Phi(0) = I$. We observe that the first and second variations can be computed using the standard formula to integrate linear differential equations. In particular, by setting $B(t) = \partial f/\partial u$ along $(x(t), u(t))$, the Fréchet derivative is:

$$\delta_1 x(T) = \Phi(T)\int_0^T \Phi^{-1}(s)B(s)\delta u(s)ds. \tag{1.4}$$

Statement and Proof of the Weak Maximum Principle

Let (x, u) be the reference trajectory defined on $[0, T]$, T fixed. Assume that $x(T)$ belongs to the boundary of the accessibility set. Then, from the open mapping theorem, the control has to be a singularity of the endpoint mapping and we must have

$$\text{rank } E'_{x_0,T}(u) < n$$

where $E'_{x_0,T}(u)$ is the Fréchet derivative at u computed according to (1.4),

$$E'_{x_0,T}(u) = \delta_1 x(T).$$

To get the weak maximum principle, we take a nonzero covector $\bar{p}$ orthogonal to the image of $E'_{x_0,T}(u)$ and we set:

$$p(t) = \bar{p}\Phi(T)\Phi^{-1}(t).$$

Hence, p is solution of the *adjoint equation*

$$\dot{p} = -p\frac{\partial f}{\partial x}(x(t), u(t))$$

and, by construction,

$$\int_0^T p(t)B(t)\delta u(t)dt = 0$$

for all variations in $\mathrm{L}^\infty([0, T])$. As a result, $p(t)B(t)$ is zero almost everywhere and we have proved the following proposition.

Proposition 1.1. *Let (x, u) be a trajectory defined on $[0, T]$ such that $x(T)$ belongs the boundary of $\mathcal{A}_{x_0,T}$, the control set being open in $\mathbf{R}^m$. There exists an absolutely continuous nonvanishing covector function p defined on $[0, T]$ such that the triple (x, p, u) is almost everywhere solution of*

$$\dot{x} = \frac{\partial H}{\partial p}(x, p, u), \;\; \dot{p} = -\frac{\partial H}{\partial x}(x, p, u)$$
$$\frac{\partial H}{\partial u}(x, p, u) = 0$$

where $H(x, p, u) = \langle p, f(x, u)\rangle$ is the Hamiltonian *of the system.*

The covector function p is called the *adjoint state*. In particular, if (x, u) is time minimizing, $x(T)$ belongs to the boundary of the accessibility set and satisfies the previous necessary conditions.

1.2.3 The Maximization Condition

Actually, the second-order variation can be used so as to derive more conditions for time-optimality as explained in [9]. Let us denote by Π the image of the Fréchet derivative of the endpoint mapping at u. As previously noticed, if the reference trajectory is optimal, the hyperplane Π is at least of codimension one. Consider now the generic case where Π is of codimension exactly one, and where the reference trajectory is differentiable at T and intersects Π transversely. The adjoint vector at T is orthogonal to Π and thus uniquely defined up to a scalar. Morevor, since the trajectory is transverse to Π at T, we can use the normalization

$$p(T)f(x(T), u(T)) > 0.$$

We introduce the *intrinsic second-order derivative* which is defined as the restriction of the second variation to the kernel K of $E'_{x_0,T}(u)$ projected to $\Pi^\perp$. It is given by

$$\delta u \in K \mapsto p(T)\delta_2 x(T)$$

with $\delta_2 x$ computed by means of (1.3). If u is time-optimal, we must have (see Fig. 1.1):

$$p(T)\delta_2 x(T) \leq 0, \;\; \delta u \in K.$$

Expliciting $\delta_2 x(T)$, one gets the additional standard *Legendre-Clebsch* condition,

$$\frac{\partial^2 H}{\partial u^2} \leq 0$$

and finally obtains the (local) maximization condition : almost everywhere,

$$H(x(t), p(t), u(t)) = \max_{v \in V_t} H(x(t), p(t), v)$$

with, for each t, V_t a neighbourhood of $u(t)$.

1.2.4 Maximum Principle, Fixed Time

Statement

Consider a system $\dot{x} = f(x, u)$ with, as before, f and $\partial f/\partial x$ continuous functions on an open subset of $\mathbf{R}^{n+m}$. The set of admissible controls $\mathcal{U}$ is again the set of locally bounded functions taking values in a fixed subset U of $\mathbf{R}^m$, and such that the responses starting at $t = 0$ from x_0 are defined on the whole interval $[0, T]$, T fixed. Let (x, u) be a reference trajectory such that the endpoint $x(T)$ belongs to the boundary of the accessibility set. Then, there exists a non-trivial covector absolutely continuous function p such that the triple (x, p, u) is almost everywhere solution of the equations

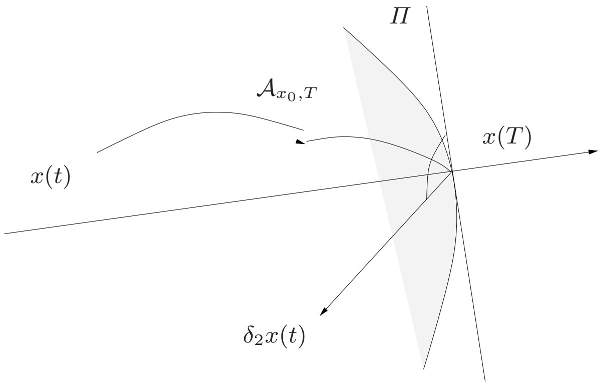

Fig. 1.1. Legendre-Clebsch condition: non-positivity of the intrinsic second-order derivative

$$\dot{x} = \frac{\partial H}{\partial p}(x, p, u), \;\; \dot{p} = -\frac{\partial H}{\partial x}(x, p, u) \tag{1.5}$$

where $H(x, p, u) = \langle p, f(x, u)\rangle$ is the Hamiltonian. Moreover, the maximization condition holds almost everywhere along the extremal triple,

$$H(x(t), p(t), u(t)) = M(x(t), p(t))$$

where $M(x, p) = \max_{v\in U} H(x, p, v)$, and $t \mapsto M(x(t), u(t))$ is constant on $[0, T]$.

The Proof of the Maximum Principle

Needle variations. The basic concept needed to prove the maximum principle is the concept of *needle variation*. Indeed, because the control domain is arbitrary, standard L^∞ variations of the reference control used when U is open have to be replaced by L^1 elementary ones of the form:

$$u_{\pi_1}(t, \varepsilon) = \begin{cases} u_1 \text{ on } [t_1 - \varepsilon l_1, t_1] \\ u(t) \text{ everywhere else on } [0, T] \end{cases}$$

where the needle variation is the triple $\pi_1 = (t_1, l_1, u_1)$, $0 < t_1 < T$, $l_1 \geq 0$, u_1 in U. For $\varepsilon > 0$ small enough, the perturbed control is a well defined admissible control with response $x_{\pi_1}(t, \varepsilon)$ starting from x_0. Clearly, $x_{\pi_1}(t, \varepsilon)$ tends to $x(t)$ uniformly on $[0, T]$ when ε tends to 0, and is continuous with respect to (π_1, t, ε). To get differentiability with respect to ε, we require that t_1 be a Lebesgue point so that

$$\int_{t_1-\varepsilon}^{t_1} f(x(t), u(t))dt = f(x(t_1), u(t_1)) + o(\varepsilon).$$

From standard integration theory, the subset $\mathcal{L}$ of Lebesgue points has full measure on $[0,T]$. If π_1 is such a needle variation, then the corresponding response defines a curve at $x(t_1)$, $\alpha(\varepsilon) = x_{\pi_1}(t_1,\varepsilon)$ whose tangent vector is

$$\dot\alpha(0) = l_1(f(x(t_1),u_1) - f(x(t_1),u(t_1))).$$

This comes from the estimate

$$x_{\pi_1}(t_1,\varepsilon) = x(t_1 - l_1\varepsilon) + \int_{t_1-l_1\varepsilon}^{t_1} f(x_{\pi_1}(t,\varepsilon),u_1)dt \tag{1.6}$$

$$= x(t_1) - \varepsilon l_1 f(x(t_1),u(t_1)) + \varepsilon l_1 f(x(t_1),u_1) + o(\varepsilon). \tag{1.7}$$

This tangent vector is called the *elementary perturbation vector* associated to the needle variation and is denoted v_{π_1}.

Remark 1.1. If t_1 is a Lebesgue point, for any positive η, from the definition one can find another Lebesgue point t such that $|t-t_1| \le \eta$ and $|f(x(t),u(t)) - f(x(t_1),u(t_1))| \le \eta$.

Parallel displacements along the trajectory. We first recall a standard but crucial result. Let $\dot x = X(x)$ be a smooth differential equation, and let $\varphi_t = \exp tX$ define the local one parameter group. If $\alpha(\varepsilon)$ is a smooth curve at x_0, then $(t,\varepsilon) \mapsto \beta(t,\varepsilon) = \varphi_t(\alpha(\varepsilon))$ is a smooth two-dimensional surface. Let $x(t) = \exp tX(x_0)$ be the reference curve and Φ_t be the matrix valued solution of the variational equation

$$\delta\dot x = \frac{\partial X}{\partial x}(x(t))\delta x \tag{1.8}$$

with $\Phi_0 = I$. Then, for each fixed t, the derivative at 0 of the curve $\varepsilon \mapsto \beta(t,\varepsilon)$ is the so-called *parallel displacement* $w(t)$ given by

$$w(t) = \Phi_t v$$

where $v = \dot\alpha(0)$. Moreover, if X is analytic and if $\dot\alpha(0) = Y(x_0)$ with Y another analytic vector field, w can be computed using the ad-formula

$$w = \sum_k \frac{t^k}{k!}\mathrm{ad}^k X \cdot Y(x(t))$$

where

$$[X,Y] = \frac{\partial X}{\partial x}(x)Y(x) - \frac{\partial Y}{\partial x}(x)X(x) \tag{1.9}$$

is the Lie bracket[3] and $\mathrm{ad}X$ is the linear operator $\mathrm{ad}X \cdot Y = [X,Y]$. Extending this result to the time depending case, we can transport an elementary

[3] Beware of the sign, here, opposite to some classical texts, *e.g.* [11].

pertubation vector v_{π_1} at $x(t_1)$ along the reference trajectory. The variational equation is

$$\dot{v} = \frac{\partial f}{\partial x}(x(t), u(t))v$$

and if p is solution of the adjoint equation then, by construction,

$$p(t)v(t) = \text{constant}.$$

We note Φ_{t,t_1} the fundamental matrix solution of the variational equation (1.8) with initial condition $\Phi_{t_1,t_1} = I$. From our previous analysis, we know that if v_{π_1} is the elementary perturbation vector, tangent to the curve $\alpha(\varepsilon)$, then for $t \geq t_1$, $\Phi_{t,t_1} v_{\pi_1}$ is the tangent vector to a curve $\beta(t, \varepsilon)$, image of α by the flow (see Fig. 1.2).

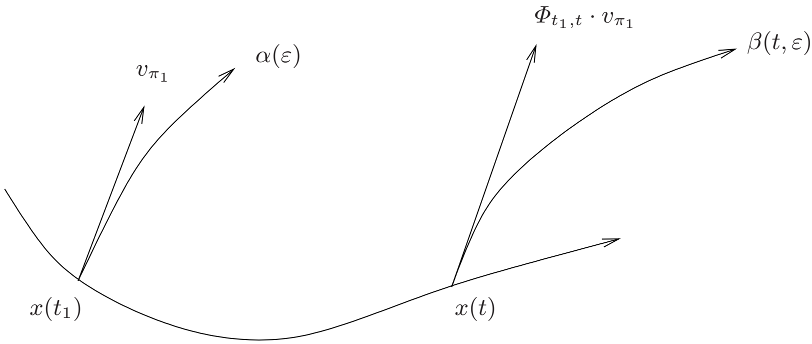

Fig. 1.2. Transport of the elementary perturbation vector by the flow

Definition 1.1. *The* first tangent perturbation cone *or* first Pontryagin cone K_t *at any time* $0 < t \leq T$ *is the smallest convex cone*[4] *in the tangent space at* $x(t)$ *that contains all parallel displacements of all elementary perturbation vectors at Lebesgue points on* $]0, t]$,

$$K_t = \text{cone}(\{\Phi_{t,t_1} v_{\pi_1},\ \pi_1 = (t_1, l_1, u_1) \in \mathcal{L} \times \mathbf{R}_+^* \times U,\ 0 < t_1 \leq t\}).$$

Let now $\pi_i = (t_i, l_i, u_i)$, $i = 1, \ldots, k$, be needle variations with distinct times t_i. Let $\pi = (\pi_1, \ldots, \pi_k)$ be the *complex variation* associated to the perturbed control

$$u_\pi(t, \varepsilon) = \begin{cases} u_i \text{ on } [t_i - \varepsilon l_i, t_i] \\ u(t) \text{ everywhere else on } [0, T] \end{cases}$$

which is well defined for ε small enough because the t_i are distinct. Clearly, the estimate (1.7) can be extended to complex variations as follows.

[4] For a subset A of $\mathbf{R}^n$, the smallest convex cone containing A is $\text{cone}\, A = \{\sum_{i=1}^k \lambda_k x_k,\ k \geq 1, \lambda_1 \geq 0, \ldots, \lambda_k \geq 0\}$.

Lemma 1.1. *Let* $v_i = \Phi_{t,t_i} v_{\pi_i}$ *be parallel displacements of elementary perturbation vectors defined by needle variations* $\pi_i = (t_i, l_i, u_i)$ *with distinct times* t_i, $i = 1, \dots, k$. *Then, the convex combination* $\lambda_1 v_1 + \cdots + \lambda v_k$, $\lambda_i \geq 0$ *and* $\sum_{i=1}^k \lambda_i = 1$, *is tangent to* $x_\pi(t, \varepsilon)$, *the response to the perturbed control* $u_\pi(t, \varepsilon)$ *where* π *is the complex variation* $((t_1, \lambda_1 l_1, u_1), \dots, (t_k, \lambda_k l_k, u_k))$*:*

$$x_\pi(t, \varepsilon) = x(t) + \varepsilon(\lambda_1 v_1 + \cdots + \lambda v_k) + o(\varepsilon).$$

Fundamental lemma. In order to prove the maximum principle, we need a technical lemma which is a consequence of the following byproduct of the Brouwer fixed point theorem [15].

Proposition 1.2. *Let* f *be a continuous mapping from the closed unit ball* B *of* $\mathbf{R}^n$ *into* $\mathbf{R}^n$. *Let* $0 < \varepsilon < 1$ *be such that, for all* x *in the unit sphere* S,

$$|f(x) - x| \leq \varepsilon.$$

Then, $f(B)$ *contains the open ball of radius* $1 - \varepsilon$ *centered at the origin.*

Lemma 1.2. *Let* v *be a nonzero vector interior to* K_t, *then* $x(t) + \varepsilon v$ *lies interior to the accessibility set* $\mathcal{A}_{x_0,t}$ *for all small enough and positive* ε.

Proof. Let v be nonzero and interior to K_t. There are independent parallel displacements $v_1, \dots, v_n$ in K_t such that v is interior to the convex set generated by $v_1, \dots, v_n$. Let Π be the hyperplane defined by these vectors. Since v is interior, any point y in the interior of the cone generated by $v_1, \dots, v_n$ can be written $y = x(t) + \varepsilon(v + r)$ with $\varepsilon > 0$, and r in a suitable open subset of the $n - 1$ dimensional vector space parallel to Π (see Fig. 1.3). For such an r, there are nonnegative scalars $\lambda_1, \dots, \lambda_n$, $\sum_{i=1}^n \lambda_i = 1$, such that $v + r = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Besides, there are needle variations $\pi_i = (t_i, l_i, u_i)$ such that $v_i = \Phi_{t,t_i} v_{\pi_i}$, $i = 1, \dots, n$, and one can assume all Lebesgue points t_i distinct (see remark 1.1). Hence, for ε small enough it is possible to define the perturbed control u_r associated to the complex variation $(t_i, \lambda_i l_i, u_i)_i$. If x_r denotes the corresponding response, Lemma 1.1 asserts that

$$\begin{aligned} x_r(t, \varepsilon) &= x(t) + \varepsilon(\lambda_1 v_1 + \cdots + \lambda_n v_n) + o(\varepsilon) \\ &= x(t) + \varepsilon(v + r) + o(\varepsilon). \end{aligned}$$

Let then define the continuous mapping $g \ : \ (\varepsilon, r) \mapsto x_r(t, \varepsilon)$ into the accessibility set $\mathcal{A}_{x_0,T}$. In coordinates (ε, r), $g(\varepsilon, r) = (\varepsilon + o(\varepsilon), r + o(1))$. As a result, $|g(\varepsilon, r) - (\varepsilon, r)|$ tends to zero when ε does so and, by Proposition 1.2, one can find ε_0, positive and small enough, such that the image by g of $[0, \varepsilon_0] \times \{|r| \leq \varepsilon_0\}$ (with g continuously extended at $\varepsilon = 0$ according to $g(0, r) = (0, r)$) contains $]0, \varepsilon_0[\times \{|r| < \varepsilon_0\}$. Therefore, $\mathcal{A}_{x_0,t}$ is a neighbourhood of $x(t) + \varepsilon v$ for $0 < \varepsilon < \varepsilon_0$, hence the result. ∎

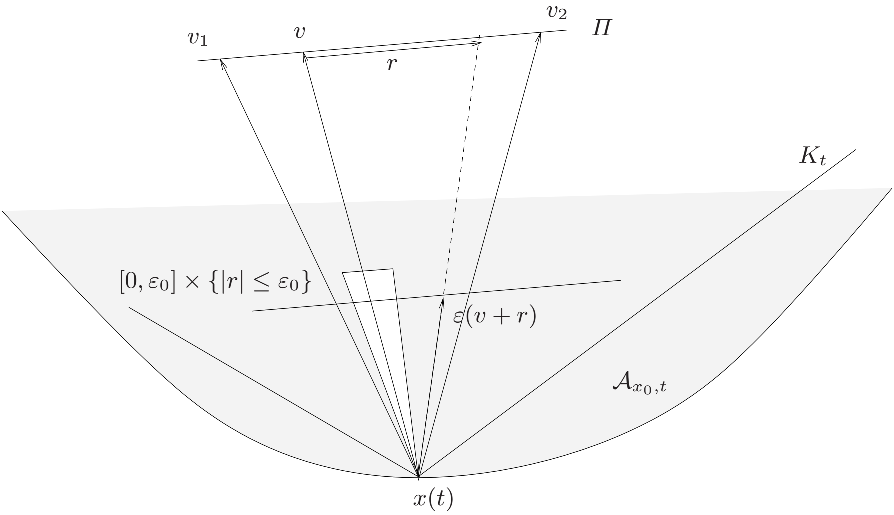

Fig. 1.3. Conical neighbourhood of vector v in the accessibility set

End of the proof. To finish the proof of the maximum principle, we just use a geometric separation argument. Indeed, if $x(T)$ belongs to the boundary of $\mathcal{A}_{x_0,T}$, then there exists a sequence of points x_n not belonging to the interior of the accessibility set, converging to $x(T)$ and such that, up to a subsequence, the unit vectors $(x_n - x(T))/|x_n - x(T)|$ have a limit v when n tends to $+\infty$. This vector v is not interior to K_T otherwise, from the fundamental Lemma 1.2, $x(T) + \varepsilon v$ would be interior to $\mathcal{A}_{x_0,T}$ for any small and positive ε, and so would be x_n for n big enough. The convex cone K_T is thus included in a half-space defined by a separating hyperplane Π. Let $\bar{p}$ be the unit normal to Π oriented outwards K_T, and let us denote p the solution of the adjoint equation

$$\dot{p} = -p\frac{\partial f}{\partial x}(x(t), u(t))$$

satisfying $p(T) = \bar{p}$. Then, the maximization condition must hold almost everywhere. Indeed, let t_1 in $]0, T[$ be a Lebesgue point, and let u_1 be in U. The elementary perturbation vector $v_{\pi_1} = f(x(t_1), u_1) - f(x(t_1), u(t_1))$ associated to $\pi_1 = (t_1, 1, u_1)$ is in K_{t_1}, so $v = \Phi_{T,t_1} v_{\pi_1}$ is in K_T and

$$\langle p(t_1), v_{\pi_1} \rangle = \langle \bar{p}, v \rangle \leq 0$$

that is $H(x(t_1), p(t_1), u_1) \leq H(x(t_1), p(t_1), u(t_1))$. Accordingly, the Hamiltonian is maximized at t_1, $H(x(t_1), p(t_1), u(t_1)) = \max_{v \in U} H(x(t_1), p(t_1), v)$, and the conclusion proceeds from the fact that the set of Lebesgue points has full measure. Standard arguments allow to prove that $t \mapsto M(x(t), p(t)) =$

$\max_{v\in U} H(x(t), p(t), u)$ is absolutely continuous with zero derivative almost everywhere : hence M is constant along (x, p).

Application to Time-optimal Control

We can apply our result to the time-optimal control problem. Indeed, assume that the reference control is time-optimal on $[0, T]$. Then, for each t in $]0, T]$, the point $x(t)$ is in $\partial A(x_0, T)$ so that $\dot{x}(t) = f(x(t), u(t))$ cannot be interior to the first order cone $K(t)$. Indeed, from the fundamental lemma, $x(t+\varepsilon)$ would be in $\mathcal{A}_{x_0,t}$ for $\varepsilon > 0$ small enough, otherwise, contradicting optimality. Hence, we have the additional condition $\langle p(t), f(x(t), u(t))\rangle \geq 0$ and the reduced Hamiltonian is constant and positive.

1.2.5 Maximum Principle, General Case

We formulate the result which can be used to analyze general finite dimensional optimal control problems. We consider a system $\dot{x} = f(x, u)$ written in local coordinates x in $\mathbf{R}^n$, where the set $\mathcal{U}$ of admissible controls is the set of locally bounded functions valued in a fixed control domain $U \subset \mathbf{R}^m$. Let M_0 and M_1 be the regular submanifolds defining the boundary conditions, and let

$$c(x, u) = \int_0^T f^0(x, u)dt$$

be the cost functional assigned to an admissible control and its response x assumed to be defined on $[0, T]$, T *free*. As before, $\widetilde{x} = (x^0, x)$ is the cost extended state and $\widetilde{f}$ the extended dynamics. We assume that $\widetilde{f}$ satisfies the previous regularity assumptions, namely that it is continuous on $\mathbf{R}^{1+n+m}$, together with its partial derivative $\partial\widetilde{f}/\partial\widetilde{x}$. Let

$$\widetilde{H}(\widetilde{x}, \widetilde{p}, u) = p^0 f^0(x, u) + \langle p, f(x, u)\rangle$$

be the extended Hamiltonian and

$$\widetilde{M}(\widetilde{x}, \widetilde{p}) = \max_{v\in U} H(\widetilde{x}, \widetilde{p}, v).$$

Theorem 1.1. *If u is an optimal control on $[0, T]$ then there exists an absolutely continuous extended adjoint covector function $\widetilde{p} = (p^0, p)$, nonzero on $[0, T]$ and such that the following equations are satisfied almost everywhere by the triple $(\widetilde{x}, \widetilde{p}, u)$:*

$$\dot{\widetilde{x}} = \frac{\partial\widetilde{H}}{\partial\widetilde{p}}(\widetilde{x}, \widetilde{p}, u),\ \dot{\widetilde{p}} = -\frac{\partial\widetilde{H}}{\partial\widetilde{x}}(\widetilde{x}, \widetilde{p}, u)$$
$$\widetilde{H}(\widetilde{x}, \widetilde{p}, u) = \widetilde{M}(\widetilde{x}, \widetilde{p}).$$

Moreover, $\widetilde{M}(\widetilde{x},\widetilde{p}) = 0$ *on* $[0,T]$ *and* p^0 *is constant and non-positive. Eventually,* p *can be selected at the extremities so as to satisfy the transversality conditions*

$$p(0) \perp T_{x(0)}M_0, \;\; p(T) \perp T_{x(T)}M_1.$$

Proof. We use the necessary conditions for the fixed time case and we extend the cone with additional directions. Indeed, since u is optimal, the endpoint $(x^0(T), x(T))$ of the extended system belongs to the boundary of the extended accessibility set. So there exists a non-trivial augmented adjoint covector $\widetilde{p} = (p^0, p)$ such that, almost everywhere,

$$\widetilde{H}(\widetilde{x},\widetilde{p},u) = \widetilde{M}(\widetilde{x},\widetilde{p})$$

and the maximized Hamiltonian $\widetilde{M}$ is constant along $(\widetilde{x},\widetilde{p})$ on $[0,T]$. In order to extend the first tangent cone $\widetilde{K}_t$ of the extended system, we proceed as follows. Since the time is not fixed, by making time variations $t + \varepsilon\delta t$ of Lebesgue points we can add to $\widetilde{K}_t$ the two vectors $v_\pm = \pm\widetilde{f}(\widetilde{x}(t), u(t))$. The two manifolds M_0, M_1 are embedded into $\mathbf{R}^{1+n}$ by taking $\widetilde{M}_0 = (0, M_0)$ and $\widetilde{M}_1 = (0, M_1)$, with respective tangent bundles $T\widetilde{M}_0$, $T\widetilde{M}_1$. Since the initial condition is relaxed to $\widetilde{M}_0$, we can add to $\widetilde{K}_t$ the parallel displacements in the tangent space to $\widetilde{M}_0$. Hence, the *second tangent perturbation cone* $\widetilde{K}'_t$, $0 < t \leq T$, is defined as the convex cone generated by the vectors:

(i) $\widetilde{\Phi}(t,0)w, \;\; w \in T_{\widetilde{x}(0)}\widetilde{M}_0$.

(ii) $\widetilde{\Phi}(t,t_1)(\widetilde{f}(\widetilde{x}(t_1),v) - \widetilde{f}(\widetilde{x}(t_1),u(t_1)))$ where v is in U and $t_1 \leq t$ is a Lebesgue point.

(iii) $\pm\widetilde{\Phi}(t,t_1)\widetilde{f}(\widetilde{x}(t_1),u(t_1))$ with $t_1 \leq t$ a Lebesgue point.

According to Lemma 1.2, for any vector w interior to $\widetilde{K}'_t$ there exists $\lambda > 0$ and a conic neighbourhood of λw included in the accessibility set $\widetilde{\mathcal{A}}_{x_0} = \cup_{t>0}\widetilde{\mathcal{A}}_{x_0,t}$. In particular, since $\widetilde{x}(T)$ is optimal, the vector $(-1,0)$ of $\mathbf{R}^{1+n}$ does not belong the interior of $\widetilde{K}'_t$, otherwise we could find an admissible control minimizing the cost even more. In order to obtain the transversality condition at the endpoint, we introduce the cone T_1 at $\widetilde{x}(T)$ which is generated by $T_{\widetilde{x}(T)}\widetilde{M}_1$ and the downward vector $(-1,0)$. The second perturbation cone $\widetilde{K}'_t$ and T_1 are separated by an hyperplane Π. Here, we can take a normal vector $\bar{\widetilde{p}} = (p^0, \bar{p})$ at $\widetilde{x}(T)$ with $p^0 \leq 0$ and

$$\bar{\widetilde{p}}\widetilde{K}'_t \leq 0, \;\; \bar{\widetilde{p}}T_1 \geq 0.$$

The corresponding solution $\widetilde{p}$ of the adjoint system with $\tilde{p}(T) = \bar{\widetilde{p}}$ satisfies the maximum principle, including the required transversality conditions. ■

Remark 1.2. The case where the time is fixed can be reduced to our previous case by introducing the time as a new space variable. The result is the same, the maximized Hamiltonian still being constant along $(\widetilde{x}, \widetilde{p})$ but not necessarily zero anymore. The non-autonomous case can be similarly analyzed.

Definition 1.2. *We call* extremal *any triple (x, p, u) solution of the Hamiltonian system and verifying the maximization condition. An extremal also satisfying the transversality conditions is called a* BC-extremal.

1.2.6 Maximum Principle and Shooting Problem

Consider any optimal control problem. It is well posed if there exists an optimal solution. This can be checked by applying the standard Filippov theorem (see [15], p. 259). We assume that there is a solution satisfying the maximum principle. If we denote by $M_i^\perp$, $i = 1, 2$, the cotangent lifts

$$M_i^\perp = \{(x, p) \in T^* M_i \mid x \in M_i,\ p \perp T_x M_i\}$$

we can define the *shooting mapping*

$$S\ :\ (x_0, p_0) \in M_0^\perp \mapsto (x(T), p(T)) \in M_1^\perp. \tag{1.10}$$

An important remark is that the Hamiltonian is linear with respect to the adjoint covector p in order that p has to be taken in the projective space $\mathbf{P}^{n-1} \subset \mathbf{R}^n$. With this normalization, the number of equations is equal to the number of variables. For instance, if we consider the time-optimal control problem with fixed extremities x_0, x_1, the shooting problem is to find a time T and an initial adjoint covector p_0 in $\mathbf{P}^{n-1}$ such that

$$x(T, u, p_0) - x_1 = 0$$

where u is computed by means of the maximization condition, and where $x(., u, p_0)$ is the solution of the Hamiltonian system (1.5) with $x(0) = x_0$ and $p(0) = p_0$.

1.2.7 Introduction to the Micro-analysis of the Extremal Solutions

Consider first the case where the control domain U is open. The maximization condition gives us the conditions

$$\frac{\partial \widetilde{H}}{\partial u} = 0,\ \frac{\partial^2 \widetilde{H}}{\partial u^2} \leq 0$$

and the *regular case* occurs when the strict Legendre-Clebsch condition is satisfied:

$$\frac{\partial^2 \widetilde{H}}{\partial u^2} < 0.$$

In this case, applying the implicit function theorem to solve $\partial \widetilde{H}/\partial u = 0$ leads to compute the reference control as a smooth *dynamic feedback*,

$$(x, p) \mapsto u(x, p). \tag{1.11}$$

By plugging (1.11) into $\widetilde{H}$, we define a true Hamiltonian function. However, $\partial \widetilde{H}/\partial u = 0$ is in general a nonlinear equation with several zeros associated to various local maxima $u_i(x, p)$ of $\widetilde{H}$. The master Hamiltonian thus defines several Hamiltonian functions H_i among which an absolute maximum must be chosen. Memory of all those Hamiltonians must be kept since, along a reference extremal, bifurcations between different local maxima may occur to provide the global maximum. This key phenomenon is crucial in the analysis of the extremal solutions, see for instance the pioneering article [8] where the problem is addressed in the framework of calculus of variations with a non-convex one-dimensional Lagrangian function.

1.2.8 Affine Control Systems

In many applications the control system is

$$\dot{x} = F_0(x) + \sum_{i=1}^{m} u_i F_i(x)$$

and, for the time-optimal control problem, the *reduced Hamiltonian* is considered:

$$H = H_0 + \sum_{i=1}^{m} u_i H_i$$

where $H_i = \langle p, F_i \rangle$, $i = 1, \ldots, m$ are the Hamiltonian lifts of the vector fields. In this case, the Hamiltonian is affine in the control and the problem is singular in the sense that

$$\frac{\partial^2 H}{\partial u^2} = 0.$$

Hence, the Legendre-Clebsch condition cannot be used to separate maxima from minima if the extremal solution is interior to the control domain and higher order conditions are required.

1.3 More Second-order Conditions

1.3.1 High-order Maximum Principle

Consider first the time-optimal control problem for a single input affine control system

$$\dot{x} = F_0(x) + uF_1(x)$$

where $|u| \leq 1$. According to the maximum principle, the extremals are solutions of

$$\dot{x} = \frac{\partial H}{\partial p}(x, p, u), \ \ \dot{p} = -\frac{\partial H}{\partial x}(x, p, u)$$

together with the maximization condition

$$H(x, p, u) = \max_{|v| \leq 1} H(x, p, v)$$

where $H(z, u) = H_0(z) + uH_1(z)$, $H_i = \langle p, F_i \rangle$ for $i = 0, 1$, and $z = (x, p)$.

Definition 1.3. *Let (z, u) be a reference extremal defined on $[0, T]$. It is called* regular *if $u(t) = \text{sign } H_1(z(t))$ almost everywhere on $[0, T]$, and* singular *if $H_1(z(t)) = 0$ for all t in $[0, T]$.*

More general extremals are concatenation of regular and singular subarcs. We begin by computing the singular controls defined by the constraint $H_1(z) = 0$.

Computation of Singular Extremals

The weak and the general maximum principle lead to the same equation, $H_1(z) = 0$. To compute the corresponding trajectories, we differentiate with respect to time and use the Hamiltonian formalism. Differentiating twice with respect to t in $[0, T]$, we get:

$$\{H_1, H_0\}(z(t)) = 0$$
$$\{\{H_1, H_0\}, H_0\}(z(t)) + u(t)\{\{H_1, H_0\}, H_1\}(z(t)) = 0$$

with the Poisson brackets given by $\{H_X, H_Y\} = \langle p, [X, Y] \rangle$ and the Lie bracket defined as in (1.9).

Definition 1.4. *A singular extremal z is said to be of* minimal order *if, everywhere on $[0, T]$,*

$$\{\{H_1, H_0\}, H_1\}(z(t)) \neq 0.$$

Proposition 1.3. *If z is a singular arc of minimal order, the corresponding singular control is the dynamic feedback*

$$u_s(t) = -\frac{\{\{H_1, H_0\}, H_0\}}{\{\{H_1, H_0\}, H_1\}}(z(t))$$

and the extremal curve is smooth and solution of

$$\dot{z} = \overrightarrow{H}_s(z)$$

contained in $H_1 = \{H_1, H_0\} = 0$ and $H_s = H_0 + u_s H_1$.

A Standard Normalization

Hence a singular arc is in general smooth. Take such an arc, $t \mapsto x(t)$. Restricting if necessary its domain of definition, we can assume that it is one-to-one. Hence, it can be identified locally with the curve $\gamma\ :\ t \mapsto (t, 0, \ldots, 0)$ and is the response of to a smooth control denoted u_γ. The control can be normalized to zero by the feedback $v = u - u_\gamma$. Then, differentiating as before $H_1(z) = 0$, one gets that, everywhere

$$\langle p(t), \mathrm{ad}^k F_0 \cdot F_1(\gamma(t))\rangle = 0,\ k \geq 0.$$

We proved the following.

Proposition 1.4. *Let z be a smooth singular extremal on $[0, T]$, corresponding to a singular control identified to zero. The maximum principle is equivalent to*

$$\mathrm{ad}^k H_0 \cdot H_1(z(t)) = 0,\ k \geq 0 \tag{1.12}$$

everywhere on $[0, T]$.

In (1.12), $\mathrm{ad}^k H_0 \cdot H_1$ denotes the k-th Poisson bracket of H_1 with H_0. This is clearly equivalent to the following lemma.

Lemma 1.3. *Let x be a trajectory defined on $[0, T]$ and associated to the zero control. Assume that, for each t, $V_1(t) = \mathrm{Span}\{\mathrm{ad}^k F_0 \cdot F_1(x(t)),\ k \geq 0\}$ is of maximum rank n. Then, for each $0 \leq t_0 < t_1 \leq T$, the linearized system along x restricted to $[t_0, t_1]$ is controllable and $x(t_1)$ belongs to the interior of the accessibility set $\mathcal{A}_{x(t_0), t_1 - t_0}$.*

This gives a simple interpretation of the maximum principle for single input affine control systems in the open control case.

The Analytic Case

Consider now the case where F_0 and F_1 are real analytic vector fields and let z be the reference extremal defined on $[0, T]$ associated to a control normalized to zero. As before, the maximum principle is subsumed by (1.12) and $V_1(T)$ is the image of the Fréchet derivative of the endpoint mapping which coincides with the first order Pontryagin cone constructed in the proof of the principle. Indeed, if $v_{\pi_1}(t)$ is an elementary perturbation vector with $\pi_1 = (t_1, 1, u_1)$, one has

$$\begin{aligned} v_{\pi_1}(t) &= (F_0 + u_1 F)(x(t_1)) - F_0(x(t_1)) \\ &= u_1 F_1(x(t_1)) \end{aligned}$$

and we can take $u_1 = \pm 1$. The parallel transport can be evaluated using the ad-formula

$$(\exp tF_0)'(F_1(x)) = \sum_k \frac{t^k}{k!} \mathrm{ad}^k F_0 \cdot F_1(x(t))$$

where $x(t) = (\exp tF_0)(x_0)$. Special variations of the reference zero control can be applied to generate the Lie brackets $\mathrm{ad}^k F_0 \cdot F_1(x(t))$. We present a computation based on the Baker-Campbell-Hausdorff formula.

Generalized Variations

To simplify, we restrict ourselves to the $\mathcal{C}^\omega$–real analytic case: $\dot{x} = F_0(x) + uF_1(x)$ where $\gamma(t) = (\exp tF_0)(x_0)$ is the reference singular trajectory associated to the control normalized to zero and defined on $[0, T]$. We assume that $|u| \leq 1$. A *positive rational polynomial* is a function of the form

$$\sum_{i=1}^{p} c_i t^{q_i}, \ c_i \geq 0, \ q_i \in \mathbf{Q}.$$

A vector W belongs to the generalized Pontryagin cone $\mathcal{E}^+$ if there exist positive rational polynomials $r_1, \sigma_1, \ldots, r_{2k}, \sigma_{2k}$ associated to a perturbation π such that:

$$\begin{aligned} \alpha_\pi(y, \varepsilon) &= \exp(\varepsilon W + o(\varepsilon))(y) \\ &= (\exp \sigma_{2k}(\varepsilon) F_0)(\exp r_{2k}(\varepsilon)(F_0 - F_1)) \\ &\quad\ (\exp \sigma_{2k-1}(\varepsilon) F_0)(\exp r_{2k-1}(\varepsilon)(F_0 + F_1)) \\ &\cdots (\exp \sigma_1(\varepsilon) F_0)(\exp r_1(\varepsilon)(F_0 + F_1)) \\ &\quad\ (\exp -(\Sigma_{i=1}^{2k} \sigma_k(\varepsilon) + r_k(\varepsilon)) F_0)(y) \end{aligned}$$

where $y = \exp TF_0(x)$. By construction, for $\varepsilon > 0$ small enough $\alpha_\pi(y, \varepsilon)$ is in $\mathcal{A}(x_0, T)$ and W is the right derivative of α at 0. Moreover, from the Baker-Campbell-Hausdorff formula, the derivative belongs to the Lie algebra generated by F_0 and F_1. As for the maximum principle, a crucial property is to have convexity. To prove this property we proceed as follows. Let $\pi_1 = (r_{i,1}, \sigma_{i,1})_i$ and $\pi_2 = (r_{i,2}, \sigma_{i,2})_i$ be two perturbations with respective tangent vectors W_1 and W_2. The composition of π_1 and π_2 is defined as:

$$(\alpha_{\pi_2}(\varepsilon))(\exp \mu(\varepsilon) F_0)(\alpha_{\pi_1}(\varepsilon))(\exp -\mu(\varepsilon) F_0)$$

where $\mu(\varepsilon) = \sum_i \sigma_{i,2}(\varepsilon) + r_{i,2}(\varepsilon)$. From the ad-formula, W_1 is a tangent vector to $(\exp \mu(\varepsilon) F_0)(\alpha_{\pi_1}(\varepsilon))(\exp -\mu(\varepsilon) F_0)$. Therefore, using the Baker-Campbell-Hausdorff formula $\exp X \exp Y = \exp(X + Y + \cdots)$ we have the lemma hereafter.

Lemma 1.4. *The sum $W_1 + W_2$ is the tangent vector corresponding to the composition of π_1 and π_2. In particular, $\mathcal{E}^+$ is a convex cone.*

Next, we prove the following additional result.

Lemma 1.5. *If $\pm W$ is in $\mathcal{E}^+$, then $\pm \mathrm{ad}^k F_0 \cdot W$ is in $\mathcal{E}^+$ as well for $k \geq 0$.*

Proof. We prove the result by recurrence on k. Let $\alpha_{\pi_\pm}(\varepsilon)$ be admissible variations with respective tangent vectors $\pm W$:

$$\alpha_{\pi_\pm}(\varepsilon) = \exp(\pm\varepsilon W + o(\varepsilon^p))$$

where $p > 1$ is a rational number. Let q in $\mathbf{Q}$ be such that $0 < q < 1$ and $pq > 1$, then:

$$\begin{aligned}&(\alpha_{\pi_+}(\varepsilon^q))(\exp \varepsilon^{1-q} F_0)(\alpha_{\pi_-}(\varepsilon^q))(\exp -\varepsilon^{1-q} F_0) = \\ &\quad (\exp(\varepsilon^q W + o(\varepsilon pq)))(\exp \varepsilon^{1-q} F_0)(\exp(-\varepsilon^q W + o(\varepsilon pq)))(\exp -\varepsilon^{1-q} F_0)\end{aligned}$$

which, because of the Baker-Campbell-Hausdorff formula, is equal to

$$\exp(W(\varepsilon^q - \varepsilon^q) + F_0(\varepsilon^{1-q} - \varepsilon^{1-q}) + \varepsilon[W, F_0] + o(\varepsilon)).$$

Thus, $[W, F_0]$ belongs to $\mathcal{E}^+$. ■

In particular, using the previous variations we can recover the conditions from the maximum principle. They concern only the linearized system. An important second-order condition is given by the result hereafter.

Proposition 1.5. *The Lie bracket $[F_1, [F_1, F_0]]$ belongs to $\mathcal{E}^+$.*

Proof. Applying the Baker-Campbell-Hausdorff formula, we get:

$$\begin{aligned}&(\exp \varepsilon^{1/3}(F_0 - F_1))(\exp 2\varepsilon^{1/3}(F_0 + F_1))(\exp \varepsilon^{1/3}(F_0 - F_1))(\exp -4\varepsilon^{1/3} F_0) \\ &\quad = \exp(2\varepsilon/3 \ \mathrm{ad}^2 F_1 \cdot F_0 - 2\varepsilon \mathrm{ad}^2 F_0 \cdot F_1 + o(\varepsilon)).\end{aligned}$$

Hence the vector $\frac{2}{3}\mathrm{ad}^2 F_1 \cdot F_0 - 2\mathrm{ad}^2 F_0 \cdot F_1$ belongs to $\mathcal{E}^+$. Since $\mathcal{E}^+$ is a convex cone containing $\pm \mathrm{ad}^2 F_0 \cdot F_1$, this proves the result. ■

As in the maximum principle, $\mathcal{E}^+$ provides an approximating cone of $\mathcal{A}_{x_0,T}$ and we obtain the following result.

Proposition 1.6. *Let x be a time-optimal trajectory defined on $[0, T]$ and associated to a control normalized to zero. Then, there exists p such that the extremal $z = (x, p)$ satisfies everywhere the conditions:*

(i) $\dot{z} = \overrightarrow{H}_0(z)$, *Hamiltonian system defined by* H_0.

(ii) $\mathrm{ad}^k H_0 \cdot H_1(z(t)) = 0, \ k \geq 0.$

(iii) $H_0(z(t)) \geq 0$.

(iv) $\{H_1, \{H_1, H_0\}\}(z(t)) \geq 0$.

Definition 1.5. *The condition (iv) is called the generalized Legendre-Clebsch condition.*

Application and Geometric Interpretation

Assume that the vector field F_1 is transverse to the trajectory. Then, we can find local coordinates in which $F_1 = \partial/\partial x_n$ so that the system is written

$$\dot{x}' = F(x', x_n),\ \dot{x}_n = F_{0,n}(x) + u$$

where $x' = (x_1, \ldots, x_{n-1})$. The system in x' where x_n is taken as the new control variable is called the *reduced system*. Let $H' = \langle p', F(x', x_n)\rangle$ be the corresponding reduced Hamiltonian, $p' = (p_1, \ldots, p_{n-1})$. A straightforward computation gives

$$\frac{d}{dt}\frac{\partial H}{\partial u}(z,u) = \{H_1, H_0\}(z) = -\frac{\partial H'}{\partial x_n}(z', x_n)$$
$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u}(z,u) = \{H_1, \{H_1, H_0\}\}(z) = -\frac{\partial^2 H'}{\partial x_n^2}(z', x_n)$$

along an extremal curve: the generalized Legendre-Clebsch condition is the Legendre-Clebsch condition for the reduced system.

Multi-Input Case, Goh Condition

Similarly, higher order variations can be applied in the multi-input case to obtain further necessary conditions. The most important are the so-called *Goh conditions* that we present now. Consider a system of the form

$$\dot{x} = F_0(x) + \sum_{i=1}^{m} u_i F_i(x).$$

If $z = (x, p, u)$ is a reference singular extremal defined on $[0, T]$ then, in order to be time-optimal, the following condition has to be satisfied:

$$\{H_v, H_w\}(z(t)) = 0,\ t \in [0, T]$$

for every pair of vector fields F_v and F_w in $\mathrm{Span}\{F_1, \ldots, F_m\} \cdots$

1.3.2 Intrinsic Second-order Derivative and Conjugate Times

In the previous section we have generated special variations to obtain further necessary conditions for affine control systems. They concern Lie brackets of the form $[F_1,[F_1,F_0]]$ (generalized Legendre-Clebsch condition), or $[v,w]$ with v,w in $\mathrm{Span}\{F_1,\dots,F_m\}$ (Goh condition). These brackets are related to the second-order derivative and to the necessary conditions for *small time* optimality. We introduce now a different concept related to the loss of optimality because of the *cumulated effect of time*. It is the concept of *conjugate time* associated to the spectral properties of the intrinsic second-order derivative, and to the notion of Lagrangian manifolds in symplectic geometry. We begin by presenting these geometric tools.

Symplectic Geometry and Lagrangian Manifolds

Linear symplectic manifolds and symplectic group. We recall some standard facts about symplectic geometry. Let (V,ω) be a linear symplectic space of dimension $2n$. We can choose a basis called *Darboux* or *canonical linear coordinates* such that $V \simeq \mathbf{R}^{2n}$ and $\omega(x,y) = {}^t x J y$ where

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}. \tag{1.13}$$

A subspace L of V is called *isotropic* if $\omega|_L = 0$. An isotropic of maximal dimension n is called a *Lagrangian subspace*. Linear isomorphisms preserving ω are called *symplectomorphisms* and, in Darboux coordinates, they are identified with the elements of the *symplectic group* $\mathrm{Sp}(n,\mathbf{R})$ of matrices S satisfying ${}^tSJS = J$. Decomposing S into $n \times n$ blocks,

$$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

we obtain the relations:

$${}^tAD = {}^tBC = I,\ \ {}^tAC = {}^tCA,\ \ {}^tBD = {}^tDB.$$

The Lie algebra $\mathfrak{sp}(n,\mathbf{R})$ of $\mathrm{Sp}(n,\mathbf{R})$ is the algebra of order $2n$ matrices H such that $\exp tH$ is in $\mathrm{Sp}(n,\mathbf{R})$. These matrices are characterized by ${}^tHJ + H\,{}^tJ = 0$ and, decomposing H into blocks,

$$H = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

we obtain the equivalent definition:

$$\mathfrak{sp}(n,\mathbf{R}) = \{H = \begin{bmatrix} A & B \\ C & -{}^tA \end{bmatrix},\ B \text{ and } C \text{ symmetric}\}\cdots$$

The symplectic group acts on Lagrangian subspaces and we have the following representation of Lagrangian subspaces. Let L be a Lagrangian subspace, and let $\Pi : (x,p) \mapsto x$ be the canonical projection written in Darboux coordinates. If the restriction to L of Π is regular, L can be represented as

$$\begin{bmatrix} x \\ Cx \end{bmatrix}$$

that is as the image of $\{x\}$ by the $2n \times n$ matrix

$$\begin{bmatrix} I \\ C \end{bmatrix}$$

where C is symmetric. More generally, let L be a Lagrangian subspace represented by the $2n \times n$ matrix

$$\begin{bmatrix} A \\ B \end{bmatrix}.$$

Then, from the definition, one must have ${}^tAB - {}^tBA = 0$ and the matrix

$$\begin{bmatrix} A & -B \\ B & A \end{bmatrix}$$

is symplectic. In particular, the symplectic group acts transitively on the Lagrangian subspaces.

Symplectic and Lagrangian manifolds on the cotangent bundle. On the cotangent bundle T^*M of any smooth manifold M exists a canonical symplectic structure associated with the *Liouville form* written in coordinates as $\alpha = pdx$, where x are coordinates on M and p the dual ones. The symplectic form is defined by $\omega = d\alpha = dp \wedge dx$. We denote by Π the standard projection, $\Pi : (x,p) \in T^*M \mapsto x \in M$. Locally, we can identify M with $\mathbf{R}^n$, but globally, an important topological invariant is the space H^1 which is the quotient of the space of closed 1-forms by the space of exact 1-forms. If L is a regular submanifold of (T^*M, ω), it is called isotropic (*resp.* Lagrangian) if at each point the tangent space is isotropic (*resp.* Lagrangian). A canonical example in $\mathbf{R}^{2n}$ is constructed as follows. Let $S : x \mapsto S(x)$ be a smooth function in $\mathbf{R}^n$ and consider the graph $L = \{p = \partial S/\partial x\}$: L is a Lagrangian manifold and the projection $\Pi : L \to \mathbf{R}^n$ is regular. We generalize now the representation result of Lagrangian manifolds obtained in the linear case.

Proposition 1.7. *Let L be a Lagrangian manifold of (T^*M, ω). Then, locally, there are Darboux coordinates (x,p) together with a smooth function S of $(x_I, p_{\bar I})$ with $I = \{1, \ldots, m\}$ and $\bar I = \{m+1, \ldots, n\}$ such that L is defined by the equations*

$$p_I = \frac{\partial S}{\partial x_I},\ x_{\bar I} = -\frac{\partial S}{\partial p_{\bar I}}\cdot$$

The mapping S is called the generating mapping of L.

Definition 1.6. *Let L be a Lagrangian submanifold of (T^*M, ω). A nonzero vector v, tangent to L at x, is called* vertical *whenever $d\Pi(x)v = 0$. The* caustic *of L is the set of points at which there exists at least one vertical tangent vector.*

Example 1.1. For any x in M, the fiber $L = T_x^*M$ is a linear Lagrangian submanifold and all tangent vectors are vertical. More generally, if M_0 is a regular submanifold of M, the submanifold $M_0^\perp$ defined by the transversality relation

$$M_0^\perp = \{(x, p) \in T^*M \mid p \perp T_x M\}$$

is a Lagrangian submanifold of M.

Hamiltonian vector fields and variational equation. In order to simplify our presentation, we use local coordinates, identifying locally M to $\mathbf{R}^n$, T^*M to $\mathbf{R}^{2n}$, and ω to the standard 2-form $dp \wedge dx$. Hence, any time dependent Hamiltonian vector field is defined by the equations

$$\dot{x} = \frac{\partial H}{\partial p}(t, z), \; \dot{p} = -\frac{\partial H}{\partial x}(t, z)$$

where $z = (x, p)$ and $H(t, z)$ is the Hamiltonian. Using J as defined by (1.13), the previous equation can be written in the compact form

$$\dot{z} = J\nabla_z H(t, z) \tag{1.14}$$

where ∇_z stands for the gradient with respect to z. When the Hamiltonian is a quadratic form

$$H(t, z) = \frac{1}{2}\,{}^t z S(t) z$$

with $S(t)$ symmetric, we get a linear Hamiltonian system

$$\dot{z} = JS(t)z = A(t)z$$

where $A(t)$ is a Hamiltonian matrix element of $\mathfrak{sp}(n, \mathbf{R})$. In order to make our geometric analysis, an important issue is the action of the group of symplectic transformations on Hamiltonian vector fields. Let $\dot{z} = J\nabla_z H(t, z)$ be a Hamiltonian vector field and consider a change of variables $z \mapsto \xi = \Phi(t, z)$. The transformation is *symplectic* if $\partial\Phi(t, z)/\partial z$ belongs to the symplectic group. Computing, one has

$$\dot{\xi} = \frac{\partial \Phi}{\partial t}(t, z) + \frac{\partial \Phi}{\partial z}(t, z)\dot{z}.$$

Since the transformation is symplectic,

$$\frac{\partial \Phi}{\partial z}(t, z) J\nabla_z H(t, z) = J\nabla_\xi \hat{H}(t, \xi)$$

where $\hat{H}(t,\xi) = H(t,z)$. Using Poincaré lemma, we can write locally

$$\frac{\partial \Phi}{\partial t}(t,z) = J\nabla_\xi R(t,\xi)$$

where R is called the *remainder function*: we have showed that any symplectic change of coordinates transforms a Hamiltonian vector field into another Hamiltonian vector field. If $z(t,t_0,z_0)$ is the solution of (1.14) starting from z_0 at t_0, then the flow $z_0 \mapsto z(t,t_0,z_0)$ is symplectic for any fixed t, t_0. Differentiating with respect to z we define the variational equation

$$\delta \dot{z} = J\nabla^2_{z^2} H(t, z(t,t_0,z_0))\delta z.$$

Symplectomorphisms induce time dependent linear symplectic isomorphisms on the corresponding variational equation. The action of the linear symplectic group on linear Hamiltonian differential equations is a standard action and numerous tensor analysis exist in the litterature. For instance, a standard result is the following.

Proposition 1.8. *Let $\dot{x} = A(t)$ be a Hamiltonian differential equation on $\mathbf{R}^{2n}$ and let $z_1, \ldots, z_n$ be n independent solutions such that $\omega(z_i, z_j) = 0$. Then, a complete set of solutions can be computed by quadrature.*

Proof. Let L be the $2n \times n$ matrix whose columns are the independent solutions. By construction, it is a one parameter Lagrangian manifold and we have

$$\dot{L} = A(t)L, \ \ {}^tL(t)JL(t) = 0.$$

Since the solution are independent, the matrix ${}^tLL$ is a non singular $n \times n$ matrix. Define the $2n \times n$ matrix $L' = JL({}^tLL)^{-1}$. Hence, ${}^tL'JL = 0$ and ${}^tLJL' = -I$. Therefore, $P = (L', L)$ is a symplectic matrix and we have

$$P^{-1} = \begin{bmatrix} -{}^tLJ \\ {}^tL'J \end{bmatrix}.$$

If we make the symplectic change of coordinates $x = Py$, we get the Hamiltonian equation

$$\dot{y} = P^{-1}(AP - \dot{P})y$$

and using the notation $\dot{x} = Ax = JS$ where S is symmetric. Decomposing $y = (u,v)$, we obtain the equations

$$\begin{aligned} \dot{u} &= 0 \\ \dot{v} &= -{}^tL'(SL' + J\dot{L}')u \end{aligned}$$

and the solution can be computed by quadrature. ∎

Similar tensor analysis can be developed to study the standard LQ problem.

Geometric analysis of linear quadratic problems. Consider the smooth linear system in $\mathbf{R}^n$ $\dot{x} = A(t) + B(t)u$ and the problem of minimizing a cost defined by

$$c(x,u) = \int_0^T ({}^t x W(t)x + {}^t u U(t)u)dt$$

with fixed time $T > 0$ and prescribed boundary conditions. The symmetric matrices $W(t)$ and $U(t)$ are smooth with respect to t and we assume that the strict Legendre-Clebsch condition holds for all t: $U(t) > 0$. By applying a proper feedback we can renormalize $U(t)$ to I. If we apply the maximum principle, the optimal solutions have to be found among the following extremals:

$$\dot{x} = A(t)x + B(t)U^{-1}(t)\,{}^t B(t)p \quad (1.15)$$

$$\dot{p} = {}^t W(t)x - {}^t A(t)p \quad (1.16)$$

which can be rewritten $\dot{z} = Hz$ with

$$H = \begin{bmatrix} A & C \\ D & -{}^t A \end{bmatrix}$$

and $C = B\,{}^t B$ ($U(t)$ being identified with I). We can assume B of full rank so that C is definite positive. In order to identify a *curvature*-like invariant connected to the optimality properties of the reference solution, a standard reduction is to write the equation as a second-order differential equation. Using the first equation, we write

$$p = C^{-1}(\dot{x} - Ax)$$

which, plugged into the second equation, gives after a left product by C,

$$\ddot{x} + \widetilde{A}\dot{x} + \widetilde{B}x = 0.$$

By setting $x(t) = S(t)X(t)$ where $S(t)$ is properly chosen it can be written

$$\ddot{X} + K(t)X = 0.$$

The matrix $K(t)$ is the curvature invariant of the problem, related to the distribution of conjugate points to be defined later. It is an invariant of the action of the symplectic subgroup of matrices of the form

$$P(t) = \begin{bmatrix} A(t) & 0 \\ B(t) & C(t) \end{bmatrix}$$

which preserves in fact the subspace δx because we must keep track of the state space. By counting the respective dimensions, a normal form contains $n(n+1)/2$ parameters which correspond to the symmetric tensor identified in our reduction. Another useful representation which will be used later is the

Riccati equation. Let Φ be the fundamental matrix solution of (1.15)-(1.16). Decomposing $\Phi(t)$ into $n \times n$ blocks

$$\Phi(t) = \begin{bmatrix} \Phi_1(t) & \Phi_3(t) \\ \Phi_2(t) & \Phi_4(t) \end{bmatrix}$$

we define the one parameter family of Lagrangian subspaces

$$L(t) = \begin{bmatrix} \Phi_3(t) \\ \Phi_4(t) \end{bmatrix}.$$

The projection $\Pi \ : \ (x,p) \mapsto x$ restricted to L is regular if and only if the matrix $\Phi_3(t)$ is invertible. We have

$$\begin{bmatrix} \dot{\Phi}_3 \\ \dot{\Phi}_4 \end{bmatrix} = \begin{bmatrix} A & B\,{}^tB \\ W & -{}^tA \end{bmatrix} \begin{bmatrix} \Phi_3 \\ \Phi_4 \end{bmatrix}.$$

In the regular case, we introduce $R(t) = \Phi_4(t)\Phi_3^{-1}(t)$ which satisfies the symmetric Riccati equation

$$\dot{R} = W - {}^tAR - RA - RB\,{}^tBR$$

whose solution is symmetric whenever $R(0)$ is symmetric.

Symplectic transformation and generating function. Let φ be a symplectomorphism. Let us prove that, locally, φ is parameterized by a *generating function*. We proceed as follows. Since the result is local, we identify the symplectic space with $\mathbf{R}^{2n}$. Let $\varphi \ : \ (x,p) \mapsto (X,P)$ be a symplectic change of coordinates. Then, the 1-form $\sigma_1 = xdp - XdP$ is closed. Assume that (p,P) define coordinates then, locally, there is a function $S_1(p,P)$ such that $\sigma_1 = dS_1$ and we get the relation

$$x = \frac{\partial S_1}{\partial p}, \ X = -\frac{\partial S_1}{\partial P}$$

which defines locally the change of coordinates. We proceed similarly with the 1-forms

$$\begin{aligned} \sigma_2 &= xdp + PdX \\ \sigma_3 &= pdx - PdX \\ \sigma_4 &= pdx + XdP \end{aligned}$$

to which we associate the generating mappings S_2, S_3, S_4. In particular, each diffeomorphism $X = \varphi(x)$ can be lifted onto a symplectomorphism $\overrightarrow{\varphi}$ given by

$$X = \varphi(x), \ p = \frac{{}^t\partial\varphi}{\partial x}(x)P$$

and defined by the generated mapping $S_4(x,P) = {}^t\varphi(x)P$. The next step is to define the geometric concept of conjugate point.

Definition 1.7. *Let $\overrightarrow{H}(t,z)$ be a smooth Hamiltonian vector field whose integral curves are the extremals of an optimal control problem with fixed time T. Let $z=(x,p)$ be a reference extremal. Then the variational equation*

$$\delta\dot{z} = \frac{\partial\overrightarrow{H}}{\partial z}(t,z(t))\delta z$$

is called the Jacobi equation. *A* Jacobi field $J=\delta z$ *is a non-trivial solution of this equation. In accordance with the Lagrangian terminology (see def. 1.6), it is called* vertical *at time t if $\delta x(t)=0$, that is if $d\Pi(z(t))J(t)=0$. The time t_c is called* conjugate *if there exists a Jacobi field vertical both at $t=0$ and t_c. In this case, $x(t_c)$ is said to be conjugate to $x(0)$ along the reference solution.*

Definition 1.8. *If $z(t,t_0,z_0)$ is the integral curve of $\overrightarrow{H}(t,z)$ with initial condition z_0 at $t=0$, the* exponential mapping *at t is defined by*

$$\exp_{x_0,t}\ :\ p_0 \mapsto \Pi(z(t,x_0,p_0)).$$

The following result is a consequence of the previous analysis.

Proposition 1.9. *Let z be a reference extremal with initial condition $z_0=(x_0,p_0)$ defined on $[0,T]$. Let L_0 be the fiber $T^*_{x_0}M$ and let L be its image by the one parameter group* $\exp t\mathbf{H}$. *Then L is a one parameter family of Lagrangian submanifolds along the reference extremal curve and t_c is conjugate if and only if (L,Π) is singular at t_c, that is if p_0 is a singular point of the exponential mapping at time t_c.*

The generalization to control problems with arbitrary initial conditions is straightforward.

Definition 1.9. *Let $\overrightarrow{H}(t,z)$ be a smooth Hamiltonian vector field whose integral curves are the extremals of an optimal control problem with fixed time T and initial manifold M_0. The time t_f is a* focal time *along the BC-extremal z if there is a Jacobi field J such that $J(0)$ is in $T_{z(0)}M_0^\perp$ and J is vertical at t_f.*

Both concepts fit in the same geometric framework: a one parameter family of Lagrangian manifolds obtained by transporting the initial submanifold with the flow. The Jacobi fields span the tangent spaces of the Lagrangian manifolds computed along the reference extremal. They are the image of the initial tangent space by the fundamental matrix of the variational equation and conjugate or focal points are obtained using a verticality test. Curvature type invariants are related to tensor analysis of each problem. The analysis in the next paragraph shows the connection between the concept of conjugate point and the intrinsic second-order derivative. We derive $\mathcal{C}^1$-sufficient second order optimality conditions in the smooth case.

Conjugate Points of Smooth Time-optimal Control Problems

Preliminaries. We restrict our presentation to a smooth time optimal control problem $\dot{x} = f(x,u)$ where u belongs to U, assumed to be an open subset of $\mathbf{R}^m$. The Hamiltonian of the problem is

$$\widetilde{H}(x,p,u) = p^0 + H(x,p,u)$$

with $H(x,p,u) = \langle p, f(x,u)\rangle$ and $p^0 \leq 0$. The scalar p^0, dual to the cost functional $c(x,u) = 1$, can be normalized to 0 or to 1. The case $p^0 = 1$ is called the *normal case*. The maximum principle asserts that time-optimal solutions satisfy $\partial H/\partial u = 0$ and $\partial^2 H/\partial u^2 \leq 0$. In the so-called regular case, the strict Legendre-Clebsch condition holds and $\partial H/\partial u = 0$ is solved by the implicit function theorem. By plugging the dynamic feedback $\hat{u} : (x,p) \mapsto \hat{u}(x,p)$ into H, a true Hamiltonian function H_r is defined:

$$H_r(x,p) = H(x,p,\hat{u}(x,p)).$$

As usual, $t \mapsto z(t,z_0)$ is the extremal solution with initial condition $z_0 = (x_0,p_0)$. Since H is linear in p, we have a first lemma.

Lemma 1.6. *The two components of an extremal solution verify*

$$x(t,x_0,\lambda p_0) = x(t,x_0,p_0),\ \ p(t,x_0,\lambda p_0) = \lambda p(t,x_0,p_0).$$

In particular, the rank of the exponential mapping $\exp_{x_0,t}$ *at a given time* t *is at most* $(n-1)$.

The aim of this section is twofold. First, thanks to the concept of conjugate point, we obtain second-order necessary and sufficient conditions for optimality, the set of controls being endowed with the L^∞ topology. Then, using standard field theory, we extend those optimality results to the $\mathcal{C}^0$ topology on the set of trajectories of the system. In order to carry out a more complete analysis applicable to affine systems, we make the following prolongation. We set $\dot{u} = v$ and extend the original system to a control affine one:

$$\begin{aligned}\dot{x} &= f(x,u)\\ \dot{u} &= v.\end{aligned}$$

If we write the system $\dot{y} = F_0(y) + \sum_{i=1}^m v_i F_i(y)$ with $y = (x,u)$, the controlled distribution is flat: $[F_i,F_j] = 0$, $i,j = 1,\ldots,m$. Our analysis also applies to control affine systems whose distribution is involutive. A prototype of such systems is the single input control system of the form: $\dot{y} = F_0(y) + vF_1(y)$. Having made our prolongation, we must change the L^∞ control topology on u into the L^1 topology on $v = \dot{u}$. According to Section 1.3.1, there is a one-to-one correspondance between the extremal solutions of the original system and the affine system obtained by prolongation. As a consequence, we shall be able to translate the relevant optimality results.

Second order sufficient optimality conditions for single input affine systems. We consider a single input affine control system

$$\dot{x} = F_0(x) + uF_1(x)$$

and we assume that the control domain is $U = \mathbf{R}$. The controlled vector field F_1 is called the *cheap direction* and time-optimal curves are to be searched among concatenation of standard extremals with jumps into this cheap direction. We compute a normal form under the action of the *feedback group*. The group acts locally with the following transformations:

(i) Change of coordinates, $y = \varphi(x)$,

$$(F_0, F_1) \mapsto (\varphi_* F_0, \varphi_* F_1).$$

(ii) Feedback transformation, $v = \alpha(x) + \beta(x)u$ where β is invertible,

$$(F_0, F_1) \mapsto (F_0 + \alpha F_1, \beta F_1).$$

The following result is standard.

Proposition 1.10. *The singularities of the endpoint mapping corresponding to extremals curves of the time-optimal control problem are feedback invariant.*

Hence, we shall use the action of the feedback group to normalize our system along a reference extremal, each change of coordinates $y = \varphi(x)$ being lifted onto a symplectic diffeomorphism $\overrightarrow{\varphi}$ acting on the extremal flow.

Geometric reduction. We proceed in two steps. We first pick a reference smooth extremal trajectory γ defined on $[0, T]$. Assuming it is one-to-one, we can identify it with $t \mapsto (t, 0, \dots, 0)$ in suitable coordinates $x = (x_1, \dots, x_n)$. A tubular neighbourhood of γ is characterized by small x_i's for $i \geq 2$. Then we consider the Taylor expansion of the pair F_0, F_1 along γ: the *jet* of order one (*resp.* two) is the collection of all linear (*resp.* quadratic) terms. The control is also normalized to zero thanks to the feedback $v = u - u(x_1)$ (see Section 1.3.1). Besides, if F_1 is tranverse to γ, we can choose the coordinates in the neighbourhood of the curve such that F_1 is identified with $\partial/\partial x_n$. From our preliminary analysis, we know that the first order Fréchet derivative of the endpoint mapping depends only upon the jet of order one, while the second-order intrinsic derivative depends only upon the jet of order two. Furthermore, all the information about first and second variations is collected by Lie brackets within the two spaces $E_1(t) = \mathrm{Span}\{\mathrm{ad}^k F_0 \cdot F_1(\gamma(t))\}$ and $E_2(t)$ which is generated by the restriction to γ of Lie brackets with at most two occurences of F_1. The second normalization is performed choosing a reference extremal meeting the generic requirements hereafter:

(i) $E_1(t)$ is of codimension one and is generated by the first $(n-1)$ brackets, $\mathrm{ad}^k F_0 \cdot F_1(\gamma(t))$, $k = 0, \ldots, n-2$, for any t in $[0, T]$.

(ii) The Lie bracket $\mathrm{ad}^2 F_1 \cdot F_0(\gamma(t))$ is not contained in $E_1(t)$ for t in $[0, T]$.

(iii)The vector field F_0 restricted to γ is tranverse to $E_1(t)$ on $[0, T]$.

This has the following implications: first, for each $0 < t_0 < t_1 \leq T$, the singularity of the endpoint mapping at the zero control defined on $[t_0, t_1]$ is of codimension one and the image of its Fréchet derivative is $E_1(t_1)$. Secondly, the adjoint covector p is unique up to a scalar and oriented in order that $H_0 = \langle p, F_0 \rangle$ be positive. The singular trajectory which is of minimal order by virtue of requirement (ii), is said to be *hyperbolic* if $\langle p, \mathrm{ad}^2 F_1 \cdot F_0(\gamma(t)) \rangle < 0$ on $[0, T]$, *elliptic* if $\langle p, \mathrm{ad}^2 F_1 \cdot F_0(\gamma(t)) \rangle > 0$. Observe that the generalized Legendre-Clebsch condition is only satisfied in the hyperbolic case. It is now crucial to notice that since the reference curve is a one-dimensional manifold, we can normalize any independent family of Lie brackets to form a frame along it. Our assumptions allow us to pick coordinates preserving the previous normalizations and defining a moving frame defined by:

$$\mathrm{ad}^k F_0 \cdot F_1(\gamma(t)) = \frac{\partial}{\partial x_{n-k}}, \; k = 0, \ldots, n-1, \; t \in [0, T].$$

Moreover, since the feedback is chosen so that u is zero along γ, we can impose the linearization condition $\mathrm{ad}^k F_0 \cdot F_1(\gamma(t)) = 0$ for $k > n-2$ and $t \in [0, T]$. These computations can be explicited. In particular, the moving frame construction amounts to a time dependent linear transformation. Having made these normalizations, we have the following.

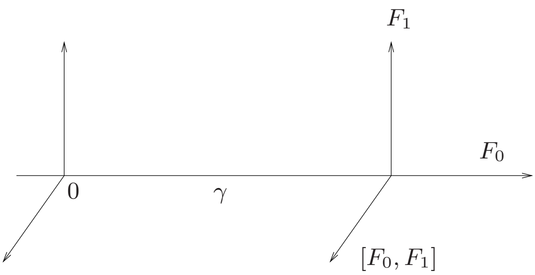

Fig. 1.4. Canonical moving frame

Proposition 1.11. *Along the reference curve, the system is feedback equivalent to the system defined by the two vector fields*

$$F_0 = \frac{\partial}{\partial x_1} + \sum_{i=2}^{n-2} x_{i+1}\frac{\partial}{\partial x_i} + \sum_{i,j=2}^{n} a_{ij}(x_1)x_i x_j \frac{\partial}{\partial x_1} + R$$
$$F_1 = \frac{\partial}{\partial x_n}$$

where the remainder $R = \sum_{i=1}^{n} R_i \frac{\partial}{\partial x_i}$ *is such that the* 1*-jets of the* R_i*'s along* γ *are zero,* $i = 1, \dots, n$*, as well as the* 2*-jets for* $i \geq 2$*.*

Definition 1.10. *The truncated system*

$$F_0 = \frac{\partial}{\partial x_1} + \sum_{i=2}^{n-2} x_{i+1}\frac{\partial}{\partial x_i} + \sum_{i,j=2}^{n} a_{ij}(x_1)x_i x_j \frac{\partial}{\partial x_1}$$
$$F_1 = \frac{\partial}{\partial x_n}$$

is called the approximating model *along* γ*.*

Properties of the model. In this model, we have gathered in one normal form all the information required to evaluate the endpoint mapping (and thus the accessibility set) up to second-order relevant terms. The adjoint covector is oriented by the condition $H_0 \geq 0$ and normalized to $p = (1, 0, \dots, 0)$ The linearized system along the reference trajectory is a constant linear system in Brunovsky normal form. Indeed,

$$\begin{aligned}
\dot{x}_1 &= 1 + q(x_1, x_2, \dots, x_n) \\
\dot{x}_2 &= x_3 \\
&\vdots \\
\dot{x}_n &= u
\end{aligned}$$

with $q(x_1, x_2, \dots, x_n) = \sum_{i,j=2}^{n} a_{ij} x_i x_j$. Setting $x(t) = t + \xi(t)$ we get

$$\begin{aligned}
\dot{\xi}_1 &= \sum_{i,j=2}^{n} a_{ij} x_i x_j \\
\dot{\xi}_2 &= \xi_3 \\
&\vdots \\
\dot{\xi}_n &= u
\end{aligned}$$

and the system describing the evolution of ξ is the linearized system. We substitute x_1 by t in the quadratic form q. The last diagonal coefficient $a_{nn}(t)$ is

$\langle p, \partial^2 F_0/\partial x_n^2\rangle(\gamma(t))$, which is also equal to the opposite of the Poisson bracket $\{\{H_1, H_0\}, H_1\}(z(t))$ involved in the generalized Legendre-Clebsh condition: it is negative in the hyperbolic case, and positive in the elliptic one. The kernel K_t of the first order derivative is $\dot\xi_2 = \xi_3, \ldots, \dot\xi_{n-1} = \xi_n$ with the boundary conditions $\xi_2(0) = \cdots = \xi_n(0) = 0$, $\xi_2(t) = \cdots = \xi_n(t) = 0$, and the quadratic form q represents in fact the intrinsic second-order derivative defined on $[0, T]$ by the restriction to K_t of

$$Q_t(\xi) = \int_0^t \sum_{i,j=2}^{n} a_{ij}(s)\xi_i(s)\xi_j(s)ds.$$

By construction, our affine system is the prolongation of a regular system in $\mathbf{R}^{n-1}$ where the control variable is $x_n = \xi_n$.

Accessory problem and intrinsic derivative. By taking ξ_n as control variable and approximating by the model, clearly, the reference extremal curve is time-optimal on $[0, T]$ if and only if Q_t is negative for each t in $]0, T]$. This leads to consider the so-called *accessory problem*, $\varepsilon Q_t \to \min$, with $\varepsilon = -1$ in the hyperbolic case, and $\varepsilon = 1$ in the elliptic one. This is a standard problem in differential operator theory. We can rewrite the intrinsic second order derivative as

$$Q_t = \int_0^T q(y(s))ds$$

with $y = \xi_2$ and where

$$q(y(t)) = \sum_{i,j=0}^{n-2} b_{ij}(t)y^{(i)}(t)y^{(j)}(t)$$

the b_{ij} being symmetric functions. The boundary conditions on $[0, T]$ define the set $\mathcal{C}_t$ of smooth curves such that $y(0) = \cdots = y^{(n-3)}(0) = 0$, $y(t) = \cdots = y^{(n-3)}(t) = 0$. Let D be the differential operator of order $2(n-2)$ defined by

$$Dy = \frac{1}{2}\sum_{i=0}^{n-2}(-1)^i \frac{d^i}{dt^i}\frac{\partial q}{\partial y^{(i)}}(y).$$

It is the *Euler-Lagrange operator* associated to the accessory minimization problem and it can be written

$$Dy = \sum_{i,j=0}^{n-2}(-1)^j \frac{d^j}{dt^j} b_{ij}(t)\frac{d^i}{dt^i}\cdot$$

Its restriction D_t to $\mathcal{C}_t$ is a self-adjoint differential operator representing the second-order intrinsic derivative. This operator is regular since $b_{n-2,n-2}(t) = \{\{H_1, H_0\}, H_1\}(\gamma(t))$ is nonzero. The following result holds.

Lemma 1.7. *The equation $Dy = 0$ is equivalent to Jacobi equation along the reference extremal and is Euler-Lagrange equation associated to the accessory problem. If J is a Jacobi field solution of the variational equation, then J is vertical at 0 and t_c if and only if $D_{t_c}J = 0$ so that*

$$Q_{t_c}(J) = \int_0^{t_c} D_{t_c}J(s) \cdot J(s)ds = 0.$$

The spectral properties of Q_t are investigated using the classical theory on linear differential operators, see [17], and we get the next proposition.

Proposition 1.12. *For each t in $]0, T]$, there exists a sequence of eigenvectors and eigenvalues $(e_{t,\alpha}, \lambda_{t,\alpha})_{\alpha \geq 1}$ such that*

(i) The eigenvectors $e_{t,\alpha}$ belong to $\mathrm{L}^2([0,T]) \cap \mathcal{C}_t$ and $D_t e_{t,\alpha} = \lambda_{t,\alpha} e_{t,\alpha}$.

(ii) Each curve y in $\mathcal{C}_t$ can be represented by its uniformly convergent Fourier series,

$$y = \sum_{\alpha \geq 1} y_\alpha e_{t,\alpha}.$$

We order the eigenvalues increasingly, $\lambda_{t,1} \leq \lambda_{t,2} \leq \ldots$, and state the Morse result about the time evolution of the spectrum of D_t.

Proposition 1.13. *Let y be in $\mathcal{C}_t$ with Fourier series $y = \sum_{\alpha \geq 1} y_\alpha e_{t,\alpha}$. Then $Q_t(y) = \sum_{\alpha \geq 1} \lambda_{t,\alpha} y_\alpha^2$ and Q_t is positive for t small enough. The first conjugate time to 0, t_{1c}, is the smallest t such that $\lambda_{t,1} = 0$. If $t < t_{1c}$, the only minimizer of Q_t on $\mathcal{C}_t$ is $y = 0$. If $t > t_{1c}$, the infimum of Q_t is $-\infty$.*

Proof. Rather than using the standard Morse theory, we make a simple proof of the loss of optimality after the first conjugate time based on the geometric argument of the Riemannian case [7]. Indeed, let t_{1c} be the first conjugate time along the reference trajectory γ. There exists a Jacobi field vertical at 0 and t_{1c} corresponding to a variation of γ with $\delta x(0) = \delta x(t_{1c}) = 0$. Then, for $t > t_{1c}$ we can construct a broken solution with the same time duration (see Fig. 1.5). But in our regular case, an optimal solution cannot be broken. In fact, by smoothing the corner we obtain a shortest path. Since the model approximates our system up to relevant terms of order two, we conclude that optimality is lost. ■

Proposition 1.14. *Consider a single input affine control system defined by the pair F_0, F_1. Under our assumptions, a reference trajectory is time minimal (*resp. *maximal) in the hyperbolic (*resp. *elliptic) case up to the first conjugate time with respect to all trajectories with the same extremities and contained in a tubular $\mathcal{C}^1$-neighbourhood of the reference extremal.*

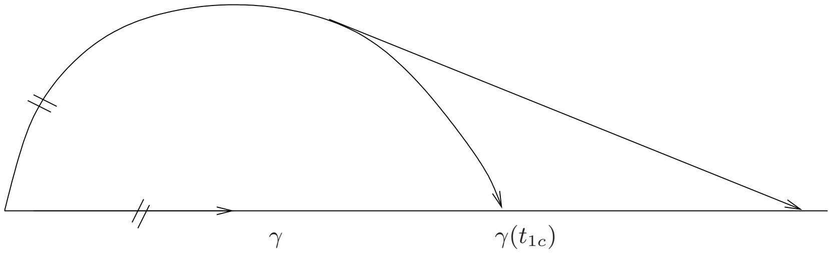

Fig. 1.5. Broken solution with same time duration and shortest path

The same optimality result holds for the restricted system, the set of controls being endowed with the L^∞-norm topology.

Computation and estimation of conjugate points. The simplest test deals with the nonlinear system $\dot{x} = f(x, u)$. We denote by $H_r(x, p)$ the smooth maximized Hamiltonian function and we restrict our equation to the level set $H_r = 1$ so as to break the symmetry due to the linearity with respect to p. This amounts to assume p in the projective space $\mathbf{P}^{n-1}$. We note $J_1, \ldots, J_{n-1}$ a basis of Jacobi fields which are vertical at 0. Let $L(t) = \mathrm{Span}\{J_1(t), \ldots, J_{n-1}(t)\}$ be the corresponding isotropic space. The numerical test for conjugate times is

$$\mathrm{rank}\ d\Pi(z(t))L(t) < n - 1 \tag{1.17}$$

which can easily be tested numerically. Moreover, since the reference trajectory is tranverse, the test is equivalent to

$$\det(d\Pi(z(t))L(t), f(x(t), u(t))) = 0. \tag{1.18}$$

In order to estimate the conjugate points we can use the curvature tensor according to the Sturm comparison theorem.

Lemma 1.8. *(Sturm) Let v be the solution of $\ddot{v} + A(t)v = 0$ with $v(0) = 0$, $\dot{v}(0) = 1$, and let w be the solution of $\ddot{w} + B(t)w = 0$ with the same initial conditions. Suppose $A(t) \leq B(t)$. If a (*resp. *b) is the first positive zero of v (*resp. *w), then $a \leq b$*

Proof. In accordance with initial conditions, v and w are positive for $0 < t < a$ and $0 < t < b$, respectively. Assume by contradiction that $a > b$. One has

$$0 = \int_0^b (v(\ddot{w} + Bw) - w(\ddot{v} + Av))dt \tag{1.19}$$

$$= [v\dot{w} - w\dot{v}]_0^b + \int_0^b (B - A)vwdt. \tag{1.20}$$

Since $A(t) \geq B(t)$ and since v and w are positive on $[0, b]$, then the integral in (1.20) is non-positive. Therefore, $[v\dot{w} - w\dot{v}]_0^b$ is nonnegative, that is $v(b)\dot{w}(b) \geq 0$. But $v(b) > 0$ and $\dot{w}(b) < 0$ (since w is not identically zero, w and $\dot{w}$ cannot be both zero at b), hence the contradiction. ■

Corollary 1.1. *Let $\delta\ddot{x} + K(t)\delta x = 0$ be the one-dimensional Jacobi equation in normal form. Assume $0 < K_1 \leq K(t) \leq K_2$. If t_{1c} is the first conjugate time, then t_{1c} belongs to $[\pi/\sqrt{K_2}, \pi/\sqrt{K_1}\,]$.*

In higher dimension, this result can be applied using the *sectional curvature*. In our problem, the Jacobi equation is identified with a differential operator and we can use a different normal form with less invariant terms than in the curvature because, while evaluating the intrinsic second-order derivative, we have reduced the terms by integrating by parts. The normal form is given by the proposition hereafter.

Proposition 1.15. *Any self-adjoint differential operator l with real coefficients is of even order and can be written according to*

$$l(y) = (p_0 y^{(q)})^{(q)} + (p_1 y^{(q-1)})^{(q-1)} \cdots + (p_q y).$$

Accordingly, l is defined by the $q+1$ functions of time $p_0, \ldots, p_q$.

We now strengthen our optimality results to get $\mathcal{C}^0$-sufficient optimality conditions.

Central field, Hamilton-Jacobi-Bellman equation and $\mathcal{C}^0$-sufficient optimality conditions. Let $(\bar{z}, \bar{u})$ be an extremal defined on $[0, T]$ of the time-optimal control problem of $\dot{x} = f(x, u)$, x in a manifold M, extremities being fixed. As previously, we make the following regularity assumptions:

(i) The strict Legendre-Clebsch conditions holds along the extremal,

$$\frac{\partial^2 H}{\partial u^2}(\bar{z}(t), \bar{u}(t)) < 0,\ t \in [0, T].$$

(ii) On each subinterval $0 < t_0 < t_1 \leq T$ the singularity is of codimension one.

(iii)We are in the normal case where $H = \langle p, f(x, u)\rangle$ is not zero and p can be chosen on the level set $\widetilde{H} = -1 + \langle p, f(x, u)\rangle = 0$.

In this case, our reference control is smooth and the reference extremal is solution of a system defined by a smooth Hamiltonian H_r. We denote by φ_t the one parameter local group

$$\varphi_t = \exp t\overrightarrow{H}_r$$

and by $L(t)$ the one parameter family of Lagrangian manifolds image of the fiber $T^*_{x_0}M$. Assume that the reference extremal curve is one-to-one and that there exists no conjugate point on $[0,T]$. Then we can imbed the reference solution into a *central field*, projection of the Lagrangian submanifolds on M.

This construction is valid in a neighbourhood of the reference curve, but it can be prolongated to a maximal open set W homeomorphic to a convex cone, each point of the domain being related to a unique point of $\Pi(L(t))$. Our aim is to prove that the reference extremal curve is optimal with respect to all trajectories with the prescribed extremities contained in this set W. The first result is the following [14].

Lemma 1.9. *(Verification lemma) Excluding* $x(0)$*, assume that there exists an open neighbourhood* N *of the reference trajectory and two smooth mappings* $S : W \to \mathbf{R}$ *and* $\hat{u} : W \to U$ *such that for each* (x,u) *in* $W \times U$ *we have the maximization condition*

$$\widetilde{H}(x, dS(x), \hat{u}(x)) \geq \widetilde{H}(x, dS(x), u)$$

and $\widetilde{H}(x, dS(x), \hat{u}(x)) = 0$*. Then the reference trajectory is optimal among all the trajectories of the system with the same extremities and contained in* W*.*

Proof. Let $0 < \bar{t}_0 \leq \bar{t}_1 \leq T$ and let (x,u) be a trajectory of the system defined on $[t_0, t_1]$, contained in W and satisfying the boundary conditions $x(t_0) = \bar{x}(\bar{t}_0)$, $x(t_1) = \bar{x}(\bar{t}_1)$. If we note $T(x,u)$ the transfer time $t_1 - t_0$, we must prove that $T(\bar{x}, \bar{u}) \leq T(x,u)$. By definition,

$$\begin{aligned} 1 &= \langle dS(x(t)), f(x(t), u(t)) \rangle - \widetilde{H}(x(t), dS(x(t)), u(t)) \\ &= dS(x(t))\dot{x}(t) - \widetilde{H}(x(t), dS(x(t)), u(t)). \end{aligned}$$

Therefore we obtain

$$T(x,u) = \int_{t_0}^{t_1} dt = S(x(t_1)) - S(x(t_0)) - \int_{t_0}^{t_1} \widetilde{H}(x(t), dS(x(t)), u(t))dt.$$

Besides, since along the reference curve we have $\widetilde{H} = 0$, one has

$$T(\bar{x}, \bar{u}) = S(\bar{x}(\bar{t}_1)) - S(\bar{x}(\bar{t}_0)).$$

Because the extremities are fixed we get

$$T(x,u) - T(\bar{x}, \bar{u}) = -\int_{t_0}^{t_1} \widetilde{H}(x(t), dS(x(t)), u(t))dt$$

which is nonnegative by virtue of the maximization condition and $T(x,u) \geq T(\bar{x}, \bar{u})$. This proves the result. ■

The construction of S is equivalent to solve the standard *Hamilton-Jacobi-Bellman* equation

$$\max_{u\in U} H(x, \frac{\partial S}{\partial x}) = 1$$

the transfer times of the extremal curves being $T(x,u) = S(x(t_1)) - S(x(t_0))$ and the adjoint covector p being $\partial S/\partial x$.

Geometric construction and concluding result. Restated in symplectic formalism, the construction of the solution S of the Hamilton-Jacobi-Bellman equation is clear. The set $L = \{p = \partial S/\partial x\}$ is a Lagrangian manifold and the restriction of Π to L is a diffeomorphism. To construct L we use the central field and we take $L = (\cup_{t>0} L(t)) \cap \{\widetilde{H} = 0\}$. Indeed, for each t the manifold $L(t)$ is Lagrangian and homogeneous with respect to p which is normalized by the condition $p^0 = 1$, p^0 being the dual to the cost. The projection gives the set of points at time t from x_0. In particular, $L(t) \cap \{\widetilde{H} = 0\}$ is an isotropic manifold of dimension $n-1$. It is straightforward to prove that it defines a Lagrangian manifold when saturated by the flow.

Proposition 1.16. *Under our assumptions, the reference extremal curve is optimal with respect to all curves solution of the system with same extremities and contained in the domain covered by the central field.*

1.3.3 Examples

Sub-Riemannian Heisenberg Case

We consider a system in $\mathbf{R}^n$ of the form

$$\dot{x} = \sum_{i=1}^{m} u_i F_i(x).$$

Such systems are called symmetric and are relevant in robotics. In particular, they are connected to motion planning: given a path $\gamma \ : \ t \mapsto \gamma(t)$, t in $[0,1]$, find an admissible trajectory (x,u) of the system with same extremities. Moreover, in order to avoid obstacles, we fix around γ a tube W in which the admissible trajectories have to stay. For such problems we can also assign a cost, *e.g.* minimum time or minimum length. We shall consider here the latter case where the length is defined by a Riemannian metric. The restriction of this metric to a the distribution generated by $F_1, \ldots, F_m$ defines a sub-Riemannian problem. If we choose an orthonormal subframe in the distribution, the problem becomes $\dot{x} = \sum_{i=1}^{m} u_i F_i(x)$ with criterion

$$\int_0^T (\sum_{i=1}^{m} u_i^2)^{1/2} dt$$

and the size of the tube W can be set by the metric. We restrict the analysis to the standard Heisenberg contact case in $\mathbf{R}^3$.

Consider the following system in $\mathbf{R}^3$:

$$\dot{q} = u_1 F_1(q) + u_2 F_2(q)$$

where $F_1 = \partial/\partial x + y\partial/\partial z$, $F_2 = \partial/\partial y - x\partial/\partial z$. If we set $F_3 = \partial/\partial z$, we get $[F_1, F_2] = 2\partial/\partial z$. Moreover, all the Lie brackets of order three or more are zero. The distribution D spanned by F_1, F_2 is a *contact distribution* defined as the kernel of the 1-form $\alpha = dz + (xdy - ydx)$. A sub-Riemannian metric is associated to a metric of the form $g = a(q)dx^2 + 2b(q)dxdy + c(q)dy^2$. By choosing suitable coordinates, the smooth functions a, b and c can be normalized to $a = c = 1$ and $b = 0$. The case $g = dx^2 + dy^2$ is called the *Heisenberg case* or the *sub-Riemannian flat contact case*.

Heisenberg sub-Riemannian geometry and the Dido problem. We observe that the previous problem can be written $\dot{x} = u_1$, $\dot{y} = u_2$, $\dot{z} = \dot{x}y - \dot{y}x$, and

$$\int_0^T (\dot{x}^2 + \dot{y}^2)^{1/2} dt \to \min$$

in order that:

(i) The length of a curve $t \mapsto (x(t), y(t), z(t))$ is the length of the projection in the xy-plane.

(ii) If a curve joins (x_0, y_0, z_0) to (x_1, y_1, z_1), $z_1 - z_0$ is proportional to the area swept by the projection of the curve on the xy-plane.

Hence, our problem is dual to the *Dido problem* whose solutions are circles: find closed curves in the plane of prescribed length and maximum enclosed area.

Computation of the extremal curves. According to Maupertuis principle, minimizing the length is the same as minimizing the energy,

$$\int_0^T (u_1^2 + u_2^2) dt \to \min$$

the final time T being fixed. The Hamiltonian is

$$\widetilde{H}(x, p, u) = \sum_{i=1}^{2} (p^0 u_i^2 + u_i P_i)$$

with the Hamiltonian lifts or *Poincaré coordinates* $P_i = \langle p, F_i \rangle$, $i = 0, \ldots, 2$ and $F_0 = \partial/\partial z$. We are in the normal case and set $p^0 = -1/2$ so that, in these coordinates, the extremals are the solutions of:

$$\dot{x} = P_1,\ \dot{y} = P_2,\ \dot{z} = P_1 y - P_2 x$$

and

$$\dot{P}_1 = 2P_3P_2,\ \dot{P}_2 = -2P_3P_1,\ \dot{P}_3 = 0.$$

By setting $P_3 = \lambda/2$, we obtain the equation of a linear pendulum, $\ddot{P}_1 + \lambda^2 P_1 = 0$ and the equations are integrable by quadrature with trigonometric functions. The integration is straightforward if we observe that

$$\ddot{z} - \frac{\lambda}{2}\frac{d}{dt}(x^2 + y^2) = 0.$$

We get the following parameterization of the extremals starting from $z_0 = (0,0,0)$. If $\lambda = 0$, $x(t) = At\cos\varphi$, $y(t) = At\sin\varphi$, $z(t) = 0$, that is we have straight lines. If $\lambda \neq 0$, $x(t) = A/\lambda\ (\sin(\lambda t + \varphi) - \sin\varphi)$, $y(t) = A/\lambda\ (\cos(\lambda t + \varphi) - \cos\varphi)$, $z(t) = (A^2/\lambda)t - (A^2/\lambda^2)\sin\lambda t$, with $A = (P_1^2 + P_2^2)^{1/2}$ and φ is the argument of the vector $(\dot{x}, -\dot{y})$.

Conjugate points, global optimality. The computation of conjugate points by means of the previous parameterization is obvious: the extremal straight lines have no conjugate points, and the extremals which project onto circles in the xy-plane have their first conjugate points after one revolution. We note $S(a,r)$ the sphere of points at sub-Riemannian distance r of the center a. Note that the sub-Riemannian distance is continuous. It is isometric to the sphere of same radius centered at the origin, $S(0,r)$, which is a surface of revolution with respect to the z-axis and symmetric with respect to the xy-plane. Eventually, we can take a unit radius by homogeneity. Standard existence theorems tell us that the sphere is made of extremity points of minimizing extremals of unit length. As a consequence of our computations, a conjugate point is also a *cut point* where a minimizer ceases to be globally optimal. This is a degenerate situation similar to the Riemannian case on $\mathbf{S}^2$. Arguably, as the sphere is a surface of revolution with respect to the z-axis, there is a one parameter family of extremal curves intersecting exactly at the same point.

The Flat Torus

Another interesting example is the flat torus $\mathbf{T}^2$ obtained by identifying points on the opposite sides of the square $[0,1] \times [0,1]$. The extremals with respect to length are the images of straight lines in $\mathbf{R}^2$ through this identification. Since the problem is flat, the curvature is zero and the Jacobi equation is trivial. There is no conjugate point and optimality is related to the topology of the torus. Actually, if x_0 is chosen at the center of the square, then an extremal is optimal until it reaches the sides where it meets another minimizing curve (there are up to four such curves at a corner). To describe the underlying topology, we can choose coordinates l_1, l_2 which are angles. Given any extremal, there are infinitely many extremals with same extremities but different rotation numbers. This property is crucial to understand the orbit transfer problem.

1.4 Time-optimal Transfer Between Keplerian Orbits

An important matter in astronautics is to transfer a satellite between elliptic orbits. Optimal criterions related to this issue are the maximization of the final mass (which amounts to minimizing the fuel consumption) or the minimization of the transfer time. We shall consider this second problem for two reasons. First, recent research projects concern orbit transfer with electro-ionic propulsion for which the thrust is very low and the transfer duration very long (up to several months). Moreover, since the transfer is always with maximum thrust, the structure of the minimum time extremals is simpler than in the minimum consumption case [10] and fit in the smooth case previously analyzed.

1.4.1 Model and Basic Properties

Let m be the mass of the satellite, and let F be the thrust of the engine. The equation describing the system in Cartesian coordinates are:

$$\ddot{q} = -q\frac{\mu}{r^3} + \frac{F}{m}$$

where q is the position of the satellite measured in a fixed frame I, J, K whose origin is the Earth center, where $r = |q| = (q_1^2 + q_2^2 + q_3^2)^{1/2}$, and where μ is the gravitation constant. The free motion with $F = 0$ is *Kepler equation*. The thrust is bounded, $|F| \leq F_{\max}$, and the mass variation is described by

$$\dot{m} = -\frac{|F|}{v_e} \tag{1.21}$$

where v_e is a positive constant (see Table 1.1). Practically, the initial value m_0 of the mass is known, and m has to remain greater than the mass of the

Table 1.1. Physical constants

Variable	Value
μ	5165.8620912 $\mathrm{Mm^3 \cdot h^{-2}}$
$1/v_e$	$1.42e-2$ $\mathrm{Mm^{-1} \cdot h}$
m_0	1500 kg
$F_{\max}$	3 N

satellite without fuel: $m \geq \chi_0$. If (1.21) is not taken into account, we have a simplified *constant mass model*. Roughly speaking, this model is sufficient for our geometric analysis, but the mass variation has to be included for numerical computation. Besides, if the thrust is maximal, maximizing the final mass reduces to minimizing the transfer time. If $q \wedge \dot{q}$ is not zero, the thrust can be decomposed in a moving frame attached to the satellite. A canonical choice consists in the *radial-orthoradial frame*: $F = u_r F_r + u_{\text{or}} F_{\text{or}} + u_c F_c$ with

$$F_r = \frac{q}{r}\frac{\partial}{\partial \dot{q}}, \quad F_c = \frac{q \wedge \dot{q}}{|q \wedge \dot{q}|}\frac{\partial}{\partial \dot{q}}$$

and $F_{\text{or}} = F_c \wedge F_r$. If $u_c = 0$, the state space is the tangent space to the osculating plane generated by q and $\dot{q}$ and we have a 2D-problem. For the sake of simplicity, we shall restrict our study to this setting which already exhibits all the relevant features of the whole system. We first recall the classical properties of the Kepler equation.

Proposition 1.17. *The two vectors below are first integrals of the Kepler equation:*

$$c = q \wedge \dot{q} \ \textit{(angular momentum)}$$
$$L = -q\frac{\mu}{r} + \dot{q} \wedge c \ \textit{(Laplace vector).}$$

Moreover, the energy $H(q, \dot{q}) = 1/2\ \dot{q}^2 - \mu/r$ is preserved and the following relations hold:

$$L.c = 0, \ \ L^2 = \mu^2 + 2Hc^2.$$

Proposition 1.18. *If $c = 0$, the motion is on a colliding line. Otherwise, if $L = 0$ the motion is circular while if $L \neq 0$ and $H < 0$ the trajectory is an ellipse:*

$$r = \frac{c^2}{\mu + |L|\cos(\theta - \theta_0)}$$

where θ_0 is the argument of the pericenter.

Definition 1.11. *The domain $\Sigma_e = \{H < 0, \ c \neq 0\}$ is filled by elliptic orbits and is called the* elliptic (2D-elliptic in the planar case) domain.

Geometric coordinates. To each pair (c, L) corresponds a unique oriented ellipse. Using these coordinates, we have a natural representation of the system. Namely,

$$\dot{c} = q \wedge \frac{F}{m}$$
$$\dot{L} = F \wedge c + \dot{q} \wedge (q \wedge \frac{F}{m})\cdot$$

Since Σ_e is a fiber bundle, one needs a coordinate to describe the evolution on the fiber itself. The fiber is $\mathbf{S}^1$ and the coordinate is the so-called *cumulated longitude*. A second representation is thus provided by the *orbit elements*. Restricting to the 2D case, one has $x = (P, e, l)$ with

- P, *semi-latus rectum* related to the semi-major axis by $P = a(1 - e^2)$.
- $e = (e_x, e_y)$, *eccentricity vector* oriented along L, that is along the semi-major axis, and whose norm is the eccentricity of the ellipse.
- l, *cumulated longitude* measured with respect to the q_1-axis.

Using the radial-orthoradial decomposition in which the dynamics is $\dot{x} = F_0 + u_r F_r + u_{\mathrm{or}} F_{\mathrm{or}}$, the vector fields are

$$F_0 = \sqrt{\frac{\mu}{P}}\frac{W^2}{P}\frac{\partial}{\partial l}$$

$$F_r = \sqrt{\frac{P}{\mu}}\left(+\sin l\frac{\partial}{\partial e_x} - \cos l\frac{\partial}{\partial e_y}\right)$$

$$F_{\mathrm{or}} = \sqrt{\frac{P}{\mu}}\left(\frac{2P}{W}\frac{\partial}{\partial P} + (\cos l + \frac{e_x + \cos l}{W})\frac{\partial}{\partial e_x} + (\sin l + \frac{e_y + \sin l}{W})\frac{\partial}{\partial e_y}\right).$$

According to the data provided by the French Space Agency (CNES), the problem is to transfer the system from a low eccentric initial ellipse towards the geostationnary orbit. The boundary conditions are given in Table 1.2. Though the longitude is free on the initial and terminal orbits, we set $l(0) = \pi$ for numerical issues[5].

Table 1.2. Boundary conditions

Variable	Initial cond.	Final cond.
P	11.625 Mm	42.165 Mm
e_x	0.75	0
e_y	0	0
h_x	0.0612	0
h_y	0	0
l	π rad	103 rad

1.4.2 Maximum Principle and Extremal Solutions

Definition 1.12. *We call SR-problem with drift the time-optimal problem for a system of the form*

[5] This amounts to start at the apocenter where the attraction is the weakest. The numerical integration is thus improved.

$$\dot{x} = F_0 + \sum_{i=1}^{m} u_i F_i$$

with $x \in \mathbf{R}^n$, $F_0, \ldots, F_m$ smooth vector fields, the control u in $\mathbf{R}^m$ being bounded by $\sum_{i=1}^m |u_i|^2 \leq 1$.

Let the H_i's be the usual Hamiltonian lifts $\langle p, F_i \rangle$, $i = 0, \ldots, m$, and let Σ be the switching surface $\{H_i = 0,\ i = 1, \ldots, m\}$. The maximization of the Hamiltonian $H = H_0 + \sum_{i=1}^m u_i H_i$ outside Σ implies that

$$u_i = \frac{H_i}{\sqrt{\sum_{i=1}^m H_i^2}},\ i = 1, \ldots, m. \tag{1.22}$$

Plugging (1.22) into H, one defines the Hamiltonian function

$$H_r = H_0 + \left(\sum_{i=1}^m H_i^2 \right)^{\frac{1}{2}}. \tag{1.23}$$

Definition 1.13. *The corresponding solutions are called* order zero extremals. *From the maximum principle, optimal extremals are contained in the level set $\{H_r \geq 0\}$. Those in $\{H_r = 0\}$ are* exceptional.

The following result is standard.

Proposition 1.19. *The order zero extremals are smooth responses to smooth controls on the boundary of $|u| \leq 1$. They are singularities of the endpoint mapping $E_{x_0,T} : u \mapsto x(T, x_0, u)$ for the* L$^\infty$*-topology when u is restricted to the unit sphere $\mathbf{S}^{m-1}$.*

In order to construct all extremals, we must analyze the behaviour of those of order zero near the switching surface. On one hand, observe that we can connect two such arcs at a point located on Σ if we respect the Weierstraß-Erdmann conditions

$$p(t+) = p(t-),\ H_r(t+) = H_r(t-)$$

where t is the time of contact with the switching surface. Those conditions, obtained in classical calculus by means of specific variations, are contained in the maximum principle. On the other hand, singular extremals satisfy $H_i = 0$, $i = 1, \ldots, m$, and are contained in Σ. They are singularities of the endpoint mapping if u is interior to the control domain, $|u| < 1$. For the 2D orbit transfer, we restrict ourselves to the constant mass model. Since Lie brackets can easily be computed in Cartesian coordinates, the system is written $\dot{x} = F_0(x) + u_1 F_1(x) + u_2 F_2(x)$ where $x = (q, \dot{q})$ and where F_0 is the Kepler vector field with $F_1 = \partial/\partial_{\dot{q}_1}$, $F_2 = \partial/\partial_{\dot{q}_2}$. The following result is straightforward.

Lemma 1.10. *Let $\mathcal{D}$ be the controlled distribution generated by F_1 and F_2. Then $[\mathcal{D},\mathcal{D}]=0$ and at each point the rank of the span of F_1, F_2, $[F_0,F_1]$ and $[F_0,F_2]$ is four.*

This allows to make a complete classification of the extremals. Differentiating along an extremal curve, one gets

$$\dot{H}_i = \{H_i, H_0\},\ i=1,2.$$

At a switching point, $H_1 = H_2 = 0$ but, since F_1, F_2, $[F_0,F_1]$ and $[F_0,F_2]$ form a frame, both Poisson brackets $\{H_1,H_0\}$ and $\{H_2,H_0\}$ cannot be vanish (this is the so-called *order one* case [2]). In order to understand the behaviour of extremals in the neighbourhood of such a point, we make a polar blowing up

$$H_1 = \rho\cos\varphi,\ H_2 = \rho\sin\varphi$$

and we get

$$\begin{aligned}
\dot{\rho} &= \cos\varphi\{H_1,H_0\} + \sin\varphi\{H_2,H_0\}\\
\dot{\varphi} &= 1/\rho\ \left(-\sin\varphi\{H_1,H_0\} + \cos\varphi\{H_2,H_0\}\right).
\end{aligned}$$

Extremals curves crossing Σ are obtained by solving $\dot{\varphi}=0$ and the following holds.

Proposition 1.20. *There are extremals curves made of two order zero extremals concatenated and crossing Σ with a given slope. The corresponding control rotates instantaneously of an angle π at the contact with the switching surface. The resulting singularity of the extremal curve is called a Π-*singularity*.*

Since these singularities must be isolated, the only extremal curves for the orbit transfer are order zero extremal or finite concatenation of such arcs with Π-singularities at the junctions. Hence, numerically, we only compute order zero extremals since Π-singularities where the switching surface is crossed with a given slope are properly handled by an integrator with adaptive step length.

1.4.3 Numerical Resolution

Shooting Function

The boundary value problem defined by the maximum principle is solved by shooting. The shooting function is defined by (1.10). Namely, when the initial and final states are fixed, $x(0)=x_0$ and $x(T)=x_1$, one has

$$S \ : \ (T, p_0) \in \mathbf{R}_+^* \times \mathbf{P}^{n-1} \mapsto \exp_{x_0,T}(p_0) - x_1 \tag{1.24}$$

where the exponential mapping at time T is as before defined on $\mathbf{P}^{n-1}$ by $\exp_{x_0,T}(p_0) = \Pi(z(T, x_0, p_0))$. In the standard transfer case, the final longitude is free and the equation involving l_f has to be replaced by the transversality condition $p_l(T) = 0$, where p_l is the adjoint state to l. Practically, the n-dimensional nonlinear equation (1.24) on $\mathbf{R}\times\mathbf{P}^{n-1}$ is treated as an equation on $\mathbf{R}^{n+1}$, the initial adjoint covector p_0 being taken in $\mathbf{R}^n$ and normalized according to $p_0 \in \mathbf{S}^{n-1}$. Alternatively, in the normal case one can also prescribe the Hamiltonian level set, *e.g.* $H_r = 1$. Since equation (1.24) is to be solved by a Newton-like method, an important issue is the regularity of the shooting mapping. Here, the extremals are smooth outside Π-singularities so the analysis consists in studying the shooting mapping in the neighbourhood of such points. To this end, we use the *nilpotent model* of the 2D problem obtained in [2]. Here, the dimension is four and a nilpotent approximation with brackets of length greater than three all vanishing is

$$\begin{array}{ll} \dot{x}_1 = 1 + x_3 & \dot{x}_2 = x_2 \\ \dot{x}_3 = u_1 & \dot{x}_4 = u_2. \end{array}$$

The coupling of the system arises from the constraint on the control, $u_1^2+u_2^2 \leq 1$. Clearly enough, the extremals of such a system are given by

$$\dot{x}_3 = \frac{at+b}{\sqrt{(at+b)^2 + (ct+d)^2}}$$
$$\dot{x}_4 = \frac{ct+d}{\sqrt{(at+b)^2 + (ct+d)^2}}$$

where $a = -p_1(0)$, $b = p_3(0)$, $c = -p_2(0)$ and $d = p_4(0)$. The *switching function* is $t \mapsto (at+b, ct+d)$, and the set Σ of initial covectors generating switch points is stratified as follows. If a and c are both nonzero, the existence of a time such that the two components of the switching function vanish simultaneously reduces to the condition $ad - bc = 0$, so the first strata is the quadric $\Sigma_1 = \{p_1 \neq 0,\ p_2 \neq 0,\ p_1p_4 - p_2p_3 = 0\}$. If a and b are zero while c is not, there is no condition on d and, by symmetry, we also get the two disjoint unions of half planes $\Sigma_2^1 = \{p_1 = p_2 = 0,\ p_3 \neq 0\}$ and $\Sigma_2^2 = \{p_3 = p_4 = 0,\ p_1 \neq 0\}$. Eventually, note that a, b, c and d all zero is impossible since no singular control is allowed here (see Section 1.4.2). As a result, Σ is partitionned into a set of codimension one, and two sets of codimension two:

$$\Sigma = \Sigma_1 \cup (\Sigma_2^1 \cup \Sigma_2^2).$$

The fact that these subsets are of codimension greater or equal to one implies that, despite the existence of the singularities illustrated below, the numerical computation is essentially reduced to the smooth case and thus tractable. Regarding Σ_1, let a and c be nonzero reals. Let then $\delta = (b/a - d/c)/2$ be the half distance between the roots of each component of the switching function. By symmetry, we can assume that a and c are equal and positive. Up to a translation of time, integrating the nilpotent model amounts to integrate

$$\dot{\xi} = \frac{t-\delta}{\sqrt{t^2+\delta^2}}$$

as well as the symmetric term ($t+\delta$ at the numerator). For nonzero δ, one gets

$$\xi_t(\delta) = \sqrt{t^2+\delta^2} - |\delta|(\operatorname{arc\,sh}\frac{t}{\delta}+1) + \text{constant}$$

where t is a fixed positive time. We have a singularity of the kind $\delta \log|\delta|$ and the shooting mapping is continuous but not differentiable at $\delta = 0$, that is when Σ_1 is crossed. Finally, as for Σ_2^2, let c and d be zero (the case of Σ_2^1 being treated analogously), and let $\sigma = -b/a$ be the unique root of the first component of the switching function when a is not zero. Assume for instance that a is positive so that the exponential is computed by integrating

$$\dot{\xi} = \frac{t-\sigma}{|t-\sigma|}.$$

Hence, for a positive fixed t, up to a constant

$$\begin{aligned} \xi_t(\delta) &= t && \text{if } \sigma < 0 \\ &= t - 2\sigma && \text{if } 0 \le \sigma \le t \\ &= -t && \text{if } \sigma > t. \end{aligned}$$

Accordingly, the function is not differentiable at $\sigma = 0$ and $\sigma = t$, and with zero derivative outside $[0, t]$. In other words, for $p_1 > 0$, the shooting mapping is singular outside the cone $\{-p_1 t \le p_3 \le 0\}$ included in Σ_2^2, and not differentiable on its boundary.

Homotopy on the Maximal Thrust

Beyond regularity issues, a delicate task is to provide the Newton method with nice initial guesses for T and p_0. Since the convergence is only local, no matter how smooth the function may be, a relevant approach is *homotopy*. Indeed, our transfer problem is naturally embedded in a family of such problems parameterized by the maximum thrust $F_{\max}$. Moreover, one can expect that given two such thrusts $F^0_{\max}$ and $F^1_{\max}$ close enough, the associated solutions T^i, p^i_0, $i = 0, 1$ also be close. This is the basic idea of homotopy which connects the simple problem with $F^0_{\max}$ big (the bigger the thrust, the shorter the transfer time and the easier the control problem) to the more intricate one with $F^1_{\max}$ smaller. One may for instance consider the sequence of intermediate problems generated by the convex homotopy $F_{\max} = (1-\lambda)F^0_{\max} + \lambda F^1_{\max}$, where λ in $[0, 1]$ is the homotopy parameter. Then *discrete homotopy* consists in picking a finite sequence $\lambda_0 = 0 < \cdots < \lambda_k < \cdots < \lambda_N = 1$ and trying to follow the associated path of zeros: the solution at step k is supposed to be an initial guess precise enough to ensure convergence of the solver at step $k+1$. More subtile alternatives where the step on the homotopic parameter is automatically adjusted are *simplicial* or *differential homotopy* [10]. In the case of discrete homotopy—often refered to as *discrete continuation*—, the minimal regularity required to ensure the process is relevant is provided by the following proposition [5].

Proposition 1.21. *The value function $F_{\max} \mapsto T(F_{\max})$ mapping to each positive maximum thrust the corresponding minimum time is right continuous for the transfer problem (2D or 3D, constant mass or not).*

As a matter of fact, we will use a decreasing sequence of thrusts bounds $(F^k_{\max})_k$. Therefore, right continuity of the value function is enough to guarantee that $T(F^k_{\max})$ tends to $T(F_{\max})$ when the thrusts decrease to $F_{\max}$. But while mere discrete homotopy is used to initialize the search for p_0, a much more precise guess for the minimum time is available. Clearly, the value function $T(F_{\max})$ is decreasing, and one can easily prove that the product $T(F_{\max}) \cdot F_{\max}$ is bounded below. Actually, it is also bounded over (see [6]) and the conjecture is that it has a limit when $F_{\max}$ tends to 0. In practice, we use the heuristic $T(F_{\max}) \cdot F_{\max} \simeq$ constant. Figure 1.6 presents the result of such a computation for a medium thrust of 3 Newtons.

Conjugate Points

In order to deal with conjugate and not focal points, we restrict ourselves to the transfer with fixed final longitude, l^f. As a result, the initial and final state are prescribed, and we are in the situation of Section 1.3.2. The shooting mapping is exactly defined by (1.24) and an extremal is easily computed by solving the shooting problem for a fixed final longitude close to the one obtained with l_f free. So as to integrate the Jacobi equation

$$\delta\dot{z} = d\overrightarrow{H}_r(z(t)) \cdot \delta z$$

along the resulting extremal, $z(t) = z(t, t_0, x_0, \bar{p}_0)$ where $\bar{p}_0$ is a root of the shooting mapping, the standard procedure is to consider the augmented system

$$z = \overrightarrow{H}_r(z),\ \delta\dot{z} = d\overrightarrow{H}_r(z) \cdot \delta z$$

with initial condition $(x_0, \bar{p}_0)$ on z. As for $\delta z = \delta x, \delta p)$, the initial Jacobi fields must span the $(n-1)$ dimensional tangent space to $\{x_0\} \times \mathbf{S}^{n-1}$ so that

$$\delta x_i(0) = 0,\ \delta p_i(0) \perp \bar{p}(0),\ i = 1, \ldots, n-1.$$

According to (1.18), conjugate times are roots of

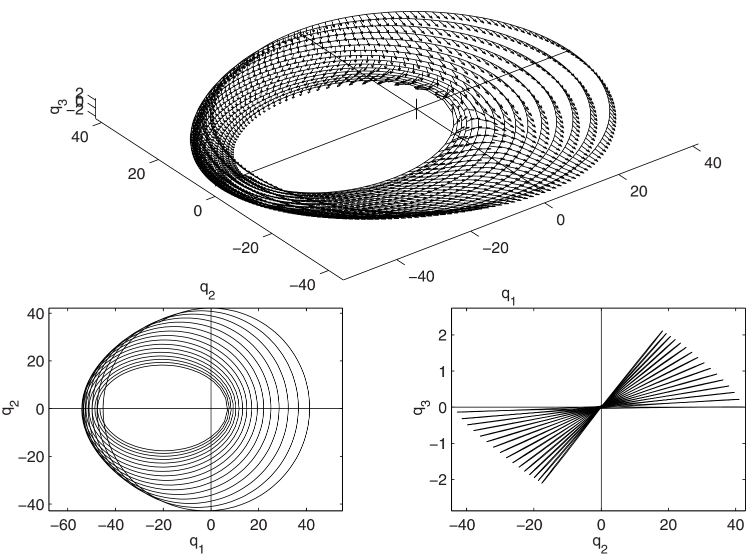

Fig. 1.6. Three dimensional transfer for 3 Newtons. The arrows indicate the action of the thrust. The main picture is 3D, the other two are projections. The duration is about twelve days.

$$\det(\delta_1 x(t), \ldots, \delta_{n-1} x(t), \dot{x}(t)) = 0.$$

This test is important since, numerically, a generic root can be detected by a change in signs. Hence, a rough estimate of conjugate times is easily obtained by dichotomy and then made more accurate, *e.g.*, by a few Newton steps. Nevertheless, it is still compulsary to refine the computation to take into account cases when the rank of the exponential mapping becomes strictly less than $n-2$ (see (1.17)). To this end, we also perform a singular value decomposition so as to check the rank of the span of $\delta x_1, \ldots, \delta x_{n-1}$ along the extremal. The counterpart is that, since the singular values are nonnegative, the detection of zeros is more intricate. An example of this kind of computation in the orbit transfer case is shown at Fig. 1.7.

1.5 Introduction to Optimal Control with State Constraints

In many applied control problems, the systems has state constraints. For instance, in the shuttle re-entry case, there is a constraint on the thermic flow in order to avoid the destruction of the ship. Whereas standard existence theorems hold, the necessary conditions become intricate. Indeed, general extremal

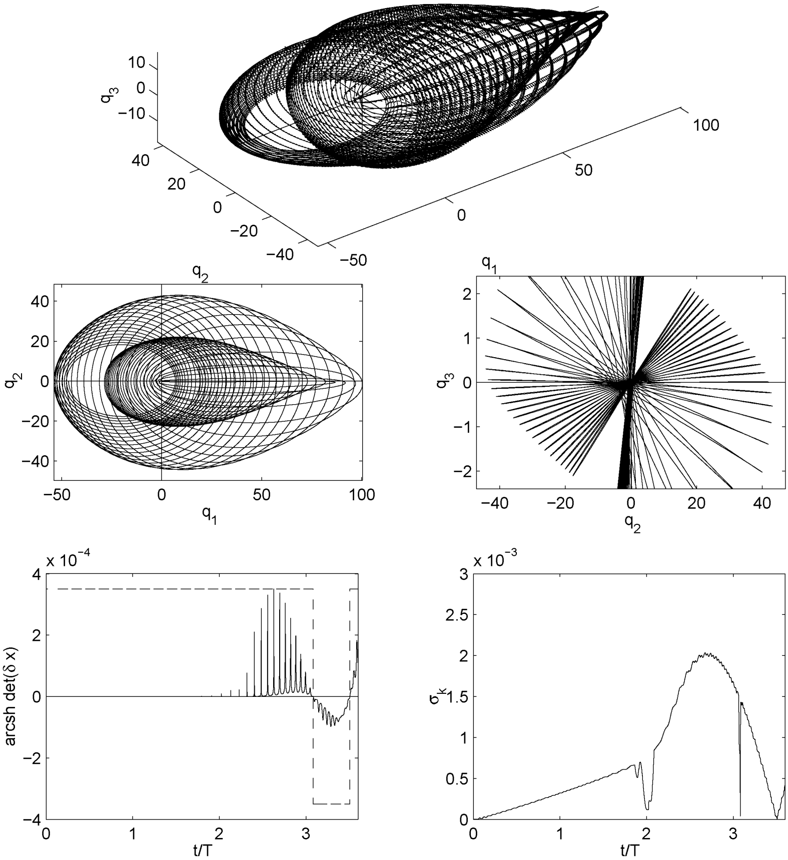

Fig. 1.7. An extremal, which is roughly the same as in Fig. 1.6 (the difference being the fixed final longitude), is extended until 3.5 times the minimum time. Bottom left, the determinant, bottom right, the smallest singular value of the Jacobi fields associated to the extremal. There, two conjugate times are detected. The optimality is lost about three times the minimum time.

curves are parameterized by measures supported by the boundary of the domain. Hence, the key point is to make a geometric analysis of the accessibility set near the constraints of order one to derive these conditions in a geometric form. The ultimate goal is to glue together extremals of the non-constrained problem with those on the boundary to provide an optimal synthesis. This kind of analysis comes from the pioneering work of Weierstraß in 1870 who solved the problem of minimizing the length of a planar curve in the presence of obstacles. The resulting necessary conditions can also be used to analyze *hybrid systems* defined by two subsystems $\dot{x} = f_1(x, u)$, $\dot{x} = f_2(x, u)$, each subsystem describing the evolution in two domains separated by a surface. In this case, Descartes-like refraction rules obtain as consequences of the maximum principle with state constraints. In the orbit transfer case, this approach allows us to take into account the eclipse phenomenon associated to electro-ionic propulsion.

1.5.1 The Geometric Framework

We consider a system of the form $\dot{x} = f(x,u)$, x in $\mathbf{R}^n$ and u in U, subset of $\mathbf{R}^m$, with a cost

$$c(x,u) = \int_0^T f^0(x,u)dt$$

in the presence of one state constraint of the form $g(x) \leq 0$, $g : \mathbf{R}^n \to \mathbf{R}$. We denote by $\widetilde{x} = (x^0, x)$ the extended state, the extended dynamics by $\widetilde{f}$, and by $\widetilde{g} = (0, g)$ the extended state constraint. In order to make a geometric analysis, we restrict our study to piecewise smooth pairs $(\widetilde{x}, \widetilde{u})$ defined on $[0, T]$. An optimal solution $\widetilde{x}$ is thus made of extremal subarcs contained in the open domain $\{g < 0\}$ where the constraint is not active and where the standard maximum principle holds, and of subarcs contained in the boundary, namely *boundary arcs*. In order to decide upon optimality, we split the problem in two.

Optimality of boundary arcs. Clearly, a boundary arc has to be optimal with respect not only to all trajectories contained in the boundary, but also to all neighbouring arcs in the open domain $\{g < 0\}$. This is illustrated by Fig. 1.8 where the hypersurface is a sphere and where the two kinds of variations are represented.

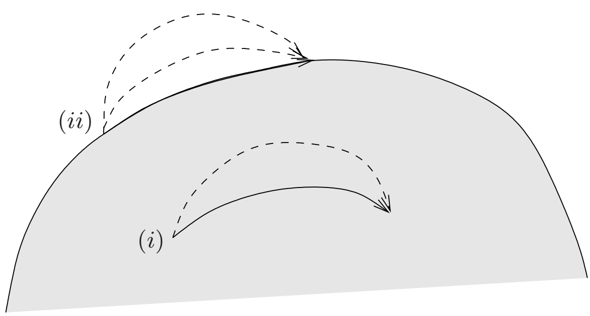

Fig. 1.8. Boundary curve (i) and neighbouring curve outside the constraint (ii)

Optimality conditions at junctions or reflections with the boundary. In this case, the matter is to glue together extremal curves of the non-constrained problem. The junction and reflection conditions were derived by Weierstraß by applying the variations of Fig. 1.9.

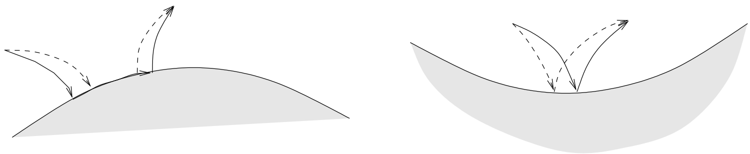

Fig. 1.9. Weierstraß variations

The two previous drawings are the keys of the geometric analysis which is organized as follows. We give first the necessary conditions of [18], illustrating them by several examples. Since the proof is technical, we present next the original proof of Weierstraß using standard calculus of variations.

1.5.2 Necessary Optimality Conditions for Boundary Arcs

Statement

For the sake of simplicity, we shall assume the control domain to be a smooth manifold with boundary defined by $q(u) \leq 0$. Differentiating the constraint along the solution, we get the Lie derivative with respect to the dynamics:

$$\begin{aligned} h(x,u) &= L_{f(x,u)} g \\ &= d/dt\,(g(x)) \\ &= (\nabla g(x) | f(x,u)). \end{aligned}$$

The crucial concept is given by the following definition.

Definition 1.14. *The pair* (x, u) *is of* order one *if*

(i) $h(x,u) = 0$

(ii)$\partial h(x,u)/\partial u \neq 0$

which corresponds to a contact of minimal order *with the boundary.*

In this case, we can define locally a system in the boundary by choosing controls such that $h(x,u)$ is zero. In order to ensure the existence of variations, we further impose the following regularity condition: if u belongs to the boundary of the control domain, $\partial h/\partial u(x,u)$ and $dq(u)/du$ are linearly independent,

$$\frac{\partial h}{\partial u} \wedge \frac{dq}{du} \neq 0.$$

Let $\widetilde{H}(\widetilde{x},\widetilde{p},u)=p^0f^0(x,u)+\langle p,f(x,u)\rangle$ be the Hamiltonian of the problem. If we maximize $\widetilde{H}$ over the set $U(x)$ defined by $h(x,u)=0$, $q(u)\leq 0$, then there are Lagrange multipliers λ and ν such that

$$\frac{\partial \widetilde{H}}{\partial u}=\lambda\frac{\partial h}{\partial u}+\nu\frac{dq}{du}\cdot \tag{1.25}$$

We can now formulate the necessary optimality conditions for boundary arcs.

Theorem 1.2. *Let (x,u) be a smooth optimal solution defined on $[0,T]$ of the problem with fixed extremities. Then, there is a continuous adjoint covector (p^0,p), nonzero, and a scalar function λ such that the following conditions are satisfied:*

$$\dot{x}=\frac{\partial \widetilde{H}}{\partial p}(\widetilde{x},\widetilde{p},u),\ \dot{p}=-\frac{\partial \widetilde{H}}{\partial x}(\widetilde{x},\widetilde{p},u)+\lambda\frac{\partial h}{\partial x}(x,u)$$
$$\widetilde{H}(\widetilde{x},\widetilde{p},u)=\max_{v\in U(x)}\widetilde{H}(\widetilde{x},\widetilde{p},v)=0$$

where, for each t, $\lambda(t)$ is a Lagrange multiplier defined by (1.25). Moreover, p^0 is non-positive, $p(0)$ can be chosen tangent to $\{g=0\}$ and, at each derivability point of λ, the vector $\dot{\lambda}(t)\nabla g(x(t))$ is zero or pointing towards the interior of the domain.

Application in Riemannian Geometry

We consider a smooth hypersurface M defined by the equation $g(x)=0$ and imbedded in the Euclidean space $\mathbf{R}^n$. The manifold M is thus Riemannian for the induced metric. Outside M, the curves of minimum length are straight lines and we can recover the geometric properties defining the extremals from Theorem 1.2. Clearly, they have to be extremal curves for the induced Riemannian metric. Besides, the convexity properties of the surface are important as illustrated by Fig. 1.10.

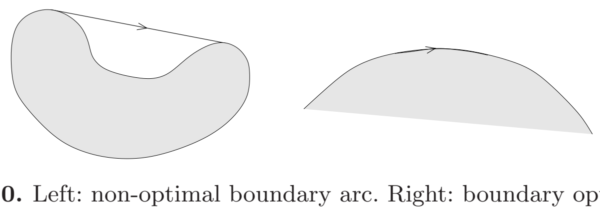

Fig. 1.10. Left: non-optimal boundary arc. Right: boundary optimal arc

The problem can be formulated as the time-optimal control problem of the system $\dot{x}=u$, x in $\mathbf{R}^n$ and u in $\mathbf{R}^n$, where the control domain is the unit sphere $\sum_{i=1}^{n} u_i^2 = 1$. One has $h(x,u) = (\nabla g(x)|u)$ and we consider an arc $t \mapsto x(t)$ such that $g(x) = 0$ and $h(x,u) = 0$. The Hamiltonian is $\widetilde{H} = p^0 + \langle p, u\rangle$ and the adjoint system satisfies

$$\dot{p} = \lambda \frac{\partial h}{\partial x} = \lambda \frac{d}{dt} \nabla g(x) \tag{1.26}$$

The cost multiplier p^0 is not zero and can be normalized to $p^0 = -1$. Hence we get $\langle p, u\rangle = 1$. Moreover, $\partial \widetilde{H}/\partial u = p = \lambda \partial h(x,u)/\partial u + 2\nu dq(u)/du$, so

$$p = \lambda \nabla g(x) + 2\nu u$$

and, multiplying by $u = \dot{x}$, we obtain

$$1 = \langle p, u\rangle = \lambda \langle \nabla g(x), u\rangle + 2\nu |u|^2.$$

Therefore, $\nu = 1/2$ and, using relation (1.26), we get:

$$\dot{\lambda} \nabla g(x) + \dot{u} = 0.$$

This relation tells us that the acceleration $\ddot{x} = \dot{u}$ is perpendicular to the tangent space of M: this is the standard characterization of the geodesic curves on the surface. Moreover,

$$\ddot{x} = -\dot{\lambda} \nabla g(x)$$

and $\ddot{x}$ is pointing outwards which is the convexity relation. Hence, we have a complete description of optimal curves on the surface thanks to Theorem 1.2. In the next paragraph, we present the junction and reflection conditions so as to provide an exhaustive portrait of optimal solutions.

1.5.3 Junction and Reflection Conditions

Statement

We shall consider an arc x defined on $[0, T]$ and meeting the boundary of the domain at a unique time $0 < \tau < T$.

Definition 1.15. *The point $x(\tau)$ is called a* junction point *if $x(t)$ is contained in the boundary for $t \geq \tau$, and a* reflection point *if the arc is contained in the interior of the domain when $t \neq \tau$.*

Let $\widetilde{p} = (p^0, p)$ be the adjoint covector associated to an optimal solution. For a junction point, the *jump condition* is

$$p(\tau+) = p(\tau-) + \mu \nabla g(x(\tau)).$$

For a reflection point, the condition is

$$p(\tau+) = p(\tau-) + \mu \nabla g(x(\tau)),\ \ \mu \geq 0.$$

Geometric Consequence

Consider the junction condition. Since for boundary arcs we can replace $p(\tau+)$ by $p(\tau+) + \nu\nabla g(x(\tau))$ by virtue of the tangency property to $\{g = 0\}$ of Theorem 1.2, at a junction point p can be normalized according to

$$p(\tau+) = p(\tau-).$$

Lemma 1.11. *At a junction point, the adjoint vector can be chosen continuous.*

Furthermore, using the junction condition one has

$$\langle \widetilde{p}(\tau), \widetilde{f}(\widetilde{x}(\tau), u(\tau-))\rangle = \langle \widetilde{p}(\tau), \widetilde{f}(\widetilde{x}(\tau), u(\tau+))\rangle = \max_{v\in U(x)} \widetilde{H}(\widetilde{x(\tau)}, \widetilde{p}(\tau), v)$$

where the maximized Hamiltonian is zero by virtue of Theorem 1.2. As a result, if the control is deduced from the maximization of the Hamiltonian is unique, it has to remain continuous when connecting the trajectory to the boundary. This is the case in the Riemannian problem.

Lemma 1.12. *In the Riemannian case, the straight lines connecting the boundary arcs are tangent to the surface at the junction points.*

1.5.4 Proof of the Necessary Conditions in the Riemannian Case

We consider the problem of minimizing in the plane

$$\int_{t_0}^{t_1} F(x, y, \dot{x}, \dot{y})dt$$

with (x, y) is in $\mathbf{R}^2$ and where F defines a Riemannian metric. In particular, F satisfies the homogeneity relation

$$F(x, y, k\dot{x}, k\dot{y}) = kF(x, y, \dot{x}, \dot{y}),\ k > 0. \tag{1.27}$$

Though homogeneity can be relaxed by imposing a parameterization $\dot{x}^2+\dot{y}^2 = 1$, we shall keep the problem in its general form. This will result in additional properties of the extremals. Now, if ξ and η are variations of the reference curve on the same interval $[t_0, t_1]$, the length variation is:

$$\begin{aligned}\delta l &= \int_{t_0}^{t_1} (F_x\xi + F_y\eta) + (F_{\dot{x}}\dot{\xi} + F_{\dot{y}}\dot{\eta})dt \\ &\simeq l(x+\xi, y+\eta) - l(x, y)\end{aligned}$$

so that, integrating by parts and using zero boundary conditions $\xi(t_0) = \xi(t_1) = 0$, $\eta(t_0) = \eta(t_1) = 0$,

$$F_x - \frac{d}{dt}F_{\dot{x}} = 0,\ F_y - \frac{d}{dt}F_{\dot{y}} = 0.$$

These are the *Euler-Lagrange* equations, not independent because of homogeneity. Indeed, differentiating (1.27) with respect to k at $k = 1$ one obtains

$$\dot{x}F_{\dot{x}} + \dot{y}F_{\dot{y}} = F(x, y, \dot{x}, \dot{y}).$$

Differentiating with respect to (x, y),

$$\begin{aligned} F_x &= \dot{x}F_{\dot{x}x} + \dot{y}F_{\dot{y}x} \\ F_y &= \dot{x}F_{\dot{x}y} + \dot{y}F_{\dot{y}y} \end{aligned}$$

then with respect to $\dot{x}$, we get

$$\begin{aligned} \dot{x}F_{\dot{x}\dot{x}} + \dot{y}F_{\dot{y}\dot{x}} &= 0 \\ \dot{x}F_{\dot{x}\dot{y}} + \dot{y}F_{\dot{y}\dot{y}} &= 0 \end{aligned}$$

and the problem is not regular because the Hessian matrix of the Legendre-Clebsch condition is not invertible. For $(\dot{x}, \dot{y}) \neq (0, 0)$, there is a function F_1 defined by

$$F_{\dot{y}\dot{y}} = \dot{x}^2 F_1,\ F_{\dot{x}\dot{x}} = \dot{y}^2 F_1$$

and

$$F_{\dot{x}\dot{y}} = -\dot{x}\dot{y}F_1.$$

If we introduce

$$T = (F_{x\dot{y}} - F_{y\dot{x}}) + F_1(\dot{x}\ddot{y} - \ddot{x}\dot{y})$$

we get

$$F_x - \frac{d}{dt}F_{\dot{x}} = \dot{y}T,\ F_y - \frac{d}{dt}F_{\dot{y}} = -\dot{x}T$$

and Euler equation is equivalent to the *Weierstraß* equation, $T = 0$, for $(\dot{x}, \dot{y}) \neq (0, 0)$.

Application: Necessary Boundary Optimality Conditions

The previous formulæ for the first variation will be used to derive the necessary boundary conditions. Assume the boundary is one dimensional, and let $(\widetilde{x}, \widetilde{y})$ be a boundary arc on $[t_0, t_1]$. We introduce the variations represented by Fig. 1.11.

At each point of the boundary, we construct a vector n with length u, orthogonal to the boundary and oriented towards the interior of the domain. Namely, $n = (\xi, \eta)$ with

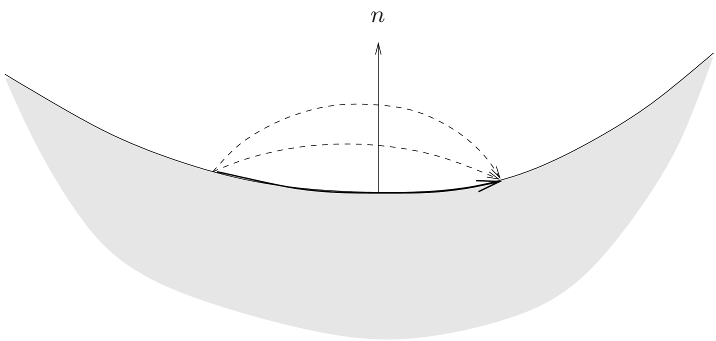

Fig. 1.11. Variation of the boundary arc

$$\xi = -\frac{u\dot{\tilde{y}}}{\sqrt{\dot{\tilde{x}}^2 + \dot{\tilde{y}}^2}}, \ \eta = \frac{u\dot{\tilde{x}}}{\sqrt{\dot{\tilde{x}}^2 + \dot{\tilde{y}}^2}}$$

and we consider the variations of the reference curve $\tilde{x}+\xi$, $\tilde{y}+\eta$. Let $u = \varepsilon p$ for ε positive and p a nonnegative function on $[t_0, t_1]$ such that $p(t_0) = p(t_1) = 0$. The associated variation has zero boundary conditions and, from the previous computation, the length variation is

$$\begin{aligned}\delta J &= \int_{t_0}^{t_1} (F_x - \dot{F}_{\dot{x}})\xi + (F_y + \dot{F}_{\dot{y}})\eta dt \\ &= -\varepsilon \int_{t_0}^{t_1} \widetilde{T} p \sqrt{\dot{\tilde{x}}^2 + \dot{\tilde{y}}^2}\, dt.\end{aligned}$$

Accordingly, if the boundary arc is optimal, one must have $\mathcal{T}$ non-positive along $(\tilde{x}, \tilde{y})$. In the Riemannian case, the Legendre-Clebsch condition $F_1 > 0$ is satisfied and we ge the curvature relation between the extremal tangent to the boundary and the boundary arc itself:

$$\frac{F_{x\dot{y}} - F_{y\dot{x}}}{F_1 \left(\dot{\tilde{x}}^2 + \dot{\tilde{y}}^2\right)^{3/2}} \leq \frac{\dot{y}\ddot{x} - \ddot{y}\dot{x}}{\left(\dot{\tilde{x}}^2 + \dot{\tilde{y}}^2\right)^{3/2}}$$

which amounts to the standard convexity relation for $F = \sqrt{\dot{\tilde{x}}^2 + \dot{\tilde{y}}^2}$.

Junction Conditions with the Boundary

In this case, the variation is represented by Fig. 1.12 and concerns the entry or exit point. The geometric situation leads us to consider two central fields associated respectively to the initial and final point (labels 0 and 1).

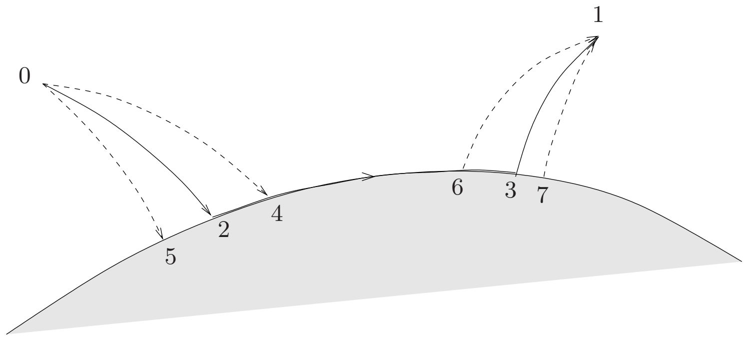

Fig. 1.12. Variations on the entry and exit junction points

In particular, consider the variation of the entry point 2 between 4 and 5. Indexing the length by the extremities of the arcs, we must estimate $l_{04} - (l_{02}+l_{24})$. This computation uses the general formula to estimate the variation of the cost between two curves. Let us denote $\gamma = (x, y)$ the extremal arc 02 defined on $[t_0, t_1]$, $\tilde{\gamma} = (\tilde{x}, \tilde{y})$ the boundary arc defined on $[t_1, t_1 + h]$ $(h > 0)$, and $\gamma + \nu$ the arc 04 also defined on $[t_0, t_1]$. Since the reference curve γ is an extremal, the length variation is, up to first order,

$$\delta l = [F_{\dot{x}}\xi + F_{\dot{y}}\eta]_{t_0}^{t_1} - F(\tilde{x}_2, \tilde{y}_2, \dot{\tilde{x}}_2, \dot{\tilde{y}}_2)h$$

where $\xi(t_0) = \eta(t_0) = 0$ since the initial point is fixed, and $\xi(t_1) = h\dot{\tilde{x}}$, $\eta(t_1) = h\dot{\tilde{y}}$ at the junction point. Hence, the length variation is

$$\begin{aligned}\delta l &= h\left[(\dot{\tilde{x}}F_{\dot{x}} + \dot{\tilde{y}}F_{\dot{y}})|_{\gamma} - F|_{\tilde{\gamma}}\right] \\ &= -hE\end{aligned}$$

where E is the *Weierstraß excess function*:

$$E(x, y, \dot{x}, \dot{y}, \dot{\tilde{x}}, \dot{\tilde{y}}) = F(\tilde{x}, \tilde{y}, \dot{\tilde{x}}, \dot{\tilde{y}}) - (\dot{\tilde{x}}F_{\dot{x}}(x, y, \dot{x}, \dot{y}) + \dot{\tilde{y}}F_{\dot{y}}(x, y, \dot{x}, \dot{y})).$$

Replacing the arc 24 by 25, we get the necessary optimality condition

$$E(x, y, \dot{x}, \dot{y}, \dot{\tilde{x}}, \dot{\tilde{y}}) = 0$$

at the entry point 2, $(\dot{x}, \dot{y})$ being the tangent to the reference extremal curve and $(\dot{\tilde{x}}, \dot{\tilde{y}})$ being the tangent to the boundary. The excess function has the following homogeneity induced by the metric:

$$E(x, y, k\dot{x}, k\dot{y}, \tilde{k}\dot{\tilde{x}}, \tilde{k}\dot{\tilde{y}}) = \tilde{k}E(x, y, \dot{x}, \dot{y}, \dot{\tilde{x}}, \dot{\tilde{y}})$$

for each positive k, $\tilde{k}$. Introducing the slopes

$$p = \frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}} = \cos\theta, \quad q = \frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}} = \sin\theta$$
$$\widetilde{p} = \frac{\dot{\widetilde{x}}}{\sqrt{\dot{\widetilde{x}}^2+\dot{\widetilde{y}}^2}} = \cos\widetilde{\theta}, \quad \widetilde{q} = \frac{\dot{\widetilde{y}}}{\sqrt{\dot{\widetilde{x}}^2+\dot{\widetilde{y}}^2}} = \sin\widetilde{\theta}$$

we get

$$E(x,y,\dot{x},\dot{y},\dot{\widetilde{x}},\dot{\widetilde{y}}) = \sqrt{\dot{\widetilde{x}}^2+\dot{\widetilde{y}}^2}\,E(x,y,p,q,\widetilde{p},\widetilde{q}).$$

In accordance with the mean value theorem, there is θ^* between θ and $\widetilde{\theta}$ such that

$$E(x,y,\cos\theta,\sin\theta,\cos\widetilde{\theta},\sin\widetilde{\theta}) = (1-\cos(\widetilde{\theta}-\theta))F_1(x,y,\cos\theta^*,\sin\theta^*).$$

In the regular case where F_1 is positive, we deduce the following.

Proposition 1.22. *In the regular case, at the entrance and exit junction points with a boundary arc, one must have $\theta = \widetilde{\theta}$: the extremal has to be tangent to the boundary.*

As a consequence, this gives the junction condition of Lemma 1.12, previously obtained as a consequence of the jump condition.

Reflection Condition on the Boundary

In this case, the variation is on the reflection point, see Fig. 1.13.

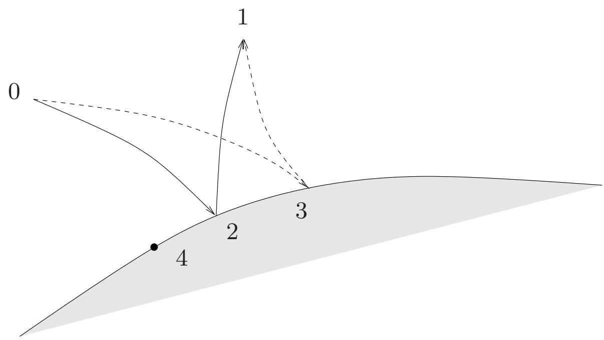

Fig. 1.13. Variations on the entry and exit junction points

The cost variation $l_{031} - l_{021}$ is evaluated by gluing together the central fields with initial point 0 and terminal point 1 along the common boundary arc 23, that is evaluating $(l_{03}-(l_{02}+l_{23}))+(l_{31}+l_{23}-l_{21})$. If h is the variation parameter on the boundary arc, we get

$$\delta l = h\left(E(x_2, y_2, \dot{x}_2^+, \dot{y}_2^+, \dot{\tilde{x}}_2, \dot{\tilde{y}}_2) - E(x_2, y_2, \dot{\bar{x}}_2, \dot{\bar{y}}_2, \dot{\tilde{x}}_2, \dot{\tilde{y}}_2)\right)$$

where $(\dot{\bar{x}}_2, \dot{\bar{y}}_2)$, $(\dot{x}_2^+, \dot{y}_2^+)$ and $(\dot{\tilde{x}}_2, \dot{\tilde{y}}_2)$ are respectively tangent to the arcs 02, 21 and 03. Hence, we get the necessary optimality condition at the reflection point in terms of the corresponding slopes:

$$E(x_2, y_2, p_2^+, q_2^+, \widetilde{p}_2, \widetilde{q}_2) = E(x_2, y_2, \bar{p}_2, \bar{q}_2, \widetilde{p}_2, \widetilde{q}_2).$$

This relation will give us the standard reflection condition if the metric is $F = \sqrt{\dot{x}^2 + \dot{y}^2}$. Indeed, $F_1(x, y, \cos\theta^*, \sin\theta^*) = 1$ and the Descartes rule is obtained:

$$\cos(\widetilde{\theta}_2 - \theta_2^-) = \cos(\widetilde{\theta}_2 - \theta_2^+).$$

The lines reflecting on the boundary must have equal angles. The same approach can be applied for the refraction rule where we glue together on the boundary two central fields with different extremal curves, see Fig. 1.14. In both cases, the rules are given by a jump condition on the adjoint state.

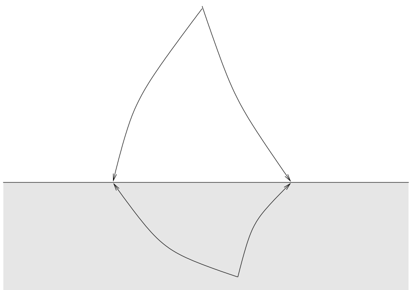

Fig. 1.14. Refraction of two central fields

Bibliographical Notes

For the maximum principle, see the introduction to the discovery in [9]. For the proof, we have followed [15]. The high order maximum principle is due to Krener [13], and the presentation using the Baker-Campbell-Hausdorff formula is inspired by [12]. Elements of symplectic geometry are borrowed from [16], and for the concept of conjugate point we use [4]. The example in sub-Riemannian geometry is excerpted from [3]. For an introduction on orbital transfer and numerical techniques, see [5, 10]. The necessary optimality conditions in the state constrained case are from [18]. For the proof in the planar case, we have followed [1].

References

1. O. Bolza (1904,). *Lectures on the calculus of variations.* Dover, New-York.
2. B. Bonnard, J.-B. Caillau, and E. Trélat (submitted,). Geometric optimal control of Keplerian orbits. *Discrete Cont. Dyn. Syst. Series B.*
3. B. Bonnard and M. Chyba (2003,). *Singular trajectories and their role in control theory.* Number 40 in Math. and Applications. Springer Verlag.
4. B. Bonnard and I. Kupka (1993,). Théorie des singularités de l'application entrée-sortie et optimalité des trajectoires singulières dans le problème du temps minimal. *Forum Mathematicum*, 5:111–159.
5. J.-B. Caillau (2000,). *Contribution à l'étude du contrôle en temps minimal des transferts orbitaux.* PhD thesis, ENSEEIHT, Institut National Polytechnique, Toulouse.
6. J.-B. Caillau, J. Gergaud, and J. Noailles (to appear in *Open Problems in Mathematical Systems and Control*, V. Blondel and A. Megretski Eds.,). *Minimum time control of the Kepler equation.* Princeton University Press.
7. M. P. Do Carmo (1993,). *Riemannian geometry.* Birkhäuser.
8. I. Ekeland (1977,). Discontinuité des champs Hamiltoniens et existence de solutions optimales en calcul des variations. *Pub. IHES*, 47:5–32.
9. R. V. Gamkrelidze (1999,). Discovery of the maximum principle. *Journal of dynamical and control systems*, 5(4):437–451.
10. J. Gergaud and T. Haberkorn (submitted,). Homotopy method for minimum consumption orbit transfer. *ESAIM COCV.*
11. C. Godbillon (1985,). *Géométrie différentielle et mécanique analytique.* Hermann, Paris.
12. H. Hermes (1978,). Lie algebras of vector fields and local approximation of attainable sets. *SIAM J. Control Optim.*, 16(5):715–727.
13. A. J. Krener (1977,). The high order maximum principle and its application to singular extremals. *SIAM J. Control Optim.*, 15(2):256–293.
14. I. Kupka. Personal communication.
15. E. B. Lee and L. Markus (1967,). *Foundations of optimal control theory.* John Wiley, New-York.
16. K. R. Meyer and G. R. Hall (1991,). *Introduction to Hamiltonian dynamical systems.* Springer Verlag.
17. M. A. Naimark (1968,). *Linear differential operators.* Frederick Ungar Publishing Co.
18. L. Pontryagin, V. Boltiansky, and R. Gamkrelidze (1974,). *Théorie mathématique des processus optimaux.* Éditions MIR, Moscow.