UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI MATEMATICA
CORSO DI LAUREA IN MATEMATICA

Nonsmooth Analysis and the Maximum Principle in Control Theory

Author:
Alexander ZASS

Supervisor:
Prof. Carlo MARICONDA

October 14, 2015

Acknowledgements

First of all, I would like to thank my supervisor, Professor Carlo Mariconda, for the patient guidance and precious advice he has provided throughout our work together. I would also like to thank the Professors of the Department who were able to transmit their passion and interest for their subject; without them this journey would have ended differently.

I would also like to express my gratitude to my father, for proofreading these pages, despite finding them full of meaningless mathematical gibberish, and to my mother, for constantly providing endless advice on bibliography and on how you are supposed to quote an author.

My friends and fellow students deserve thanks for helping me smile when the exams were having the best of me. I would like to thank Marta for always being there when times were rough, you are a good friend.

I would like to thank Michela, for being a unique and refreshing person who has helped me keep things in perspective, and for always taking the time to listen to my crazy problems.

Finally, I would like to give another thank you note to Marta and Michela, for helping me stay rational when I was sure I had lost my chance of graduating.

Preface

The purpose of this dissertation is to offer a first look into the broad subject of *Control Theory*. The object we will be studying is the *controlled differential equation*

$$\dot{x}(t) = f\big(t, x(t), u(t)\big), \; u(t) \in U, \tag{1}$$

where u is a measurable function; the couple (f, U) is known as the *control system*. The idea is that of choosing the control u in such a way as to minimize a certain cost functional $J(x, u)$, subject to an endpoint constraint $x(b) \in E \subseteq \mathbb{R}^n$, where E is called the *target set*.

We choose to require only that E be a closed set. In this way, all the more specific results, in which E is either a point, the whole space, a smooth manifold (as is the case of the original Pontryagin's Maximum Principle), or a manifold with boundary will all be special cases of what we have studied. We will not consider, however, more sophisticated versions, in which also the function f is nonsmooth.

In order to do so, though, we will need new tools, those of the so-called *nonsmooth analysis*. In particular, we will introduce the notions of *Clarke's generalized subgradient*, *proximal subdifferential*, *proximal* and *limiting normal cone*.

After this we will be ready to enter the world of Control Theory. Before actually discussing any of the main results, we will look at some of the key aspects of the *Calculus of Variations*, since it is from here that Optimal Control Theory was born.

This dissertation comes to a close with two examples of minimum-time problems; the first one consists of finding the control u which allows us to bring a spaceship to a soft landing on the lunar surface in the least time.

The second one uses a nonsmooth target set, and we will find that the minimum-time solution is not differentiable. I would like to thank Professor Rampazzo for providing me with this example.

The main reference for this work is Francis Clarke's book *Functional Analysis, Calculus of Variations and Optimal Control* (Springer - 2013).

Other books I will be referring to are *Optimal Control* (Birhhäusen - 2002) by Richard Vinter and *Control Theory* (Springer) by Alberto Bressan and Benedetto Piccoli.

Contents

1 Non-Smooth Analysis **9**

1.1 Subdifferential Calculus 9

1.2 Generalized Gradients 11

1.3 Proximal Calculus 13

1.4 Proximal Geometry 17

2 Optimal Control **21**

2.1 The Calculus of Variations 23

2.2 Controllability 35

2.3 Pontryagin Maximum Principle 37

2.4 Problems with variable time 40

Chapter 1

Non-Smooth Analysis

The purpose of this chapter is not that of giving a complete manual of Nonsmooth Analysis, but rather that of providing the reader with the tools which are necessary to fully understand the later chapters of this dissertation.

In particular, we are interested in the notions of *limiting normal cone*, and *Clarke's gradient*. For the purpose of this dissertation, we will consider X to be the normed space $\mathbb{R}^n$, with the usual euclidean norm.

1.1 Subdifferential Calculus

Definition 1.1. Let $f : X \to \mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$ be a given function, where X is the normed space $\mathbb{R}^n$, and let x be a point in $\mathrm{dom}(f) = \{w \in X : \ f(w) < +\infty\}$. We define a *subgradient* of f at x as an element ζ of X^* that satisfies the following *subgradient inequality*:

$$f(y) - f(x) \geqslant \langle \zeta, y - x \rangle, \ y \in X.$$

The set of all subgradients of f at x is called the *subdifferential* of f at x, and it is denoted by $\partial f(x)$. By definition it is a convex set, closed for the euclidean topology (in fact, for each $y \in X$, the set $\{\zeta \in X^* : \zeta$ satisfies the subgradient inequality$\}$ is closed and convex).

Remark. We call *affine* a function that differs from a linear functional only by a constant, that is if it has the form

$$g(y) = \langle \zeta, y \rangle + c;$$

and the linear functional ζ is called the *slope* of g.
Furthermore, the affine function

$$y \mapsto f(x) + \langle \zeta, y - x \rangle$$

is said to *support* f at x, and this means that it lies everywhere below f, and that at x we have equality.

We now want to illustrate the geometry of subgradients. First we need the notion of *epigraph*.

Definition 1.2. Given a function $f : X \to \mathbb{R}_\infty$, the *epigraph* of f is the set of points on or above the graph:

$$\mathrm{epi}(f) = \{(x, r) \in X \times \mathbb{R} : \ f(x) \leqslant r\}.$$

Example 1.3. We now wish to illustrate the geometry of subgradients with the help of Figure 1.1, which we think of as depicting the epigraph of a convex function $f : \mathbb{R}^n \to \mathbb{R}_\infty$.

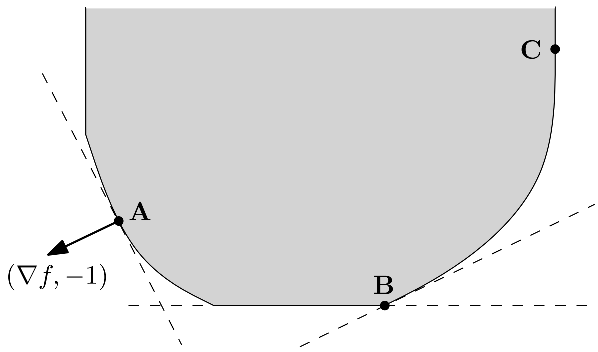

Figure 1.1: The epigraph of a convex function, and some supporting hyperplanes

The function is smooth near the point $A \equiv \big(x_1, f(x_1)\big)$ on the boundary of its epigraph. This means that there is a unique affine function $y = \langle \zeta, x \rangle + c$ that supports f at the point x_1; its slope is given by $\nabla f(x_1)$. The vector $\big(\nabla f(x_1), -1\big)$ is then orthogonal to the corresponding supporting hyperplane, and generates the normal cone to epi(f) at $\big(x_1, f(x_1)\big)$, which is, in this case, a ray.

At the point $B \equiv \big(x_2, f(x_2)\big)$, the function has a corner. The consequence is that there are infinitely many supporting affine functions of f at B; the set of all such functions constitutes $\partial f(x_2)$, by definition.

The supporting hyperplane to epi(f) at $C \equiv \big(x_3, f(x_3)\big)$ fails to define a subgradient, since it does not correspond to the graph of an affine function of x (the subdifferential of f is empty at x_3).

We now provide the following

Proposition 1.4. Let $f : X \to \mathbb{R}_\infty$ be a convex function, and $x \in \mathrm{dom}(f)$. Then

$$\partial f(x) = \big\{\zeta \in X^* : f'(x, v) \geqslant \langle \zeta, v \rangle \ \forall\, v \in X\big\}.$$

Proof. It is easy to show that a convex function f admits directional derivative, $\forall\, v \in X$,

$$f'(x, v) = \inf_{t>0} \frac{f(x + tv) + f(x)}{t},$$

since the function $g(t) = (f(x + tv) + f(x))/t$ is non-decreasing in the domain $t > 0$.

($\subseteq$) If $\zeta \in \partial f(x)$, then we have

$$f(x + tv) - f(x) \geqslant \langle \zeta, tv \rangle \ \forall\, v \in X,\ t > 0,$$

by the subgradient inequality. It follows that $f'(x, v) \geqslant \langle \zeta, v \rangle\ \forall\, v$.

($\supseteq$) Conversely, if this last condition holds, then we have

$$f(x + v) - f(x) \geqslant \inf_{t>0} \frac{f(x + tv) - f(x)}{t} \geqslant \langle \zeta, v \rangle,\ \forall\, v \in X,$$

which implies $\zeta \in \partial f(x)$. □

1.2 Generalized Gradients

The following theory serves the purpose of developing a generalized calculus on Banach spaces, which reduces to differential calculus for smooth functions. We will limit ourselves to the case in which our Banach space is $\mathbb{R}^n$.

Throughout this chapter, X denotes $\mathbb{R}^n$. First of all, we recall the following

Definition 1.5. $f : X \to \mathbb{R}$ is said to be Lipschitz of rank K near a given point $x \in X$ if, for some $\varepsilon > 0$, we have

$$|f(y) - f(z)| \leqslant K\|x - z\| \quad \forall y, z \in B(x, \varepsilon).$$

Let $f : X \to \mathbb{R}$ be *Lipschitz of rank K* near a given point $x \in X$.

These are the functions we decide to consider as an environment for our nonsmooth problems.

Definition 1.6. The *generalized directional derivative* of f at x in the direction v, denoted $f^\circ(x, v)$, is defined as follows

$$f^\circ(x, v) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{f(y + tv) - f(y)}{t},$$

where y lives in X and t is a positive scalar. Note that this definition does not presuppose the existence of any limit, that it involves only the behaviour of f arbitrarily near x, and that it differs from the traditional definition of the directional derivative in that the base point y of the difference quotient varies.

Proposition 1.7. Let f be Lipschitz of rank K near x. Then:

a. The function $v \mapsto f^\circ(x, v)$ is finite, positively homogeneous, and sub-additive on X, and satisfies $|f^\circ(x, v)| \leqslant K\|v\|,\ v \in X$

b. For every $v \in X$, the function $(u, w) \mapsto f^\circ(u, w)$ is upper semi-continuous at (x, v); the function $w \mapsto f^\circ(x, w)$ is Lipschitz of rank K on X

c. We have $f^\circ(x, -v) = (-f)^\circ(x, v),\ v \in X$

Proof. In view of the Lipschitz condition, the absolute value of the difference quotient in the definition of $f^\circ(x, v)$ is bounded by $K\|v\|$ when y is sufficiently near x and t sufficiently near 0. It follows that $|f^\circ(x, v)|$ admits the same upper bound. The fact that $f^\circ(x, \lambda v) = \lambda f^\circ(x, v)$ for any $\lambda \geqslant 0$ is immediate. We now look to prove the sub-additivity of f. We calculate:

$$f^\circ(x, v+\mu) = \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y+tv+t\mu)-f(y)}{t} \leqslant$$

$$\leqslant \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y+tv+t\mu)-f(y+t\mu)}{t} + \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y+t\mu)-f(y)}{t} = f^\circ(x,v)+f^\circ(x,\mu)$$

which proves a.

Now let $x_i \xrightarrow{n} x$ and $v_i \xrightarrow{n} v$ be arbitrary sequences converging to x and v, respectively. By definition of *upper limit*, $\forall\, i\ \exists\, y_i$ in X and $t_i \geqslant 0$ such that $\|y_i + x_i\| + t_i < \frac{1}{i}$ and

$$f^\circ(x_i; v_i) - \frac{1}{i} \leqslant \frac{f(y_i+t_iv_i)+f(y_i)}{t_i} = \frac{f(y_i+t_iv_i)+f(y_i)}{t_i} =$$

$$= \frac{f(y_i+t_iv)-f(y_i)}{t_i} + \frac{f(y_i+t_iv_i)-f(y_i+t_iv)}{t_i} \leqslant \frac{f(y_i+t_iv_i)-f(y_i)}{t_i} + K\|v_i - v\|.$$

So if we take the upper limit for $i\to\infty$ we get:

$$\limsup_{i\to\infty} f^\circ(x_i; v_i) \leqslant f^\circ(x,v),$$

which proves the first assertion of b.

We now have to prove that the function $w \mapsto f^\circ(x.w)$ is Lipschitz of rank K on X. By the Lipschitz property we have:

$$f(y+tv)-f(y) \leqslant f(y+t\mu)-f(y)+tK\|v-\mu\|$$

for all y near x and positive t near 0.
We then divide by t and consider the limits as $y \to x,\ t\downarrow 0$:

$$f^\circ(x,v) \leqslant f(y+t\mu) + K\|v-\mu\|$$

and viceversa:

$$f^\circ(x,\mu) \leqslant f(y+tv) + K\|v-\mu\|$$

proving the second assertion of b.

We now only have to prove c:

$$f^\circ(x,-v) = \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y-tv)-f(y)}{t} = \limsup_{\substack{z\to x\\ t\downarrow 0}} \frac{(-f)(z+tv)-(-f)(z)}{t}, \text{ where } z := y-tv$$

$$= (-f)^\circ(x,v).$$

□

Definition 1.8. *Clarke's generalized gradient* of the function f at x, denoted $\partial_C f(x)$, is the unique non-empty compact convex subset of X^* whose support function is $f^\circ(x,\cdot)$, that is

$$\zeta \in \partial_C f(x) \iff f^\circ(x,v) \geqslant \langle \zeta, v\rangle\ \forall\, v\in X$$

$$f^\circ(x,v) = \max\{\langle \zeta, v\rangle : \zeta \in \partial_C f(x)\}\ \forall\, v \in X.$$

Remark. Following directly from the definitions of subgradient and generalized gradient we have

$$\partial_C f(x) = \partial f^\circ(x, \cdot)(0),$$

where $\partial f^\circ(x,\cdot)(0)$ is the subdifferential of the function $w \mapsto f^\circ(x, w)$ at $x = 0$.

Theorem 1.9. If f is continuously differentiable near x then $\partial_C f(x) = \{f'(x)\}$. If f is convex and lower semicontinuous, and if $x \in \mathrm{dom}(f)$, then $\partial_C f(x) = \partial f(x)$.

The following result shows that if f is Lipschitz, its derivative can be used to generate its generalized gradient. More precisely, $\partial_C f(x)$ can be generated by the values of $\nabla f(u)$ at nearby points u at which $f'(u)$ exists. Furthermore, points u which belong to any prescribed set of zero measure can be ignored in the construction.

Theorem 1.10 (Gradient formula)**.** Let $x \in \mathbb{R}^n$, and let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near x. Let E be any subset of zero measure in $\mathbb{R}^n$, and the E_f be the set of point at which f fails to be differentiable. Then

$$\partial_C f(x) = \mathrm{co}\big\{ \lim_{i\to\infty} \nabla f(x_i) : \ x_i \to x, \ x_i \notin E \cap E_f \big\},$$

where $\mathrm{co}(S)$ is the convex hull of the set S.

1.3 Proximal Calculus

We now want to extend the use of subgradients for functions that are not necessarily convex.
We consider a given function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ a given function, $x \in \mathrm{dom}(f)$.

Definition 1.11. An element $\zeta \in \mathbb{R}^n$ is said to be a *proximal subgradient* of f at x if there exist a $\sigma \equiv \sigma(x, \zeta) \geqslant 0$, and a neighbourhood $V \equiv V(x, \zeta)$ of x, such that

$$f(y) - f(x) + \sigma\|y - x\|^2 \geqslant \langle \zeta, y - x\rangle \ \forall\, y \in V. \tag{1.1}$$

The *proximal subdifferential* of f at x is the set of all such ζ, and is denoted $\partial_P f(x)$.

For this to be an extension of the notion of subdifferential, we of course have

Proposition 1.12. Let f be convex. Then $\partial_P f(x) = \partial f(x)$.

Proof. ($\supseteq$) It follows directly from the definition of subgradient for a convex function.
($\subseteq$) Let $\zeta \in \partial_P f(x)$; we need to show that $\zeta \in \partial f(x)$.
Note that, for such a ζ, the convex function

$$y \mapsto g(y) = f(y) + \sigma\|y - x\|^2 - \langle \zeta, y\rangle,$$

by definition of proximal subgradient, has a local minimum at $y = x$. This means that $0 \in \partial g(x)$.
We then obtain $\zeta \in \partial f(x)$, as required. □

Remark. In the same way we did with subgradients of convex functions, we want to give a geometrical interpretation of proximal subgradients.
When f is convex, and element $\zeta \in \partial f(x)$ satisfies 1.1 globally, and with $\sigma = 0$. We have seen in 1.3 that this corresponds to the epigraph of f having a supporting hyperplane at $(x, f(x))$.

In the case of a generic f, however, the proximal subgradient provides only the information that *locally* f is bounded below by the function

$$y \mapsto f(x) + \langle \zeta, y - x \rangle - \sigma \|y - x\|^2.$$

The graph of this function corresponds to a downward facing parabola, passing trough the point $(x, f(x))$ with derivative ζ. Proximal subgradient are then the slopes at x of locally supporting parabolas to the epigraph of f.

We refer to Figure 1.2 for the following example.

Example 1.13. In a neighbourhood of $A = \big(x_A, f(x_A)\big)$, the function f coincides with the smooth function $x \mapsto x + k$, for some constant k. This means that there are infinitely many parabolas that locally support the epigraph at A, and they all have slope 1 at $x_A \implies \partial_P f(x_A) = \{1\}$.

Near the point $B = \big(x_B, f(x_B)\big)$ the epigraph of f is the same as that of the function

$$g(x) = \begin{cases} 0 & \text{if } x \leqslant x_B, \\ +\infty & \text{if } x > x_B. \end{cases}$$

It follows that $\partial_P f(x_B) = [0, +\infty)$, and these are the slopes at x_B of all possible locally supporting parabolas at B.

At x_C the function has a concave corner (f is locally of the form $-|x - x_C| + k$). As a consequence, no parabola can locally support the epigraph of f at $C \implies \partial_P f(x_C) = \emptyset$.

The point D, instead, corresponds to a convex corner (f is locally of the form $|x - x_D| + k$). In this case we have $\partial_P f(x_D) = \partial(|x - x_D|)(x_D) = [-1, 1]$.

At x_E, f has infinite slope, which precludes the existence of any supporting parabolas. This means that $\partial_P f(x_E) = \emptyset$.

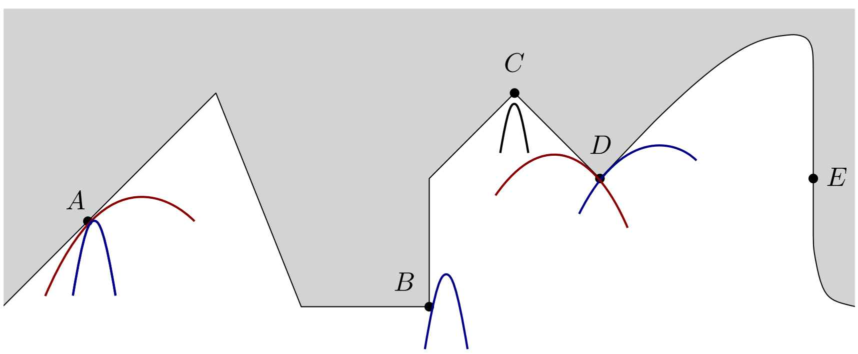

Figure 1.2: The epigraph of a function, and some locally supporting parabolas

Relation to derivatives. We may now see what the relations between proximal subgradients and derivatives are. We recall that a function $f : X \to Y$ is *Gateaux differentiable at* x if the directional derivative $f'(x,v)$ exists for all $v \in X$, and if there exists $\Lambda \in L_C(X,Y)$ such that, for all $v \in X$, we have

$$f'(x,v) = \langle \Lambda, v \rangle.$$

It follows that the element Λ is unique, and it is denoted as $f'_G(x)$, and referred to as the *Gateaux derivative* of f at x.

Proposition 1.14. If f is Gateaux differentiable at x, then $\partial_P f(x) \subseteq \{f'_G(x)\}$.

Proof. Let $v \in X$ be fixed. Setting $y = x + tv$ in the proximal subgradient inequality we obtain:

$$(f(x+tv) - f(x))/t \geqslant \langle \zeta, v \rangle - \sigma t \|v\|^2, \ \forall\, t > 0 \text{ sufficiently small.}$$

Passing to the limit for $t \downarrow 0$ we get $\langle f'_G(x), v \rangle \geqslant \langle \zeta, v \rangle, \forall v \in X$. Then we must have $\zeta = f'_G(x)$. □

Proposition 1.15. Let $x \in \mathrm{dom}(f)$, and let $g : X \to \mathbb{R}$ be differentiable in a neighbourhood of x, with g' Lipschitz near x. Then

$$\partial_P(f+g)(x) = \partial_P f(x) + \{g'(x)\}.$$

Proof. We will first prove that there exist $\delta > 0$ and a constant M such that

$$u \in B(x,\delta) \implies |g(u) - g(x) - \langle g'(x), u - x \rangle| \leqslant M\|u - x\|^2. \tag{1.2}$$

Indeed, thanks to the Lipschitz hypothesis on g' we find a $\delta > 0$ and M such that

$$y, z \in B(x,\delta) \implies \|g'(y) - g'(z)\| \leqslant M\|y - z\|^2.$$

By the mean value theorem, for any $u \in B(x,\delta)$, there exists $z \in B(x,\delta)$ such that $g(u) = g(x) + \langle g'(z), u - x \rangle$. Then, by the Lipschitz condition for g', we have

$$|g(u) - g(x) - \langle g'(x), u - x \rangle| = |\langle g'(z) - g'(x), u - x| \leqslant M\|z - x\|\|u - x\| \leqslant M\|u - x\|^2,$$

which proves the assertion.

($\subseteq$) Now let $\zeta \in \partial_P(f+g)(x)$. Then, for some $\sigma \geqslant 0$ and for a neighbourhood V of x, we have

$$f(y) + g(y) - f(x) - g(x) + \sigma\|y - x\|^2 \geqslant \langle \zeta, y - x \rangle, \ \forall\, y \in V.$$

It follows from (1.2) that

$$f(y) - f(x) - g(x) + (\sigma + M)\|y - x\|^2 \geqslant \langle \zeta - g'(x), y - x \rangle, \ \forall\, y \in V_\delta := V \cap B(x,\delta).$$

We then conclude, by definition, that $\zeta - g'(x) \in \partial_P f(x)$.

($\supseteq$) Conversely, if $\psi \in \partial_P f(x)$, then, for some $\sigma \geqslant 0$ and neighbourhood V of x, we have

$$f(y) - f(x) + \sigma\|y - x\|^2 \geqslant \langle \psi, y - x \rangle, \ \forall\, y \in V.$$

From (1.2) we deduce that

$$f(y) + g(y) - f(x) - g(x) + (\sigma + M)\|y - x\|^2 \geqslant \langle \psi + g'(x), y - x \rangle, \ \forall\, y \in V_\delta,$$

from which we conclude that $\psi + g'(x) \in \partial_P(f+g)(x)$. □

By taking $f \equiv 0$ in the Proposition above, we obtain

Corollary 1.16. Let $g : X \to \mathbb{R}$ be differentiable in a neighbourhood of x, with g' Lipschitz near x. Then $\partial_P g(x) = \{g'(x)\}$.

We now provide the following results, which we will need later on.

Theorem 1.17 (Proximal Density)**.** Let X be a Hilbert space, and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous. Let $x \in \mathrm{dom}(f)$ and $\varepsilon > 0$ given. Then there exists a point $y \in x + \varepsilon B(x, \delta), \delta > 0$, satisfying $\partial_p f(y) \neq \emptyset$ and $|f(y) - f(x)| \leqslant \varepsilon$.

Proposition 1.18. Let f be Lipschitz of rank K near x, and let x_i and ζ_i be sequences in X and X^* respectively such that

$$x_i \to x,\ \zeta_i \in \partial_C f(x_i)\ \forall\, i.$$

If ζ is a weak* cluster point of the sequence ζ_i (e.g. if $\zeta_i \to \zeta$ in X^*), then we have $\zeta \in \partial_C f(x)$.

Proposition 1.19. Let f be Lipschitz near x. Then $\partial_P f(x) \subseteq \partial_C f(x)$. If f is $\mathcal{C}^1$ near x and f' is Lipschitz near x (i.e. if f is $\mathcal{C}^2$ near x), then we have

$$\partial_P f(x) = \{f'(x)\} = \partial_C f(x).$$

Proof. An immediate consequence of the proximal subgradient inequality is that, for any given v, for all $t > 0$ sufficiently small, we have

$$(f(x+tv) - f(x))/t \geqslant \langle \zeta, v \rangle - \sigma t |v|^2.$$

We then have $f^\circ(x, v) \geqslant \langle \zeta, v \rangle$, whence ζ belongs to $\partial C f(x)$ by definition of generalized gradient.

Let f be $\mathcal{C}^1$ near x. Then by Theorem 1.9 we have $\partial_C f(x) = \{f'(x)\}$. The fact that equality holds when f' is Lipschitz near x follows from Corollary 1.16. □

Definition 1.20. The *limiting subdifferential* $\partial_L f(x)$ is defined by applying a closure operation to the proximal subdifferential:

$$\partial_L f(x) = \{\zeta = \lim_{i \to \infty} \zeta_i \ :\ \zeta_i \in \partial_P f(x_i),\ x_i \to x,\ f(x_i) \to f(x)\}.$$

In the same way we did for the proximal subdifferential, we now see the relations to derivatives.

Proposition 1.21. Let f be Lipschitz near x. Then $\emptyset \neq \partial_L f(x) \subseteq \partial_C f(x)$. If f is $\mathcal{C}^1$ near x, we have $\partial_L f(x) = \{f'(x)\}$. If f is convex, we have $\partial_L f(x) = \partial f(x)$.

Proof. Thanks to Theorem 1.17, we have that there exist points x_i converging to x, which admit elements $\zeta_i \in \partial_P f(x_i)$ such that $|f(x_i) - f(x)| < 1/i$. Let K be a Lipschitz constant for f in a neighbourhood of x. Then, for all i sufficiently large, we have

$$\zeta_i \in \partial_P f(x_i) \subseteq \partial_C f(x_i) \subseteq B(0, K).$$

Thus ζ_i is a bounded sequence, and we can then suppose (by passing to a subsequence) that $\zeta_i \to \zeta$ for some $\zeta \in \partial_L f(x)$.
Thanks to Propositions 1.18 and 1.21, we have that $\partial_L f(x) \subseteq \partial_C f(x)$.

Let us now suppose $f \in \mathcal{C}^1$ near x. We know that in this case

$$\emptyset \neq \partial_L f(x) \subseteq \partial_C f(x) = \{f'(x)\},$$

which implies $\partial_L f(x) = \{f'(x)\}$.

Let f be convex. Then we can apply Proposition 1.12 and get $\partial_P f(x) = \partial f(x)$. This means, since by definition of $\partial_L f(x)$, $\zeta = \lim_{i\to\infty} \zeta_i$, that

$$f(y) - f(x - i) \geqslant \langle \zeta_i, y - x_i \rangle, \ \forall\, y \in \mathbb{R}^n.$$

Passing to the limit we get that $\zeta \in \partial f(x) \implies \partial_L f(x) \subseteq \partial f(x) = \partial_P f(x) \subseteq \partial_L f(x)$, from which we conclude. □

1.4 Proximal Geometry

Let S be a non-empty closed subset of $\mathbb{R}^n$, $x \in S$; we are going to introduce the notion of *proximal normal and limiting normal cone* to S at x.

Definition 1.22 (proximal normal cone)**.** Let $x \in S$. A vector $\zeta \in \mathbb{R}^n$ is said to be a *proximal normal* to S at x if and only if there exists a constant $\sigma = \sigma(x,\zeta) \geqslant 0$ such that

$$\langle \zeta, u - x \rangle \leqslant \sigma |u - x|^2 \ \forall\, u \in S. \tag{1.3}$$

The set of all such ζ, denoted $N_S^P(x)$, defines the *proximal normal cone* to S at x.

Despite the global nature of the proximal normal inequality (1.3), proximal normals are a local construct. This is shown in the following

Proposition 1.23. Suppose that there exist a $\sigma \geqslant 0$ and a $\delta > 0$ such that

$$\langle \zeta, u - x \rangle \leqslant \sigma |u - x|^2 \ \forall\, u \in S \cap B(x,\delta).$$

Then $\zeta \in N_S^P(x)$.

Proof. Let $\zeta \notin N_S^P(x)$. Then for each $i \in \mathbb{Z}$ there is a point $u_i \in S$ such that $\langle \zeta, u_i - x \rangle > i|u_i - x|^2$. It follows that $u_i \to x$. But then, for i sufficiently large, we have $|u_i - x| < \delta$ and $i > \sigma$.
From this we get

$$\sigma |u_i - x|^2 < i|u_i - x| < \langle \zeta, u_i - x \rangle,$$

and this concludes the proof. □

Proposition 1.24. Let $x \in S$. Then

$$\partial_P I_S(x) = N_S^P(x) = \{\lambda\zeta : \ \lambda \geqslant 0, \ \zeta \in \partial_P d_S(x)\},$$

where $d_S(x) = \inf_{y\in S} \|y - x\|$ and $I_S(x) = \begin{cases} 0 & \text{if } x \in S \\ +\infty & \text{if } x \notin S \end{cases}$.

Proof. Observe that the first equality is a direct consequence of the previous Proposition. We then turn to the second one.

Let $\zeta \in N^P_S(x)$. From the proximal normal inequality we get that, for a certain $\sigma \geqslant 0$, the function

$$y \mapsto \phi(y) := -\langle \zeta, y\rangle + \sigma|y-x|^2$$

attains a minimum relative to S at $y = x$. Fix any $\varepsilon > 0$. Since, sufficiently near x, the function ϕ is Lipschitz of rank $|\zeta| + \varepsilon$, thanks to a result on generalized gradients, we have that x is a local minimum of the function $\phi(y) + (|zeta| + \varepsilon)d_S(y)$.
From this we get

$$0 \in \partial_P\big\{-\langle \zeta, y\rangle + \sigma|y-x|^2 + (|\zeta|+\varepsilon)d_S(y)\big\} = -\zeta + (|\zeta|+\varepsilon)\partial_P d_S(x),$$

by Proposition . Thus $\dfrac{\zeta}{|\zeta|+\varepsilon} \in \partial_P d_S(x)$ for every $\varepsilon > 0$.

The opposite inclusion follows easily from the definition of $\zeta \in \partial_P d_S(x)$, which immediately implies the proximal normal inequality. □

A geometrical interpretation. We will now show how proximal normals to a point x correspond to *closest point* directions emanating outwards from x, and that they are generated by the projection onto the set.

We recall the notion of projection of a point y onto the set S, as

$$P_S(y) = \big\{s \in S : |y-s| = d_S(y)\big\}.$$

Proposition 1.25. A nonzero vector ζ satisfies the proximal normal inequality (1.3) if and only if $x \in P_S(y)$, where $y := x + \dfrac{\zeta}{2\sigma}$.

Proof.

$$\begin{aligned} x \in P_S\left(x + \frac{\zeta}{2\sigma}\right) &\implies \left|\frac{\zeta}{2\sigma}\right| \leqslant \left|x + \frac{\zeta}{2\sigma} - z\right| \ \forall z \in S \implies \\ &\implies \left|\frac{\zeta}{2\sigma}\right|^2 \leqslant \left|x + \frac{\zeta}{2\sigma} - z\right|^2 \ \forall z \in S \implies \\ &\implies 0 \leqslant |x-y|^2 + \left\langle \frac{\zeta}{\sigma}, x-y\right\rangle \ \forall z \in S \implies \\ &\implies \langle \zeta, x-y\rangle \leqslant \sigma\,|x-y|^2 \ \forall z \in S. \end{aligned}$$

□

Remark. More generally, $\zeta \in N^P_S(x) \iff \exists\, y \notin S : \ x \in P_S(y)$ and $\zeta = t(y-x)$ for some $t > 0$.

This characterization of $N^S_P(x)$ allows us to give a geometrical interpretation of the proximal normal, as seen in Figure 1.3.

Definition 1.26 (limiting normal cone)**.** We define the *limiting normal cone* as the set

$$N^L_S(x) = \{\zeta = \lim_{i\to\infty} \zeta_i \ : \ \zeta_i \in N^P_S(x_i), x_i \xrightarrow{i\to\infty} x, \ x_i \in S\}.$$

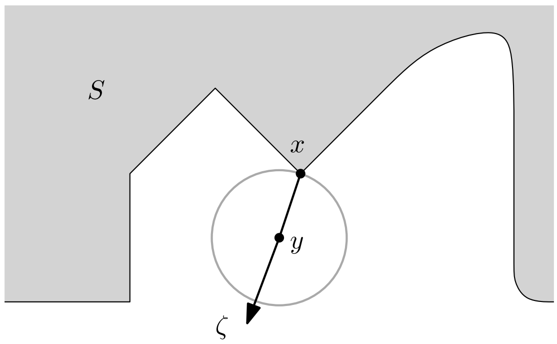

Figure 1.3: A proximal normal direction ζ to S at x

That is, the limiting normal cone is defined as the closure of the proximal normal cone.

In much the same way as for the proximal normal cone in Proposition 1.24, we have the following for the limiting normal cone:

Proposition 1.27. Let $\xi \in S$. Then

$$\zeta \in N_S^L(x) \iff \zeta \in |\zeta| \partial_L d_S(x),$$

and

$$N_S^L = \left\{ \lambda\zeta : \lambda \geqslant 0, \ \zeta \in \partial_L d_S(x) \right\} = \partial_L I_S(x).$$

Chapter 2

Optimal Control

Differential equations have proved to be an effective mathematical model to describe a wide range of physical phenomena. Systems of the form

$$\dot{x}(t) = f(t, x(t)) \tag{2.1}$$

are routinely used in areas as diverse ad aeronautics, robotics, economics and natural resources. If the state of the system is known at some initial time t_0, the future behaviour for $t > t_0$ can be determined solving what is known as a *Cauchy Problem*, consisting of (2.1), combined with the initial condition

$$x(t_0) = x_0 \tag{2.2}$$

In this case, we are taking the *spectator*'s point of view: the mathematical model only allows us to understand and predicts the evolution of a portion of the physical world; we have no means of altering its behaviour.

Control theory works in a different way: in this paradigm, we assume the presence of an external agent, who can actively influence the evolution of the system.
This is done by introducing an explicit *control variable* in the differential equation, that can be chosen as to attain a certain preassigned goal - steer the system from one state to another, maximize the value of a certain parameter or minimize a certain cost functional, etc..

And this is the main object of this dissertation, the *controlled differential equation*

$$\dot{x}(t) = f(t, x(t), u(t)), \ u(t) \in U, \tag{2.3}$$

where u is a measurable function; the couple (f, U) is known as the *control system*.

Example 2.1 (boat on a river)**.** [1] Consider a river with a straight course.
Using a set of planar coordinates, assume that it occupies the horizontal strip

$$S = \{(x_1, x_2) : \ (x_1, x_2) \in \mathbb{R}^2, \ -1 \leqslant x_2 \leqslant 1\}.$$

Moreover, assume that the speed of the water is given by the velocity vector $\mathbf{v}(x_1, x_2) = (1 - x_2^2, 0)$.

If a boat on the river is merely dragged along by the current, its position will be determined by the differential equation

$$(\dot{x}_1, \dot{x}_2) = (1 - x_2^2, 0).$$

On the other hand, if the boat is powered by an engine, then its motion can be modelled by the control system

$$(\dot{x}_1, \dot{x}_2) = \mathbf{v} + \mathbf{u} = (1 - x_2^2 + u_1, u_2),$$

where the vector $\mathbf{u} = (u_1, u_2)$ describes the velocity of the boat relative to the water. The set of admissible controls consists of all measurable functions $u : \mathbb{R} \to \mathbb{R}^2$ taking values inside the closed disc

$$\mathbf{U} = \left\{ \omega_1, \omega_2) : \ \sqrt{\omega_1^2 + \omega_2^2} \leqslant M \right\},$$

where the constant M accounts for the maximum speed (in any direction) that can be produced by the engine.

We wish to choose the control u in such a way as to minimize the cost functional $J(x, u)$ defined by

$$J(x, u) = l(x(b)) + \int_a^b \Lambda(t, x(t), u(t))dt, \tag{2.4}$$

subject to the endpoint constraint $x(b) \in E$, where $E \subseteq \mathbb{R}^n$ is the *target set*, and Λ (the *running cost*) and l (*endpoint cost*) are given functions.

So we have the standard *Optimal Control Problem*:

$$\begin{cases} \text{minimize} & J(x, u) = l(x(b)) + \displaystyle\int_a^b \Lambda(t, x(t), u(t))dt \\ \text{subject to} & \dot{x}(t) = f(t, x(t), u(t)), \ t \in [a, b] \ a.e. \\ & u : [a, b] \to U, \text{ measurable} \\ & x(a) = x_0, \ x(b) \in E \end{cases} \tag{OC}$$

A *process* is a couple (x, u) that satisfies (OC).

We will work under what are called the *classical regularity hypotheses*:

* the function l is continuously differentiable;
* the functions f and Λ are continuous, and admit derivatives relative to x which are continuous in all variables (t, x, u).

The purpose of this dissertation is that of providing the reader with a set of necessary conditions for the solution of the *Optimal Control Problem*.
This result is known as the *Maximum Principle*. We will discuss two different versions of the Maximum Principle: the classic Pontryagin one, which requires the *classical regularity hypotheses* to be valid, and a Variable Time Principle, which, under the same hypotheses, does not require for the endtime b to be fixed.

Classically, this problems are studied in the case of a target set E that is a smooth manifold. Thanks to the tools we have prepared in the previous chapter, however, we can extend the classical theory to the nonsmooth case. We choose to consider here Clarkes' version of the Pontryagin Maximum Principle, where the target set E is any closed subset of $\mathbb{R}^n$. The cases in which the target set is either a point, the whole space, a smooth manifold, or a manifold with boundary will then all be a special case of what we have studied.

We chose not to consider here more refined versions of the Pontryagin Maximum Principle, allowing the dynamics to be nonsmooth.

Optimal Control owes its origin to the *Calculus of Variations*, as many of the current developments in the field have resulted from marrying old ideas from the Calculus of Variations and modern analytical techniques, as the ones found in the first chapter of this work.
Although they should be familiar to most readers, let us review some of the classic results of the Calculus of Variations.

2.1 The Calculus of Variations

The basic problem in what is known as the *calculus of variations* is that of finding an arc $\bar{x}$ which minimizes the value of an integral functional of the form

$$J(x) = \int_a^b \Lambda(t, x(t), x'(t))\, dt$$

over some class of arcs x defined on the interval $[a, b]$ and which take prescribed values at a and at b, where $[a, b]$ is a given interval in $\mathbb{R}$, and where $\Lambda = \Lambda(t, x, v)$ is a function of three variables (time, state, velocity) referred to as the *Lagrangian*.

Remark. Notice that this is a particular case of Optimal Control Problem, in which we minimize the funtional $J = \int_a^b \Lambda(t, x, u)dt$, with $x' = u$.

Example 2.2 (The Brachistochrone Problem [3]). The following, circulated by Johann Bernoulli in the late 17th century, is an early example of such a problem. Positive numbers s_f and x_f are given. A frictionless bead, initially located at the point $(0, 0)$, slides along a wire under the force of gravity. The wire, which is located in a fixed vertical plane, joins the points $(0, 0)$ and (s_f, x_f).

The question is: what should the shape of the wire be, for the beam to arrive at the point (s_f, x_f) in the least amount of time?

This is also a particular case of Control system: minimize $\int_0^{s_f} u(t, x, u)dt$, with $x' = u$.

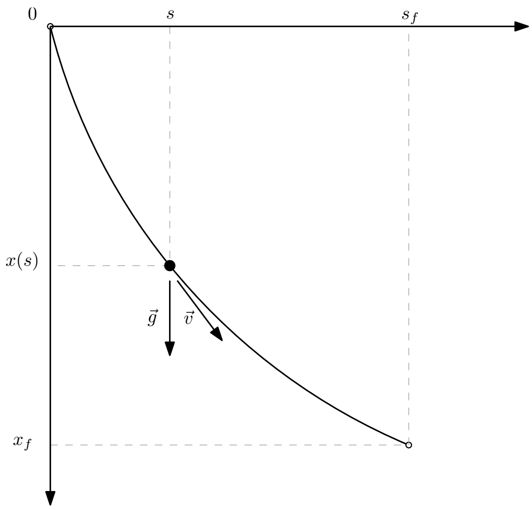

Figure 2.1: The Brachistochrone Problem

Denote by s and x the horizontal and vertical distances of a point on the path of the bead (vertical distances are measured downward). We restrict our attention to wires describable as the graph of a suitably regular function $x(s)$, $0 \leqslant s \leqslant s_f$. For any such function x, the velocity $v(s)$ is related to the downward displacement $x(s)$, when the horizontal displacement is s, according to the formula

$$mgx(s) = \frac{1}{2}mv^2(s).$$

Denoting the time variable as t, we have

$$v(s) = \frac{\sqrt{1 + |dx(s)/ds|^2}}{dt(s)/ds},$$

and by integrating we get

$$\int_0^{s_f} dt = \int_0^{s_f} \frac{\sqrt{1 + |dx(s)/ds|^2}}{v(s)} ds.$$

By eliminating v in the preceding expression and setting

$$L(s, x, w) := \frac{\sqrt{1 + |w|^2}}{\sqrt{2gx}},$$

we arrive at the following formula for the transit time:

$$J(x) = \int_0^{s_f} L\big(s, x(s), x'(s)\big) ds.$$

The problem is then that of minimizing the functional $J(x)$ over some class of arcs x satisfying $x(0) = 0$ and $x(s_f) = x_f$. This is an example of the *basic problem*.

Suppose that we seek a minimizer in the class of absolutely continuous arcs. It can be shown that the minimum time t^* and the minimizing arc $\big(x(t), s(t)\big)$, $0 \leqslant t \leqslant t^*$ (expressed in parametric form with independent variable time t) are given by the formulae

$$x(t) = a\left(1 - \cos\sqrt{\frac{g}{a}}t\right)$$
$$s(t) = a\left(\sqrt{\frac{g}{a}}t - \sin\sqrt{\frac{g}{a}}t\right),$$

where the constants a and t^* are uniquely determined by the conditions

$$x(t^*) = x_f,$$
$$s(t^*) = t_f,$$
$$0 \leqslant \sqrt{\frac{g}{a}} \leqslant 2\pi.$$

The minimizing curve is a *cycloid*, with infinite slope at the point of departure: it coincides with the locus of a point on the circumference of a disc of radius a, which rolls without slipping along a line of lenght t_f.

The Basic Problem. In this chapter we will work under the following hypotheses:

* the variables of time, space and velocity be one-dimensional
* $\Lambda : \mathbb{R}^3 \to \mathbb{R}$ be a twice continuously differentiable function
* $x : [a, b] \to \mathbb{R}$ be in $C^2[a, b]$

and we will consider the *Basic Problem*:

$$\text{minimize } J(x) : \ x \in C^2[a, b], \ x(a) = A, \ x(b) = B. \tag{B}$$

A function $x : [a, b] \to \mathbb{R}$ is said to be *admissible* if it lies in $C^2[a, b]$ and satisfies all of the boundary constraints in (B). $J(x)$ is called the *cost* corresponding to the function x.
A *solution* x_* of (B) (also referred to as a *minimizer* for the problem) is an unmissable function x_* such that

$$J(x_*) \leqslant J(x) \text{ for all other admissible functions } x$$

Necessary conditions for the basic problem. We introduce the notion of a *variation*, that is of a function $h \in C^2[a, b]$ such that $h(a) = h(b)$. The *Gateaux derivative* of the integral functional J in the direction of $h(\cdot)$ is defined as

$$dJ[x(\cdot)]h(\cdot) := \frac{d}{d\lambda}J(x(\cdot) + \lambda h(\cdot))_{|\lambda=0}$$

Theorem 2.3 (Euler - 1744)**.** If $\bar{x}$ is a solution of (B), then $\bar{x}$ satisfies the so called ***Euler equation***

$$\frac{d}{dt}\left\{\Lambda_v(t,\bar{x}(t),\bar{x}'(t))\right\} = \Lambda_x(t,\bar{x}(t),\bar{x}'(t))\ \forall\ t \in [a,b] \tag{2.5}$$

Proof. The proof uses the idea of *variation*. If we define the single variable function

$$g(\lambda) = J(\bar{x} + \lambda h) \tag{2.6}$$

we get that the Gateaux derivative is equal to

$$0 = \frac{d}{d\lambda} g(\lambda)_{|\lambda=0} = \frac{d}{d\lambda} J(\bar{x} + \lambda h)_{|\lambda=0} = \int_a^b [\Lambda_x(t,\bar{x}(t),\bar{x}')h + \Lambda_{x'}(t,\bar{x},\bar{x}')h']\ dt$$

We now set $\alpha(t) = \Lambda_x(t,\bar{x}(t),\bar{x}')$ and $\beta(t) = \Lambda_{x'}(t,\bar{x},\bar{x}')$ and we have:

$$\int_a^b [\alpha(t)h(t) + \beta(t)h(t)]\ dt = 0$$

Integrating by parts we then get:

$$\int_a^b [\alpha(t) - \beta'(t)]h(t)\ dt = 0$$

And the conclusion follows. □

A function $x \in \mathcal{C}^2([a,b])$ which satisfies Euler's equation is referred to as an *extremal*.

We wish now to develop an analogous result when looking for *local* minima of the integral functional J. A function $\bar{x}$, admissible for (B) is said to provide a *weak local minimum* if, for some $\varepsilon > 0$, for all admissible $x : \|x - \bar{x}\| \leqslant \varepsilon$ and $\|x' - \bar{x}'\| \leqslant \varepsilon$, we have

$$J(x) \geqslant J(\bar{x}).^1$$

The proof of Theorem 2.3 can still be followed for a local minimizer, with the only difference being that the function g will attain a *local* minimum at 0, rather than a global one.

We say that the Lagrangian Λ is *autonomous* if it has no explicit dependence on the t variable. The following necessary condition for extremals is a consequence of the Euler equation.

Proposition 2.4. Let $\bar{x}$ be a weak local minimizer for B, where Λ is autonomous. Then $\bar{x}$ satisfies the *Erdmann condition*: for some constant h, we have

$$\bar{x}'(t) \cdot \Lambda_v(\bar{x}(t),\bar{x}'(t)) - \Lambda(\bar{x}(t),\bar{x}'(t)) = h\ \forall\ t \in [a,b].$$

[1] the norm will always be that of $L^\infty[a,b]$. Thus $\|x - \bar{x}\| = \max\{|x(t) - \bar{x}(t)| : t \in [a,b]\}$.

Proof. The conclusion follows once we have proven that the derivative of the function on the left side is zero. This is a consequence of the Euler equation.

$$\begin{aligned}
&\frac{d}{dt}\Big(\bar{x}'(t)\cdot\Lambda_v\big(\bar{x}(t),\bar{x}'(t)\big)-\Lambda\big(\bar{x}(t),\bar{x}'(t)\big)\Big)=\\
&=\bar{x}''(t)\cdot\Lambda_v\big(\bar{x}(t),\bar{x}'(t)\big)+\bar{x}'(t)\cdot\frac{d}{dt}\Big\{\Lambda_v\big(\bar{x}(t),\bar{x}'(t)\big)\Big\}-\frac{d}{dt}\Big\{\Lambda\big(\bar{x}(t),\bar{x}'(t)\big)\Big\}=\\
&=\bar{x}''(t)\cdot\Lambda_v\big(\bar{x}(t),\bar{x}'(t)\big)+\bar{x}'(t)\cdot\Lambda_x\big(t,\bar{x}(t),\bar{x}'(t)\big)+\\
&\quad-\bar{x}'(t)\cdot\Lambda_x\big(\bar{x}(t),\bar{x}'(t)\big)-\bar{x}''(t)\cdot\Lambda_v\big(\bar{x}(t),\bar{x}'(t)\big)=0.
\end{aligned} \tag{2.7}$$

□

Example 2.5 (A minimal surface problem)**.** Max lives in Prague; he likes to watch as street artists create soap bubbles in the square. One day a juggler does something which Max has never seen: he picks up two rings, submerges them in soap, and then creates a beautiful surface made out of soap.

A classical example of the basic problem is that of finding the shape of the curve $r(x)$ joining (a, A) to (b, B) whose associated surface of rotation has minimal area.

When a soap surface is spanned by two concentric rings of radius A and B, the resulting surface will be a surface of rotation of a curve $r(x)$, and the area of the surface will be a minimum[2].

Let OP be a generic point of the surface. In cylindrical coordinates we have

$$OP(x,y,z)\equiv OP(x,\phi)=\begin{cases} x &= x,\\ y &= r(x)\cos\phi,\\ z &= r(x)\sin\phi.\end{cases}$$

The surface element is

$$d\mathcal{A}=\left|\frac{\partial OP}{\partial x}\times\frac{\partial OP}{\partial \phi}\right|=r(x)\sqrt{1+r'(x)^2}$$

Causing the total area of the surface to be

$$\mathcal{A}=\int_0^{2\pi}d\mathcal{A}d\phi=2\pi\int_a^b r(x)\sqrt{1+r'(x)^2}dx$$

So the problem now consists of the following

$$\begin{gathered}\text{minimize}\int_a^b r(x)\sqrt{1+r'(x)^2}dx\\ \text{subject to } r(a)=A,\ r(b)=B.\end{gathered}$$

[2]thanks to *D'Alembert's principle* for potential energy, which affirms that in static equilibrium, the observed configuration minimizes total potential energy

Or, in the notations of the basic problem,

$$\text{minimize} \int_a^b x(t)\sqrt{1+x'(t)^2}dt$$
$$\text{subject to } x(a)=A,\ x(b)=B.$$

That is, we have $\Lambda(t,x,v)=x\sqrt{1+v^2}$.

Suppose that $\bar{x}$ is a local minimizer for the problem, with $\bar{x}(t)>0$ for all t. The Euler equation is given by

$$\begin{aligned} 0 &= \frac{d}{dt}\left\{\Lambda_v(\bar{x}(t),\bar{x}'(t))\right\} - \Lambda_x(t,\bar{x}(t),\bar{x}'(t)) = \\ &= \frac{d}{dt}\frac{\bar{x}\bar{x}'}{\sqrt{1+\bar{x}'^2}} - \sqrt{1+\bar{x}'^2} = \frac{\bar{x}'^4+\bar{x}'^2+\bar{x}\bar{x}''}{(1+\bar{x}'^2)^{3/2}} - \sqrt{1+\bar{x}'^2} = \\ &= \frac{\bar{x}'^4+\bar{x}'^2+\bar{x}\bar{x}''-(1+\bar{x}'^2)^2}{(1+\bar{x}'^2)^{3/2}}. \end{aligned}$$

This gives us the equivalent

$$\begin{aligned} &0 = \bar{x}'^4+\bar{x}'^2+\bar{x}\bar{x}''-1-\bar{x}''^4-2\bar{x}'^2 = -1-\bar{x}'^2+\bar{x}\bar{x}' \iff \\ &\iff \bar{x}''(t) = \left(1+\bar{x}'(t)^2\right)/x(t). \end{aligned}$$

We deduce from this that $\bar{x}'$ is strictly increasing (thus $\bar{x}$ is strictly convex). Since Λ is autonomous, we may invoke the Erdmann condition (Proposition 2.4), yielding the existence of a constant h such that

$$\begin{aligned} &\bar{x}'\frac{\bar{x}\bar{x}'}{\sqrt{1+\bar{x}'^2}} - \bar{x}\sqrt{1+\bar{x}'^2} = h \iff \\ &\iff \left(\bar{x}'(t)\right)^2 = \frac{\bar{x}^2(t)}{h^2} - 1,\ \forall\, t\in[a,b]. \end{aligned}$$

We immediately notice that the constant $\bar{x}=h$ is a solution, corresponding to the cylinder of radius $R=h$.

By separation of variables we get, if we assume $\bar{x}'$ to be positive throughout $[a,b]$,

$$\frac{h\,dx}{\sqrt{x^2-h^2}} = dt \implies \bar{x}(t) = h\cosh\left(\frac{t+c}{h}\right).$$

This type of curve is called a *catenary*.

If $\bar{x}'$ is negative on $[a,b]$, instead, we get, in much the same way,

$$\bar{x}(t) = k\cosh\left(\frac{t+d}{k}\right),$$

for some constants k and d, at first sight different from h and c.

In the general case, we have that $\bar{x}'$ is negative up to a certain $\tau \in [a, b]$, and then positive thereafter. This means that $\bar{x}$ is a catenary (with constants k, d), followed by another catenary (with constants h, c). It can be shown that the smoothness of $\bar{x}$ forces the constants to coincide; we can then simply assert that the solution to the minimal surface problem is a catenary.
Such a catenary exists if the two rings are not too far apart with respect to their radius. Otherwise the soap bubble "breaks"; this may be interpreted as if the solution is *nonsmooth*[3].

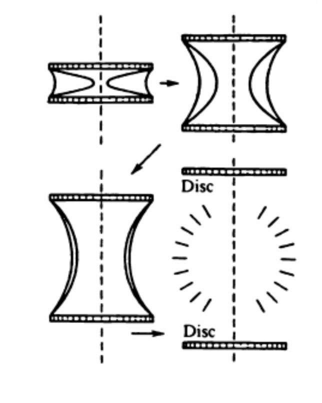

Figure 2.2: Minimal surface solutions

We now wish to study second order conditions. To do so, we will strengthen the regularity hypotheses on the Lagrangian, by assuming that Λ is C^3. We can then provide the following

Theorem 2.6 (Legendre's necessary condition - 1876)**.** Let x_* be a local minimizer for (B). Then

$$\Lambda_{vv}(t, \bar{x}(t), \bar{x}'(t)) \geqslant 0 \; \forall \; t \in [a, b].$$

The transversality condition. We now focus on variational problems in which the endpoint values of the functions x are not fully prescribed. This extra flexibility at the boundary gives us what is known as the *transversality conditions*.
Let us consider the problem of minimizing

$$l(x(b)) + \int_a^b \Lambda\big(t, x(t), x'(t)\big)\, dt$$

over functions $x \in \mathcal{C}^2\big([a, b]\big)$, satisfying the initial condition $x(a) = A$. The given function l (which we take to be continuously differentiable) corresponds do an extra cost term that depends on the value of $x(b)$, which is not prescribed.

[3]when $a = -l$, $b = l$, $A = B = R$, the problem is of the form $\frac{R}{l}\zeta = \cosh \zeta$. Depending on the length of l we have the pictured situations

Theorem 2.7. Let $\bar{x}$ be a weak local minimizer of the above problem. Then $\bar{x}$ is an extremal for Λ, and $\bar{x}$ also satisfies the following *transversality condition*:

$$-\Lambda_v\big(b, \bar{x}(b), \bar{x}'(b)\big) = l'\big(\bar{x}(b)\big).$$

Proof. If we impose $B = \bar{x}(b)$, it is clear that $\bar{x}$ is a weak local minimizer for the original problem B. Thanks to Theorem 2.3 we then have that $\bar{x}$ is an extremal.

Let us now choose any function $y \in \mathcal{C}^2\big([a,b]\big)$ for which $y(a) = 0$. We define g as follows:

$$g(\lambda) = l\big(\bar{x}(b) + \lambda y(b)\big) + J(\bar{x} + \lambda y),$$

where, as usual, we have set $J(x) = \int_a^b \Lambda\big(t, x(t), x'(t)\big)\,dt$.
It follows that g has a local minimum at $\lambda = 0$; thus $g'(0) = 0$. Following the proof of Theorem 2.3, we get

$$l'\big(\bar{x}(b)\big)y(b) + \int_a^b \big[\alpha(t)y(t) + \beta(t)y'(t)\big]\,dt = 0.$$

From the Euler equation we have $\alpha = \beta'$, and integrating by parts we get

$$\int_a^b \big[\alpha(t)y(t) + \beta(t)y'(t)\big]\,dt = \beta(b)y(b).$$

Therefore we derive

$$\big[l'\big(\bar{x}(b)\big)\big]y(b) = 0.$$

For the arbitrariness of $y(b)$, we deduce $l'\big(\bar{x}(b)\big) + \beta(b) = 0$, which is our conclusion. □

Generally, the solution to a problem of minimum in the Calculus of Variations does not belong in $\mathcal{C}^2$ nor in $\mathcal{C}^1$, as seen in the following

Example 2.8. Consider the problem

$$\text{minimize} \int_0^1 (\dot{x}^2 - 1)^2 dx : \; x(0) = 0, \; x(1) = 0.$$

It does not have a minimum in $\mathcal{C}^1$, but instead has a Lipschitz minimum, as in the Figure below.

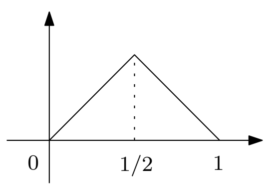

Figure 2.3: Lipschitz minimum

The appropriate space in which to study the existence of minimums is, in fact, the space of AC functions.

We then consider the basic problem

$$\text{minimize } J(x): \ x \in AC\big([a,b]\big),\ x(a) = A,\ x(b) = B, \tag{P}$$

The following Theorem is extremely useful in this regard:

Theorem 2.9 (Tonelli, 1915)**.** Let the Lagrangian $\Lambda(t,x,v)$ be continuos, convex, and coercive of degree $r > 1$: for certain constants $\alpha > 0$ and β we have

$$\Lambda(t,x,v) \geqslant \alpha|v|^r + \beta \ \forall\ (t,x,v) \in [a,b] \times \mathbb{R}^n \times \mathbb{R}^n.$$

Then the basic problem (P) admits a solution in the class $AC\big([a,b]\big)$.

Autonomous Lagrangians. The purpose of this paragraph is that of proving that, under certain hypotheses, solutions to the problem (P), which we have seen to be AC a priori, are Lipschitz.

In order to do so, we will need the following notions:

Definition 2.10. A function $\bar{x}$, admissible for the basic problem (B), is said to provide a *strong local minimum* if there exists an $\varepsilon > 0$ such that, for all admissible functions x satisfying $\|x - \bar{x}\| \leqslant \varepsilon$, we have $J(x) \geqslant J(\bar{x})$.

Remark. We observe that a strong local minimizer is automatically a weak local minimizer.

Definition 2.11. We say that a Lagrangian Λ has *Nagumo growth* along $\bar{x}$ if there exists a function $\theta : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$, satisfying $\lim_{t\to+\infty} \dfrac{\theta(t)}{t} = +\infty$, such that

$$t \in [a,b],\ v \in \mathbb{R}^n \implies \Lambda\big(t, \bar{x}(t), v\big) \geqslant \theta\big(\,|v|\,\big).$$

The following result, due to Clarke and Vinter, is a fundamental result of the Calculus of Variations for autonomous Lagrangians, and is a great is a great application of Nonsmooth Analysis.

A Lagrangian Λ is said to exhibit the *Lavrentiev phenomenon* if the infimum taken over the set of absolutely continuous trajectories is strictly lower than the infimum taken over the set of Lipschitzian trajectories, with fixed boundary conditions. The occurrence of this phenomenon prevents the possibility of computing the minimum.

This Theorem gives a set of sufficient condition for this phenomenon not to occur and, as such, is of enormous importance:

Theorem 2.12 (Clarke-Vinter, 1985)**.** Let $\bar{x} \in AC\big([a,b]\big)$ be a strong local minimizer for the basic problem

$$\text{minimize } J(x): \ x \in AC\big([a,b]\big),\ x(a) = A,\ x(b) = B, \tag{P}$$

where the Lagrangian is continuous, autonomous, convex in v, and has Nagumo growth along $\bar{x}$. Then $\bar{x}$ is Lipschitz.

Before attempting to prove this result, we need some instruments from the *multipliers theory*.

The multiplier rule. Consider the following problem of optimization, in the convex case:

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & g(x) \leqslant 0, \\ & h(x) = 0, \\ & x \in S. \end{cases} \tag{C}$$

where S is a convex subset of a real vector space X, the functions

$$f : S \to \mathbb{R} \ \text{ and } \ g_i : S \to \mathbb{R}\ (i = 1, \dots, m)$$

are convex, and the functions

$$h_j : S \to \mathbb{R},\ j = 1, \dots, n$$

are affine.

The following is a Theorem of the Multiplier rule where, as opposed to the classic version (see Theorem 9.1 in [5]), we do not require for $\bar{x}$ to lie in the interior of S.

Theorem 2.13 (Kuhn-Tucker)**.** Let $\bar{x}$ be a solution of (C). Then there exists $(\eta, \gamma, \lambda) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^n$, referred to as a *multiplier in the convex sense*, satisfying:

the **nontriviality condition**

$$(\eta, \gamma, \lambda) \neq 0,$$

the **positivity** and **complementary slackness conditions**

$$\eta = \begin{cases} 0 \\ 1 \end{cases},\ \gamma \geqslant 0,\ \big\langle \gamma, g(\bar{x}) \big\rangle = 0,$$

and the **minimization condition**

$$\big\{\eta f + \langle \gamma, h \rangle\big\}(x) \geqslant \big\{\eta f + \langle \gamma, g \rangle + \langle \lambda, h \rangle\big\}(\bar{x}) = \eta f(\bar{x})\ \forall\, x \in S.$$

We are now ready to prove Theorem 2.12.

Proof. Let $\bar{x}$ be a solution of (P) relative to $\|x - \bar{x}\| \leqslant \varepsilon$. By uniform continuity, there exists $\delta \in (0, 1/2)$ such that

$$t, \tau \in [a, b],\ |t - \tau| \leqslant (b - a)\delta/(1 - \delta) \implies |\bar{x}(t) - \bar{x}(\tau)| \leqslant \varepsilon.$$

Step 1. We consider a measurable function $\alpha : [a, b] \to [1 - \delta, 1 + \delta]$, satisfying the equality $\int_a^b \alpha(t)\,dt = b - a$.

For any such α, the relation $\tau(t) = a + \int_a^t \alpha(s)\,ds$ defines a bi-Lipschitz one-to-one mapping of $[a, b]$ onto itself; it follows that the inverse mapping $t(\tau)$ satisfies

$$\frac{d}{d\tau} t(\tau) = \frac{1}{\alpha\big(t(\tau)\big)},\ |t(\tau) - \tau| \leqslant (b - a)\frac{\delta}{1 - \delta}\ \text{a.e.}.$$

We now proceed to define and arc $y(\tau) = \bar{x}\big(t(\tau)\big)$. Then y is admissible for the problem (P), and satisfies $\|y - \bar{x}\| \leqslant \varepsilon$. But $\bar{x}$ is a strong local minimizer, and so we get

$$\int_a^b \Lambda\big(y(\tau), y'(\tau)\big)\,d\tau \geqslant J(\bar{x}).$$

Applying the change of variables $\tau = \tau(t)$ to the integral on the left, and since $y'(\tau) = x'\big(t(\tau)\big)/\alpha\big(t(\tau)\big)$ a.e., we obtain:

$$\int_a^b \Lambda\big(y(\tau), y'(\tau)\big)\,d\tau = \int_a^b \Lambda\Big(y\big(\tau(t)\big), y'\big(\tau(t)\big)\Big)\tau'(t)\,dt =$$
$$= \int_a^b \Lambda\big(\bar{x}(t), \bar{x}'(t)/\alpha(t))\big)\alpha(t)\,dt \geqslant J(\bar{x}).$$

We note that equality holds for $\bar{\alpha} \equiv 1$. We will now show that α solves a certain minimization problem.
We now introduce the function $\Phi(t, \alpha) = \Lambda\big(\bar{x}(t), \bar{x}'(t)/\alpha\big)\alpha$. The function $\Phi(t, \cdot)$ is then convex on the interval $(0, +\infty)$, for all $t \in [a, b]$. Consider the functional f given by

$$f(\alpha) = \int_a^b \Phi\big(t, \alpha(t)\big)\,dt,$$

which is well defined when α is measurable and has values in the interval $[1 - \delta, 1 + \delta]$. It also follows that f is a convex function.

By continuity we have that, for almost every t there exists a $\delta(t) \in (0, \delta]$ such that, for all $\alpha \in [1 - \delta(t), 1 + \delta(t)]$,

$$\Phi(t, 1) - 1 \leqslant \Phi(t, \alpha) \leqslant \Phi(t, 1) + 1.$$

Thanks to arguments of measurable selection theory, we can take $\delta(\cdot)$ to be measurable.

We define S to be the convex subset of $X := L^\infty\big([a, b]\big)$ whose elements α satisfy the condition $\alpha(t) \in [1 - \delta(t), 1 + \delta(t)]$ a.e..

Consider now the optimization problem on the vector space X which consists of minimizing f over S, subject to the equality constraint

$$h(\alpha) = \int_a^b \alpha(t)\,dt.(b - a) = 0.$$

Let's call (Q) this problem. This means that $\bar{\alpha} \equiv 1$ solves (Q).

Step 2. We now apply Theorem 2.13, obtaining a nonzero vector $\zeta = (\eta, \lambda) \in \mathbb{R}^2$ (whith $\eta = 0$ or 1) such that

$$\eta f(\alpha) + \lambda h(\alpha) \geqslant \eta f(\bar{\alpha}) \ \forall\, \alpha \in S.$$

We now want to show that $\eta = 1$.
Suppose $\eta = 0$. We then have

$$\eta f(\alpha) + \lambda h(\alpha) \geqslant \eta f(\bar{\alpha}) \ \forall\, \alpha \in S \iff \lambda h(\alpha) \geqslant 0 \ \forall\, \alpha \in S$$

But $\lambda \neq 0$, so we have

$$h(\alpha) \geqslant 0 \;\forall\, \alpha \in S$$
or
$$h(\alpha) \leqslant 0 \;\forall\, \alpha \in S$$

The condition $h(\alpha) \geqslant 0 \;\forall\, \alpha \in S$ is equivalent to

$$\int_a^b \alpha(t) - 1 \, dt \geqslant 0 \, \forall \alpha \in S,$$

from which we have a contradiction since, taken $\alpha(t) := 1 - \delta(t)$, we have $\alpha \in S$ and

$$\int_a^b \big(\alpha(t) - 1\big) dt = -\delta(t)(b-a) < 0.$$

In the same way we have that also the condition $h(\alpha) \leqslant 0 \;\forall\, \alpha \in S$ leads to a contradiction.

We can then suppose $\eta = 1$. We then have, for any $\alpha \in S$, the inequality

$$\int_a^b \Big\{\Lambda\big(\bar{x}(t), \bar{x}'(t)/\alpha(t)\big)\alpha(t) + \lambda\alpha(t)\Big\} dt \geqslant \int_a^b \Big\{\Lambda\big(\bar{x}(t), \bar{x}'(t)\big) + \lambda\Big\} dt.$$

Invoking a result on multifunction (Theorem 6.31 [2]), we deduce that, for almost every t, the function

$$\alpha \mapsto \theta_t(\alpha) := \Lambda\big(\bar{x}(t), \bar{x}'(t)/\alpha\big)\alpha + \lambda\alpha$$

attains a minimum over the interval $[1-\delta(t), 1+\delta(t)]$ at the interior point $\alpha = 1$. Let us fix such a value of t. Then the generalized gradient of θ_t at 1 must contain 0. From nonsmooth calculus we have

$$\Lambda\big(\bar{x}(t), \bar{x}'(t)\big) - \big\langle \bar{x}'(t), \zeta(t)\big\rangle = -\lambda \text{ a.e.}, \tag{1}$$

where $\zeta(t)$ lies in the subdifferential at $\left|\bar{x}'(t)\right|$ of the convex function $v \mapsto \Lambda\big(\bar{x}(t), v\big)$.

Step 3. What is left is only to prove that $\left|\bar{x}'(t)\right|$ is bounded for a.e. $t \in [a, b]$. Suppose we are in the conditions for (1) to hold. Then we have, using the subgradient inequality,

$$\begin{aligned} &\Lambda\Big(\bar{x}(t), \bar{x}'(t)\big\{1 + |\bar{x}'(t)|\big\}^{-1}\Big) - \Lambda\big(\bar{x}(t), \bar{x}'(t)\big) \geqslant \\ &\geqslant \Big[\big\{1 + |\bar{x}'(t)|\big\}^{-1} - 1\Big]\big\langle \bar{x}'(t), \zeta(t)\big\rangle = \\ &= \Big[\big\{1 + |\bar{x}'(t)|\big\}^{-1} - 1\Big]\Big\{\Lambda\big(\bar{x}(t), \bar{x}'(t)\big) + \lambda\Big\}. \end{aligned} \tag{2}$$

We now choose M to be a bound for all values of Λ at points of the form $\big(\bar{x}(t), w\big)$, with $t \in [a, b]$ and $w \in B(0, 1)$, using the hypothesis of Nagumo growth and inequality (2), we have:

$$\theta\big(|\bar{x}'(t)|\big) \leqslant \Lambda\big(\bar{x}(t), \bar{x}'(t)\big) \leqslant M + \big(M + |\lambda|\big)|\bar{x}'(t)|.$$

From this we can conclude that $\left|\bar{x}'(t)\right|$ is bounded for a.e. $t \in [a, b]$. □

2.2 Controllability

The question we have to ask ourselves before going any further is this one: given the initial point x_0 and a *target set* $S \subseteq \mathbb{R}^n$, does there exist a control u that can steer the system to S in a finite time?
To simplify the discussion, we will consider the problem of driving the system to the origin, that is we will suppose $S = \{0\}$. In other words, we will ignore any pay-off criterion, and instead focus on the pure existence of a control that can steer the system to the origin.

Definition 2.14. We define the *controllable set for time t* to be

$$\mathcal{C}(t) = \text{set of initial points } x_0 \text{ for which there exists a control such that } x(t) = 0,$$

and the overall *controllable set* to be

$$\mathcal{C} = \text{set of initial point s } x_0 \text{ for which there exists a control such that } x(t) = 0 \text{ for some finite time } t.$$

Clearly, we have

$$\mathcal{C} = \bigcup_{t \geqslant 0} \mathcal{C}(t).$$

To further simplify the matter, we will assume for the rest of this section that the ODEs we are studying are linear in both the state $x(\cdot)$ and the control $u(\cdot)$:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ x(0) = x_0, \end{cases}$$

where $A \in \mathbb{M}^{n \times n}$ and $B \in \mathbb{M}^{n \times m}$. We also assume that the set U of control parameters is a cube in $\mathbb{R}^m$: $U = [-1, 1]^m$.

Recalling the *variation of parameters formula* for the solution of ODEs, we have (for a given control $u(\cdot)$):

$$x(t) = e^{tA}x_0 + e^{tA} \int_0^t e^{-sA}(s)Bu(s)ds.$$

If $\mathcal{U} = \{u : \mathbb{R} \to U, \text{ measurable}\}$, then we have that

$$x_0 \in \mathcal{C}(t) \Leftrightarrow \exists\, u(\cdot) \in \mathcal{U} : \ x(t) = 0 \Leftrightarrow x_0 = -\int_0^t e^{-sA}Bu(s)ds \text{ for some } u(\cdot) \in \mathcal{U}.$$

Theorem 2.15 (structure of the controllable set)**.** With the above notations and assumptions we have that:

a) the controllable set $\mathcal{C}$ is symmetric;

b) the controllable set $\mathcal{C}$ is convex;

c) if $x_0 \in \mathcal{C}(\bar{t})$, then $x_0 \in \mathcal{C}(t)\ \forall t \geqslant \bar{t}$.

Proof. *a*) Let $t \geqslant 0$ and $x_0 \in \mathcal{C}(t)$. Then $x_0 = -\int_0^t e^{-sA}Bu(s)ds$ for some admissible control $u(\cdot) \in \mathcal{U}$.

Therefore $-x_0 = -\int_0^t e^{-sA}B(-u(s))ds$, with $-u \in \mathcal{U}$, since U is a symmetric set. This means that $-x_0 \in \mathcal{C}(t)$, and so also the set $\mathcal{C}(t)$ is symmetric. It follows that $\mathcal{C}$ is symmetric.

b) Take $x_0, \hat{x}_0 \in \mathcal{C}$; so that $x_0 \in \mathcal{C}(t), \hat{x}_0 \in \mathcal{C}(t)$ for appropriate times $t, \hat{t} \geqslant 0$.

Assume $t \leqslant \hat{t}$. Then:

$$x_0 = -\int_0^t e^{-sA}Bu(s)ds \qquad \text{for some control } u(\cdot) \in \mathcal{U},$$

$$\hat{x}_0 = -\int_0^{\hat{t}} e^{-sA}B\hat{u}(s)ds \qquad \text{for some control } \hat{u}(\cdot) \in \mathcal{U}.$$

If we now define a new control

$$\tilde{u}(s) := \begin{cases} u(s) \text{ if } 0 \leqslant s \leqslant t, \\ 0 \text{ if } s > t, \end{cases}$$

then we have

$$x_0 = -\int_0^{\hat{t}} e^{-sA}B\tilde{u}(s)ds,$$

hence $x_0 \in \mathcal{C}(\hat{t})$. Now let $0 \leqslant \lambda \leqslant 1$, and observe

$$\lambda x_0 + (1-\lambda)\hat{x}_0 = -\int_0^{\hat{t}} e^{-sA}B(\lambda\tilde{u}(s) + (1-\lambda)\hat{u}(s))ds.$$

Therefore $\lambda x_0 + (1-\lambda)\hat{x}_0 \in \mathcal{C}(\hat{t}) \subseteq \mathcal{C}$, which proves convexity.

c) Follows from b), if we take $\bar{t} = \hat{t}$.

□

Definition 2.16. We say that a linear system is **controllable** if $\mathcal{C} = \mathbb{R}^n$.

Definition 2.17. The **controllability matrix** of a linear system of the form $\dot{x} = Ax + Bu$ is the $n \times (mn)$ matrix $G = G(A, B) := [B, AB, A^2B, \dots, A^{n-1}B]$.

Theorem 2.18. Let $\mathcal{C}^\circ$ for the interior of the set $\mathcal{C}$. Then we have

$$\text{rank } G = n \iff 0 \in \mathcal{C}^\circ. \tag{2.8}$$

For a proof of this Theorem see [4] page 19.

Theorem 2.19. Let U be the cube $[-1, 1]^n \subseteq \mathbb{R}^n$. Suppose that rank $G = n$ and Re $\lambda < 0$ for each eigenvalue λ of A. Then the system $\dot{x}(t) = Ax(t) + Bu(t)$ is controllable.

Proof. [4] Thanks to Theorem (2.18), since rang $G = n$, we know that $\mathcal{C}$ contrains some ball $\mathcal{B}$ centred at 0. Now take any $x_0 \in \mathbb{R}^n$ and consider the evolution

$$\begin{cases} \dot{x}(t) = Ax(t). \\ x(0) = x_0. \end{cases}$$

Since Re $\lambda < 0$ for each eigenvalue λ of A, then the origin is asymptotically stable (Lyapunov).

So there exists a time T such that $x(T) \in \mathcal{B} \subseteq \mathcal{C}$; and hence there exists a control $u(\cdot) \in \mathcal{U}$ steering $x(T)$ into 0 in a finite time. □

2.3 Pontryagin Maximum Principle

We are now ready to write our first *maximum principle*, whose first version dates back to the 1960s.

Let us first recall our working hypotheses, the *classical regularity hypotheses*:

- the function l is continuously differentiable;
- the functions f and Λ are continuous, and admit derivatives relative to x which are continuous in all variables (t, x, u).

We also recall the problem we are trying to solve, the standard *Optimal Control Problem*:

$$\begin{cases} \text{minimize} & J(x,u) = l(x(b)) + \int_a^b \Lambda(t, x(t), u(t))dt \\ \text{subject to} & \dot{x}(t) = f(t, x(t), u(t)),\ t \in [a,b] \text{ a.e.} \\ & u : [a,b] \to U, \text{ measurable} \\ & x(a) = x_0,\ x(b) \in E \end{cases} \tag{OC}$$

The following is Clarke's version of the Pontryagin Maximum Principle. The original version requires the target set E to be a smooth manifold. After Pontryagin, there have been many versions that required always less regularity hypotheses (namely by Clarke, Sussmann, etc.). In this version E is any closed subset of $\mathbb{R}^n$.
Clarke went on to further generalize this result, by not requiring that f be regular.

Theorem 2.20 (Clarke's nonsmooth version of Pontryagin Maximum Principle)**.** Let the process $(\bar{x}, \bar{u})$ be a minimizer for the problem (OC), under the classical regularity hypotheses, and where U is a bounded set.
Then there exist an *absolutely continuous* arc $p : [a,b] \to \mathbb{R}^n$ and a scalar η equal either to 0 or 1 such that the following conditions are satisfied:

the **transversality condition**

$$-p(b) \in \eta \nabla l\big(\bar{x}(b)\big) + N_E^L\big(\bar{x}(b)\big) \tag{2.9}$$

the **adjoint equation**

$$-\dot{p}(t) = p(t) \cdot f_x\big(t, \bar{x}(t), p(t), \bar{u}(t)\big) - \eta \Lambda_x\big(t, \bar{x}(t), \bar{u}(t)\big) \text{ a.e.} \tag{2.10}$$

the **maximum condition**

$$p(t) \cdot f\big(t, \bar{x}(t), \bar{u}(t)\big) = \max_{u \in U} p(t) \cdot f(t, \bar{x}, u) \text{ a.e.} \tag{2.11}$$

The adjoint equation and the Maximum condition can be rewritten in terms of the *Unmaximized Hamiltonian* associated to the problem OC:

$$\mathcal{H}^{\eta}(t, x, p, u) := p(t) \cdot f(t, x, u) - \eta\Lambda(t, x, u)$$

so as to obtain the following:

$$-\dot{p}(t) = D_x \mathcal{H}^{\eta}\big(t, \bar{x}(t), p(t), \bar{u}(t)\big) \text{ a.e.} \tag{2.13*}$$

$$\mathcal{H}^{\eta}\big(t, \bar{x}(t), p(t), \bar{u}(t)\big) = \max_{u \in U} \mathcal{H}^{\eta}\big(t, \bar{x}(t), p(t), u\big) \text{ a.e..} \tag{2.14*}$$

and the *state equation*

$$\dot{\bar{x}} = f\big(t, \bar{x}(t), \bar{u}(t)\big) = D_p \mathcal{H}^{\eta}\big(t, \bar{x}(t), p(t), \bar{u}(t)\big) \tag{2.12}$$

Remark. Writing the adjoint and state equations in this form

$$\dot{x} = D_p \mathcal{H}^{\eta}(t, x, p, u), \ -\dot{p} = D_x \mathcal{H}^{\eta}(t, x, p, u)$$

emphasizes their affinity with a classical Hamiltonian system of differential equations, with an extra *control term* present.

In the abnormal case $\eta = 0$, we notice that the two components, l and Λ, of the cost, do not explicitly appear in the conclusions of the maximum principle. We want to show that this case is indeed a pathology, and does not happen when the final state value $x(b)$ is (locally) unconstrained.

Corollary 2.21. Under the hypotheses of Theorem 2.20, suppose that $E = \mathbb{R}^n$ or, more generally, that $\bar{x}(b) \in \text{int}(E)$. Then the maximum principle holds with $\eta = 1$.

Proof. Let us suppose that the maximum principle holds with $\eta = 0$, and obtain from this a contradiction.
If $\bar{x}(b) \in \text{int}(E)$, we have $N_E^L(\bar{x}(b)) = 0$. The transversality condition implies that $p(b) = 0$. When $\eta = 0$, the adjoint equation reduces to the following linear differential equation for p:

$$-\dot{p}(t) = p(t) \cdot f_x\big(t, \bar{x}(t), p(t), \bar{u}(t)\big).$$

But any solution p of such linear differential equation that vanishes at one point ($p(b) = 0$) necessarily vanishes everywhere, violating the nontriviality condition of Theorem 3.1: a contradiction. □

The Pontryagin Maximum Principle provides a practical method for finding solutions to the problem OC. We first have to define the function $\bar{u}$ in terms of the maximum condition

$$p(t) \cdot f(t, \bar{x}(t), \bar{u}(t) = \max_{u \in U} p(t) \cdot f(t, \bar{x}, u).$$

Thanks to Theorem 2.20 we know that, if an optimal control u does exist, it must be found among the solutions of the $2n$ differential equation system in the variables x and p:

$$\begin{cases} \dot{x} = f(t, x, u(t)) \\ \dot{p} = -p(t) \cdot f_x(t, \bar{x}(t), p(t), \bar{u}(t)) + \eta \Lambda_x(t, \bar{x}(t), \bar{u}(t)) \end{cases} \tag{2.13}$$

The equations 2.13 do not constitute a Cauchy problem on $\mathbb{R}^n$, so a solution is not easily found, unless the equations for p and x can be uncoupled. This is what happens in some particular cases.

In all of the examples that we provide, we are going to assume the existence of an optimal solution.

Consider the following:

Example 2.22 (Linear pendulum with external force)**.** [1] Let q be the position of a linearized pendulum, controlled by an external force u, with magnitude constraint $|u(t)| \leqslant 1$, $\forall t$.

For simplicity's sake, let us assume that the initial position and velocity are both zero, and that the motion is determined by the equations

$$\ddot{q}(t) + q(t) = u(t), \; q(0) = \dot{q}(0) = 0.$$

We wish to maximize the displacement $q(b)$ at a fixed terminal time b.

Introducing the variables $x_1 = q$, $x_2 = \dot{q}$, the optimization problem can be formulated as

$$\max_{u \in U} x_1(b, u),$$

where the dynamics is described by

$$\begin{cases} \dot{x_1} = x_2 & x_1(0) = 0 \\ \dot{x_2} = -x_1 + u & x_2(0) = 0, \end{cases}$$

and the set of admissible control is

$$\mathcal{U} = \{u : [0, b] \to [-1, 1], u \text{ measurable}\}.$$

We then have:

$$f(t, x, u) = \begin{pmatrix} x_2 \\ -x_1 + u \end{pmatrix}, \qquad D_x f = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

In this case the adjoint equations take the form

$$\begin{cases} \dot{p_1} = p_2 & p_1(b) = 1, \\ \dot{p_2} = -p_1 & p_2(b) = 0. \end{cases}$$

These equations can be solved for p independently of x, yielding

$$p_1(t) = \cos(b - t), \quad p_2(t) = \sin(b - t).$$

By the maximum condition we have that the optimal control $\bar{u}$ satisfies

$$p_1x_2 + p_2(-x_1 + \bar{u}) = \max_{|u|\leqslant 1} \{p_1x_2 + p_2(-x_1 + u)\}.$$

Therefore, the optimal control is

$$\bar{u} = \operatorname{sgn}(p_2(t)) = \operatorname{sgn}(\sin(b - t)).$$

Notice that the trajectories corresponding to the constant controls $u \equiv 1$ or $u \equiv -1$ are circles centred at $(1, 0)$ or $(-1, 0)$, respectively.

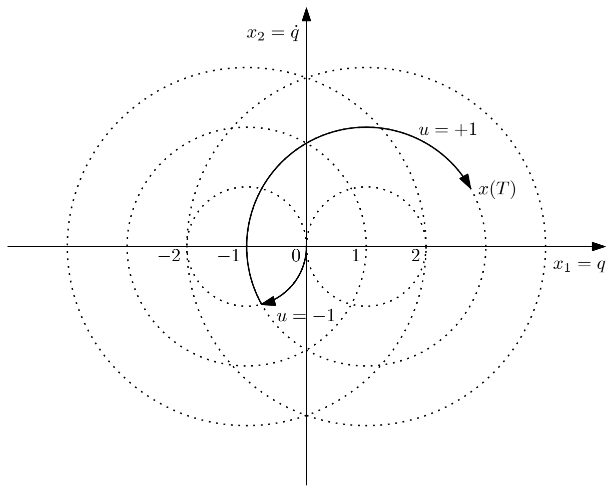

Figure 2.4: Trajectories in the phase plane for the linear pendulum

The optimal control for when $b = \frac{3\pi}{2}$ is then:

$$\bar{u}(t) = \begin{cases} -1 & \text{if } 0 < t < \pi/2, \\ 1 & \text{if } \pi/2 < t < 3\pi/2. \end{cases}$$

as highlighted in Figure 2.4.

2.4 Problems with variable time

What we have seen thus far are Optimal Control Problems in which the underlying interval is fixed. An important feature of certain problems is that the interval is itself a choice variable. Such a problem, in which the interval $[a, b]$ can vary, is referred to as a *variable-time problem*.

In this section we will consider the following form of the *optimal control problem* (note that f and Λ *do not* depend on the time variable t. This means that we can choose the initial point to be 0, without loss of generality):

$$\begin{cases} \text{minimize} & J(\tau, x, u) = l(\tau, x(\tau)) + \int_0^\tau \Lambda(x(t), u(t))dt \\ \text{subject to} & \tau \geqslant 0 \\ & \dot{x}(t) = f(x(t), u(t)),\ t \in [0, \tau]\ a.e. \\ & u : [0, \tau] \to U, \text{measurable} \\ & x(0) = x_0,\ (\tau, x(\tau)) \in S \end{cases} \tag{VT}$$

Remark. If the set S is of the form $\{T\} \times E$ then the problem (VT) reduces to a fixed-type problem like those we have already studied, since the terminal point τ is uniquely determined.

Example 2.23. The *minimal-time problem* consists of finding the process (x, u) on an interval $[0, \tau]$, such that $x(\tau) = 0$ and τ is the least time for which this is possible.

Of course, this corresponds to finding the quickest trajectory to the origin. In the notations of (VT) we have $S = \mathbb{R}_+ \times \{0\}$, $l \equiv 0$, $\Lambda \equiv 1$ (or $l \equiv \tau$ and $\Lambda \equiv 0$).

Since we now have a varying terminal point, we need to extend the concept of *local* minimum to variable-time problems.

If two arcs, x_1 and x_2 are defined on two different intervals, $[0, \tau_1]$ and $[0, \tau_2]$, we need to be able to measure their *distance*. We define the expression

$$\|x_1 - x_2\| := \max_{t \geqslant 0} |x_1(t) - x_2(t)|,$$

where we have extended the arcs by setting $x_1(t) = x_1(\tau_1)\ \forall\ t \geqslant \tau_1$, and $x_2(t) = x_2(\tau_2)\ \forall\ t \geqslant \tau_2$.

Definition 2.24. A process $(\bar{x}, \bar{u})$, defined on the interval $[0, \bar{\tau}]$, and satisfying the constraints of (VT), is said to be a **local minimizer** if there exist an $\varepsilon > 0$ such that, for all processes (x, u) on an interval $[0, \tau]$ which satisfy the constraints of (VT) and such that $|\tau - \bar{\tau}| \leqslant \varepsilon$ and $\|x - \bar{x}\| \leqslant \varepsilon$, we have $J(\bar{\tau}, \bar{x}, \bar{u}) \leqslant J(\tau, x, u)$.

Recalling the definitions and hypotheses of Theorem 2.20 we have the following:

Theorem 2.25 (Variable-Time Maximum Principle)**.** Let the process $(\bar{x}, \bar{u})$, defined on the interval $[0, \bar{\tau}]$, $\bar{\tau} > 0$, be a local minimizer for the problem (VT), under the classical regularity hypotheses, and where U is a bounded set.
Then there exist an arc $p : [0, \bar{\tau}] \to \mathbb{R}^n$ and a scalar η equal either to 0 or 1 such that the following conditions are satisfied:

the **non-triviality condition**

$$\big(\eta, p(t)\big) \neq 0\ \forall t \in [0, \bar{\tau}] \tag{2.14}$$

the **adjoint equation**

$$-\dot{p}(t) = D_x \mathcal{H}^\eta\big(t, \bar{x}(t), p(t), \bar{u}(t)\big)\ a.e. \tag{2.15}$$

the **maximum condition**

$$\mathcal{H}^\eta\big(t,\bar{x}(t),p(t),\bar{u}(t)\big) = \max_{u\in U}\mathcal{H}^\eta\big(t,\bar{x}(t),p(t),u\big) \text{ a.e..} \tag{2.16}$$

and such that, for some constant h, we have

the **constancy of the Hamiltonian**

$$\mathcal{H}^\eta\big(t,\bar{x}(t),p(t),\bar{u}(t)\big) = \max_{u\in U}\mathcal{H}^\eta\big(t,\bar{x}(t),p(t),u\big) = h \text{ a.e..} \tag{2.17}$$

as well as

the **transversality condition**

$$\big(h,-p(b)\big) \in \eta\nabla l\big(\bar{\tau},\bar{x}(\bar{\tau})\big) + N_S^L\big(\bar{\tau},\bar{x}(\bar{\tau})\big). \tag{2.18}$$

Remark. When S is of the form $\{T\}\times E$, the transversality condition reduces to the transversality condition of Theorem 2.20.

Example 2.26 (soft landing)**.** [2] This simple model is an interesting case of the minimal-time problem. The goal is to bring a spacecraft to a soft landing on the lunar surface in the least time.

We consider the dynamics

$$\ddot{x}(t) = u(t) \in [-1,1].$$

We want to find the control u that steers the initial state/velocity pair (x_0,v_0) to rest at the origin ($x=\dot{x}=0$) in the least time.

As usual, we introduce a second state variable, so that the second-order equation above takes the form of a first-order system:

$$\begin{pmatrix}\dot{x}(t)\\ \dot{y}(t)\end{pmatrix} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}x(t)\\ y(t)\end{pmatrix} + \begin{pmatrix}0\\ 1\end{pmatrix}u(t),\ u(t)\in[-1,1].$$

This is a linear system, with $n=2$ and $m=1$. We take $l(\tau,x)=\tau,\ \Lambda=0,\ S=\mathbb{R}_+\times\{(0,0)\}$. We now apply Theorem (2.25), admitting the existence of a solution.

The Hamiltonian is $\mathcal{H}^\eta(x,y,p,q,u) = py+qu$, so the *adjoint system* is given by

$$\dot{p}(t)=0,\ -q'(t)=p(t).$$

This means that $p(t)=p_0$ for some constant p_0, and that q is an affine function.
The *transversality condition* implies $h=\eta$.
If $q\equiv 0$, then also p is identically 0; but then we would have $h=\eta=0$, which violates the *non-triviality condition*. We conclude that q is not identically zero and thus, being affine, changes sign at most once in $[0,\tau]$.

The *constancy of the Hamiltonian* yields $p_o y(t) + q(t)u(t) = h$ for some constant h.

The *maximum condition* then implies that the optimal control is $u = \pm 1$, known as a *bang-bang* solution. The plus or minus sign depends on the sign of q, that is the optimal control is equal to 1 almost everywhere up to a certain point, then -1 thereafter, or else the reverse. In other words u is piecewise constant, with values in $\{-1, 1\}$, and exhibits at most one change in sign.

The trajectories (x, y) for the constant control value $u = 1$ lie on parabolas of the form $2x = y^2 + c$, since we have $2\dot{x} - 2y\dot{y} = 0$; the movement is upward since $\dot{y} = 1$. Similarly, the trajectories (x, y) for $u = -1$ correspond to parabolas of the form $2x = -y^2 + c$, with a downward motion. If a time-optimal strategy does exist (we have seen in Section 2.2 that the system is completely controllable) then it is described as follows:

$$\bar{u} = \begin{cases} +1 & \text{if } (x, y) \text{ lies to the left of } \Sigma, \\ -1 & \text{if } (x, y) \text{ lies to the right of } \Sigma, \end{cases}$$

where Σ is the *switching curve*, defined as $\Sigma = \{(-y^2/2, y) : y \geqslant 0\} \cup \{(y^2/2, y) : y \leqslant 0\}$, is obtained, for example, by beginning at a point on the positive y-axis, and following a downward parabola until its intersection with the unique upward parabola passing through $(0, 0)$; finally following that one to the origin.

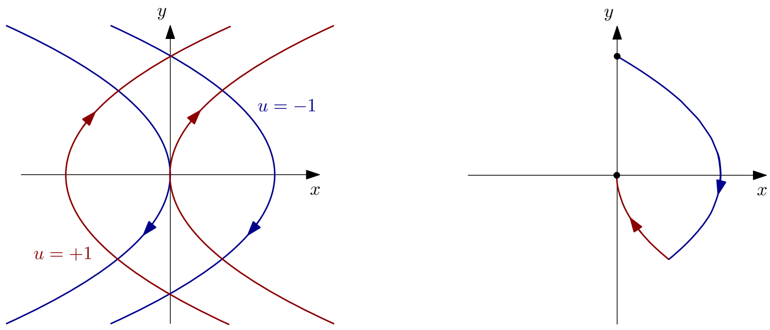

Figure 2.5: Sample trajectories for $u = +1$ and for $u = -1$, and an optimal trajectory starting from the positive y-axis

We can now compute the corresponding optimal time

$$\tau(x, y) = \begin{cases} -y + \sqrt{2y^2 - 4x} & \text{if } (x, y) \text{ lies to the left of } \Sigma, \\ +y + \sqrt{2y^2 + 4x} & \text{if } (x, y) \text{ lies to the right of } \Sigma. \end{cases}$$

We have said, in the Preface of this dissertation, that we would study the general case in which the target set E is a closed subset of $\mathbb{R}^n$. The following example shows that this generalization serves an actual purpose: here the target set will not be a manifold.

Example 2.27. Consider the system

$$\begin{cases} (\dot{x}_1, \dot{x}_2) = (2u_1, u_2), \quad (u_1, u_2) \in \mathbb{R}^2, \quad |(u_1, u_2)| \leqslant 1, \\ (x_1(0), x_2(0)) = (\underline{x}_1, \underline{x}_2) \in \mathbb{R}^2 \backslash \mathcal{S}, \\ (x_1(\tau), x_2(\tau)) \in \mathcal{S} \end{cases} \tag{2.19}$$

where $\mathcal{S} = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \leqslant |x_2|\}$, and τ is the first instant for which the trajectory is in $\mathcal{S}$.

The problem is that of minimizing τ. As usual, we admit the existence of a solution. Under the notations of the Maximum Principle we have the following

$$\begin{cases} \text{minimize} & \int_0^\tau 1 dt, \\ \text{subject to} & \tau \geqslant 0 \\ & (\dot{x}_1, \dot{x}_2) = (2u_1, u_2) \\ & (u_1, u_2) \in \mathbb{R}^2 \\ & |(u_1, u_2)| \leqslant 1, \\ & (x_1(0), x_2(0)) = (\underline{x}_1, \underline{x}_2) \in \mathbb{R}^2 \backslash \mathcal{S}, \\ & (x_1(\tau), x_2(\tau)) \in \mathcal{S} \end{cases} \tag{2.20}$$

This is a minimum-time problem, and so we take $l = 0$ and $\Lambda = 1$. Let $x = (x_1, x_2)$, $u = (u_1, u_2)$ and $p = (p_1, p_2)$.

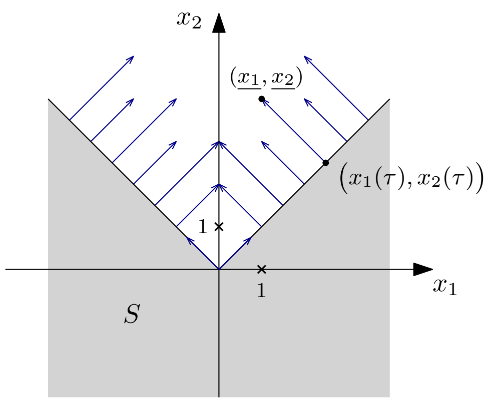

Figure 2.6: Some proximal normal directions to S

We now consider the Unmaximized Hamiltonian

$$H^\eta(x, p, u) = p(t) \cdot f(t, x, u) - \eta\Lambda(t, x, u) = 2p_1u_1 + p_2u_2 - \eta.$$

From the maximum condition we obtain:

$$2p_1\bar{u}_1 + p_2\bar{u}_2 - \eta = \max_{u_1^2 + u_2^2 \leqslant 1} 2p_1u_1 + p_2u_2 - \eta = h,$$

that is

$$2p_1\bar{u}_1 + p_2\bar{u}_2 = \max_{u_1^2 + u_2^2 \leqslant 1} 2p_1u_1 + p_2u_2.$$

Since the limiting normal cone is as portrayed in Figure 2.6, the transversality condition yields:

$$-p(\tau) \in N_S^L(\bar{x}(\tau)) \in \left\{ \lambda \begin{pmatrix} +1 \\ +1 \end{pmatrix}, \lambda \begin{pmatrix} -1 \\ +1 \end{pmatrix} \right\}, \ \lambda \geqslant 0.$$

From the adjoint equation we obtain that $-\dot{p} = 0$, which means that p_1 and p_2 are constant.

We then have two situations:

First Case. $-p_1 = \lambda \geqslant 0 \implies p_1 = -\lambda, \ p_2 = \lambda.$

$$\max_{|u| \leqslant 1} H^\eta(x, u) \longleftrightarrow \min_{u_1^2 + u_2^2 \leqslant 1} 2u_1 + u_2.$$

We know that the minimum is taken for points in the border of the disc $\{u_1^2 + u_2^2\} \leqslant 1$, and so we have that the values of u_1 and u_2 that obtain the minimum are such that:

$$\begin{cases} \begin{vmatrix} 2 & u_1 \\ 1 & u_2 \end{vmatrix} = 0 \\ u_1^2 + u_2^2 = 1 \end{cases} \iff \begin{cases} 2u_2 = u_1 \\ u_1^2 + u_2^2 = 1 \end{cases} \iff \begin{cases} u_1 = 2u_2 \\ u_2 = \pm 1/\sqrt{5} \end{cases}$$

Since we want to minimize the function $2u_1 + u_2 = 5u_2$, we have $u_2 = -1/\sqrt{5}$ and $u_2 = -2/\sqrt{5}$.

We can then write the equations of the motion

$$\begin{cases} \dot{x}_1 = 2u_1 = -4/\sqrt{5} \\ \dot{x}_2 = u_2 = -1/\sqrt{5} \end{cases} \iff \begin{cases} x_1(t) = -4/\sqrt{5}t + \underline{x}_1 \\ x_2(t) = -1/\sqrt{5}t + \underline{x}_2 \end{cases}$$

We now wish to compute the minimal time τ. Since $x_2(\tau) = |x_1(\tau)|$, we have the following situation:

$$\begin{cases} x_2(\tau) = x_1(\tau) \\ \quad \text{or} \\ x_2(\tau) = -x_1(\tau) \end{cases} \iff \begin{cases} -4/\sqrt{5}\tau + \underline{x}_1 = -1/\sqrt{5}\tau + \underline{x}_2 \\ \quad \text{or} \\ -4/\sqrt{5}\tau + \underline{x}_1 = 1/\sqrt{5}\tau - \underline{x}_2 \end{cases} \iff \begin{cases} \underline{x}_1 - \underline{x}_2 = 3/\sqrt{5}\tau \\ \quad \text{or} \\ \underline{x}_1 + \underline{x}_2 = \sqrt{5}\tau \end{cases}$$

But $\underline{x}_1 - \underline{x}_2 < 0$, so the only admissible solution is

$$\tau = \frac{\underline{x}_1 + \underline{x}_2}{\sqrt{5}}.$$

Second Case. $-p_1 = -\lambda \leqslant 0 \implies p_1 = \lambda, \ p_2 = -\lambda.$
In much the same way as previously, we have that the values of u_1 and u_2 that obtain minimum are such that:

$$\begin{cases} \begin{vmatrix} 2 & u_1 \\ -1 & u_2 \end{vmatrix} = 0 \\ u_1^2 + u_2^2 = 1 \end{cases} \iff \begin{cases} -2u_2 = u_1 \\ u_1^2 + u_2^2 = 1 \end{cases} \iff \begin{cases} u_2 = \pm 1/\sqrt{5} \\ u_1 = \mp 2/\sqrt{5} \end{cases}$$

Since we want to minimize the function $2u_1 - u_2 = -5u_2$, we have $u_2 = -1/\sqrt{5}$ and $u_2 = 2/\sqrt{5}$.

We can then write the equations of the motion

$$\begin{cases} \dot{x}_1 = 2u_1 = 4/\sqrt{5} \\ \dot{x}_2 = u_2 = -1/\sqrt{5} \end{cases} \iff \begin{cases} x_1(t) = 4/\sqrt{5}t + \underline{x}_1 \\ x_2(t) = -1/\sqrt{5}t + \underline{x}_2 \end{cases}$$

We now wish to compute the minimal time τ. We have the following:

$$\begin{cases} x_2(\tau) = x_1(\tau) \\ \quad\text{or} \\ x_2(\tau) = -x_1(\tau) \end{cases} \iff \begin{cases} 4/\sqrt{5}\tau + \underline{x}_1 = 1/\sqrt{5}\tau - \underline{x}_2 \\ \quad\text{or} \\ 4/\sqrt{5}\tau + \underline{x}_1 = -1/\sqrt{5}\tau + \underline{x}_2 \end{cases} \iff \begin{cases} -\underline{x}_1 - \underline{x}_2 = 3/\sqrt{5}\tau \\ \quad\text{or} \\ -\underline{x}_1 + \underline{x}_2 = \sqrt{5}\tau \end{cases}$$

But $-\underline{x}_1 - \underline{x}_2 < 0$, so the only admissible solution is

$$\tau = \frac{\underline{x}_2 - \underline{x}_1}{\sqrt{5}},$$

We then have two possible minimum times, and we wish to compare them:

$$\frac{\underline{x}_2 - \underline{x}_1}{\sqrt{5}} \leqslant \frac{\underline{x}_1 + \underline{x}_2}{\sqrt{5}} \iff \underline{x}_1 \geqslant 0.$$

In conclusion we have:

- If $\underline{x}_1 \geqslant 0$, then $\tau = \dfrac{\underline{x}_2 - \underline{x}_1}{\sqrt{5}}$, with optimal control $\bar{u} = (2/\sqrt{5}, -1/\sqrt{5})$
- If $\underline{x}_1 \leqslant 0$, then $\tau = \dfrac{\underline{x}_1 + \underline{x}_2}{\sqrt{5}}$, with optimal control $\bar{u} = (-2/\sqrt{5}, -1/\sqrt{5})$

We observe that this solution for the minimal-time problem is **not** differentiable for $\underline{x}_1 = 0$.

Bibliography

[1] Alberto Bressan and Benedetto Piccoli, *Introduction to the Mathematical Theory of Control*, American Institute of Mathematical Sciences (2007).

[2] Francis Clarke, *Functional Analysis, Calculus of Variations and Optimal Control*, Springer (2013).

[3] Richard Vinter, *Optimal Control*, Birkhäuser (2000).

[4] Lawrence C. Evans, *An Introduction to Mathematical Optimal Control Theory Version 0.2*, University of California, Berkeley.

[5] Franco Cardin, *Sistemi Dinamici Meccanici*, Università degli Studi di Padova (2014).

[6] Alessandro Ferriero, *The Lavrentiev phenomenon in the Calculus of Variations*, Università degli Studi di Milano-Bicocca (2004).

[7] Hector Sussmann, *Maximum Principle*, published in the Proceeding of the 46th IEEE CDC, New Orleans, LA (December 12-14 2007).