Universidade do Minho
Escola de Ciências

Sofia Oliveira Lopes

Nondegenerate forms of the Maximum Principle for Optimal Control Problems with State Constraints

Outubro, 2008

Universidade do Minho
Escola de Ciências

Sofia Oliveira Lopes

Nondegenerate forms of the Maximum Principle for Optimal Control Problems with State Constraints

Doutoramento em Ciências

Área de Conhecimento:
Matemática

Trabalho efectuado sob a orientação do
Professor Doutor Fernando A. C. C. Fontes
e da **Professora Doutora Maria do Rosário de Pinho**

Outubro, 2008

DECLARAÇÃO

Nome : Sofia Oliveira Lopes

Endereço electrónico: sofialopes@mct.uminho.pt Telefone: 253 510 429 / 91 4014961

Número do Bilhete de Identidade: 10715093

Título dissertação □/tese ☒: Nondegenerate forms of the Maximum Principle for Optimal Control Problems with State Constraints

Orientador(es): Fernando A. C. C. Fontes e Maria do Rosário de Pinho Ano de conclusão: 2008

Designação do Mestrado ou do Ramo de Conhecimento do Doutoramento: Doutoramento em Ciências

Universidade do Minho, 03/10/2008

Assinatura: Sofia Oliveira Lopes

Acknowledgements

I would like to express my deep and sincere gratitude to my supervisor Professor Fernando A. C. C. Fontes for his advice, his supervision and his crucial contributions which made him a backbone of this research and thus of this thesis.

Besides I am deeply grateful to Professor Maria do Rosário de Pinho, for her detailed and constructive comments, and for her important support throughout this work.

In addition to that I wish to thank Professor Helene Frankowska for giving me the opportunity to work with her and share all her knowledge.

My gratitude also goes to Professor Delfim F. M. Torres and Professor Emmanuel Trelat for their support.

During this work I have collaborated with many colleagues for whom I hold high regard, and I wish to extend my warmest thanks to all those who have helped me with my work in the Department of Mathematics for Science and Technology.

Words fail me to express my appreciation to my husband Emanuel whose love and persistent confidence in me have taken much off my shoulders. His company and support is my source of strength during this strenuous PhD journey. My special gratitude is due to my sister and parents, who have always supported me and have kept me focused.

And last, but definitely not least, I'd like to thank my son and nephew. It is to them that this thesis is dedicated.

The financial support from Project HPMT-CT-2001-00278 of CTS - Control Traning Site, from projecto "Optimização e Controlo" of FCT-Program and from Projecto FCT POSI/EEA-SRI/61831/2004 "Controlo Óptimo com Restrições e suas Aplicações" are gratefully acknowledged.

Abstract

Nondegenerate forms of the Maximum Principle for Optimal Control Problems with State Constraints

The Maximum Principle (MP) plays an important role in the characterization of solutions to optimal control problems. It typically identifies a small set of candidates where the minimizers belong.

However, for some optimal control problems with constraints, it may happen that the MP is unable to provide any useful information; for example, if the set of candidates to minimizers that satisfy a certain MP coincides with the set of all admissible solutions. When this happens, we say that *the degeneracy phenomenon* occurs.

One of ours main goals, is preventing the degeneracy phenomenon to occur by imposing additional terms to the MP. In this context, we developed new strengthened forms of the MP, for optimal control problems and in particular for optimal control problems with higher index state constrains.

Another case where the MP is unable to provide any useful information happens when the scalar multiplier associated with the objective function is equal to zero. So, the MP merely states a relation between the constraints and does not use the objective function to select candidates to minimizers. We have also developed strengthened forms of the MP such that the MP can be written with the multiplier associated with the objective function not zero, the so-called *normal forms* of the MP, for optimal control problems.

These two types of strengthened forms of the MP can be applied when the problem satisfies additional hypotheses, known as *constraint qualifications,* and therefore the constraint qualifications are also object of our study.

The nondegenerate forms of MP, that were developed in this thesis, are valid for new types of optimal control problems with state constraints both by addressing problems with less restrictions on its data, and also by developing new constraint qualifications that are verified for more problems or are easier to verify whether they are satisfied.

Resumo

Formas não degeneradas do Princípio do Máximo para Problems de Control Óptimo com Restrições de Estado

O Princípio do Máximo (PM) tem um papel fundamental na caracterização de soluções de problemas de controlo óptimo. O PM tipicamente identifica um pequeno conjunto de candidatos entre os quais se encontram o(s) óptimos.

Contudo, para alguns problemas de controlo óptimo com restrições, o PM poderá não fornecer qualquer informação útil; por exemplo, se o conjunto de candidatos a mínimos que satisfaz o PM coincide com o conjunto de todas as soluções admissíveis. Quando tal acontece, dizemos que o *fenómeno de degeneração* ocorre.

Um dos nossos principais objectivos, é garantir a não ocorrência do fenómeno de degeneração impondo condições adicionais ao PM. Neste contexto, desenvolvemos formas fortalecidas do PM para problems de controlo óptimo e em particular para problemas de controlo óptimo com restrições de estado de "elevado" índice.

Outro caso em que o PM não fornece informação útil, ocorre quando o multiplicador associado à função objectivo é igual a zero. Neste caso o PM é uma mera relação entre as restrições e portanto não usa a função objectivo para seleccionar um conjunto de candidatos a mínimos. Desenvolvemos, também, formas fortalecidas do PM de modo a que possam ser escritas com o multiplicador associado à função objectivo não nulo, denominadas por PM *normais*, para problemas de controlo óptimo.

Estes dois tipos de condições fortalecidas são aplicáveis apenas quando o problema satisfaz hipóteses adicionais, conhecidas como *qualificações de restrição*, e portanto as qualificações de restrição são também objecto do nosso estudo.

As formas não degeneradas do PM, desenvolvidas nesta tese, são válidas para novos tipos de problems de controlo óptimo com restrições de estado, simultaneamente por permitirem problemas com menos restrições nos dados, e também por desenvolverem qualificações de restrição que são verificadas para um maior número de problemas ou são mais fáceis de verificar.

Contents

Notation

NCO	Necessary Conditions of Optimality
CQ	Constraint Qualification
OCP	Optimal Control Problem
CVP	Calculus of Variation Problem
MPP	Mathematical Programming Problem
MP	Maximum Principle
$W^{1,1}(I : \mathbb{R})$	absolutely continuous functions from I to $\mathbb{R}$
$C([I] : \mathbb{R})$	continuous functions from I to $\mathbb{R}$
$C^*([I] : \mathbb{R})$	dual space of the space of continuous functions $C([I] : \mathbb{R})$
$C^{1,1}$	class of functions which are continuously differentiable with locally Lipschitz continuous derivatives
$\mathbb{B}$	closed unit ball in Euclidean space
$x + \delta\mathbb{B}$	ball of radius δ centred at x

$\|x\|$	Euclidean norm of x
$\|x\|_X$	norm of x in the space X
$\dot{x}(t)$	total derivative with respect to t
dom f	domain of f
int (C)	interior of C
bdy (C)	boundary of C
$\bar{C}$	closure of C
co C	convex hull of a set C
$d_C(x)$	Euclidean distance of x to the set C
epi f	epigraph of f
supp $\{\mu\}$	support of a measure μ
$a \cdot b$	inner product of a and b
$(\bar{x}, \bar{u})$	optimal minimizer of a optimal control problem
$f_x(x)$	derivative of f with respect to x
C^-	negative polar cone
$T_C(x)$	contingent cone to a set C at x

$C_C(x)$	Clarke's tangent cone to the set C at x
$N_C(x)$	Clarke's normal cone to the set C at x
$N_C^L(x)$	limiting normal cone to the set C at x
$\partial^L f(x)$	limiting subdifferential of f at x
$\tilde{\partial} f(x)$	Clarke's subdifferential of f at x
$\partial_x^> h(t,x)$	hybrid partial subdifferential
H	(pseudo-) Hamiltonian function

List of Constraint Qualifications

Optimal Control Problems

CQ to guarantee Nondegeneracy

CQ to guarantee Normality

Calculus of Variations Problems

List of Figures

Chapter 1

Introduction

1.1 Scope and Motivation

In this thesis we deal with both the “Calculus of Variations Theory” and the “Optimal Control Theory”. Our study focus on a set of conditions (necessary conditions of optimality — NCO) that allow identify a small set of candidates to minimizers among the overall set of admissible solutions.

In the 1950’s, these conditions were proved for problems with high regularity on the data, but the continuous developments in this area allowed establishing conditions for problems with: “nonsmooth” data (data that can be non differentiable), more general end-point constraints, state constraints, and other refinements.

Almost all optimization problems arising in practice really have constraints and these constraints are limitations on our decisions. For example, operations may be limited to so many hours in day, a plane to fly in security must have constraints on altitude or velocity, chemical reactors have to be limited by maximum temperature or pressure, a vehicle or robot has to avoid obstacles, amongst many others. However, for optimal control problems with state constraints the standard NCO could not, in same cases, provide useful information to select candidates to minimizers. This happens when the set of candidates to minimizers that satisfy certain NCO coincides with the set of all admissible solutions or when the scalar multiplier associated with objective function is equal to zero. It is possible, nevertheless, to avoid such phenomenon by strengthening the NCO.

As in [Fon99], we emphasise the importance attached to nondegenerate conditions by reference to their history in Mathematical Programming ([Aba67, Man69]). The Kuhn-Tucker conditions are best known optimality conditions for Mathematical Programming problems with inequality constraints. However, these conditions are just a strengthened version of previous Fritz John conditions, imposing the multiplier associated with the objective function to be positive, or simply equal to 1. Nowadays, the Kuhn-Tucker conditions are one of the most cited results in optimization. This illustrates the significance of nondegenerate versions of necessary conditions of optimality.

On other hand, the nondegenerate and normality results are important to establish the regularity of optimal trajectories and controls, and also in establishing links between NCO and Hamilton-Jacobi equations.

In this thesis we developed new strengthened forms (nondegenerate and normal forms) of necessary condition of optimality for optimal control problems with state constraints.

1.2 Overview

This thesis is organized as follows. In Chapter 2, we introduce the classical necessary conditions for calculus of variations and optimal control problems. We also introduce here some recent developments that will be of use later in this thesis and we finish this chapter with some concepts of regularity.

In Chapter 3, we review the main literature of strengthened necessary conditions for mathematical programming and optimal control problems.

Chapter 4 contains the normality result for calculus of variations problems that was developed in the author's master thesis and a discussion of the relative merits of necessary conditions of optimality that were developed for optimal control problems, in the particular case of calculus of variations problems.

The main contributions of this thesis are given from Chapter 5 to Chapter 9. Chapter 5 involves a new normality result for optimal control problems via a linearization of control systems, while the results introduced on Chapter 6 to Chapter 9 are nondegenerate results.

In Chapter 6, we propose a nondegenerate maximum principle (MP) valid under a constraint qualification of integral-type.

The nondegenerate MP, provided in the Chapter 7, is valid under constraints qualifications that are easier to verify than some appearing in previous literature.

In Chapter 8, the main result guarantees the nondegeneracy for problems that satisfies an easier verifiable integral-type constraint qualification.

In the Chapter 9, we developed a new constraint qualification for optimal control problems with state constraints that have higher index (i.e. their first derivative with respect to time does not depend on the control).

We conclude this thesis by providing a summary of contributions and posing some related open questions to motivate further research.

Finally, we offer in the Appendix a brief review of relevant background material in functional analysis and nonsmooth analysis.

Chapter 2

Background

Since necessary conditions of optimality (NCO) are the main tools of this thesis, we present in this chapter classical results on the subject for calculus of variations and optimal control problems in an informal setting. We also introduce some recent developments that will actually be of use in this thesis. We finish this chapter with some concepts of regularity.

2.1 NCO for Calculus of Variations Problems

The basic calculus of variations problem (CVP) is to find an absolutely continuous function $\bar{x}$ that solves the following problem:

$$(CVP_1) \quad \begin{cases} \text{Minimize} & J[x] = \displaystyle\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \\ \text{subject to} & x(t_0) = x_0 \\ & x(t_1) = x_1. \end{cases}$$

The interval $[t_0, t_1]$, the Lagrangian function $L : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, the initial state x_0 and the final state x_1 are given as part of the problem statement.

We say that x is an admissible trajectory if x is an absolutely continuous function on the interval $[t_0, t_1]$, satisfying the constraints of the problem, $x(t_0) = x_0$ and $x(t_1) = x_1$ and such that $L(t, x(t), \dot{x}(t))$ is a Lebesgue integrable function in this interval. The minimizer for the problem is an admissible trajectory $\bar{x}$ in the interval $[t_0, t_1]$, that satisfies

$$J[\bar{x}] \leq J[x],$$

for any admissible trajectory x in $[t_0, t_1]$.

The Calculus Variations theory is an important tool in laws of physics that identified states of nature with minimizing curves and surfaces lengthened, as Fermat's principle, Dirichlets's principle, principle of least actions, among others, (see for example [Vin00] and [Loe93]).

The best known NCO for CVP are the Euler-Lagrange and the Weierstrass Condition (see for example [Cla89], [Vin00]). They assert the existence a function $p \in W^{1,1}([t_0, t_1] : \mathbb{R}^n)$ such that

Euler-Lagrange Condition:

$$(\dot{p}(t), p(t)) = L_{x,u}(t, \bar{x}(t), \dot{\bar{x}}\ (t)),$$

Weierstrass Condition:

$$p(t) \cdot \dot{\bar{x}}\ (t) - L(t, \bar{x}(t), \dot{\bar{x}}\ (t)) = \max_{u \in \mathbb{R}^n} [p(t) \cdot u - L(t, \bar{x}(t), u)].$$

2.2 NCO for Optimal Control Problems

2.2.1 The Problems

From a modern perspective, optimal control is a generalization of the calculus of variations.

As the name indicates, optimal control problems involve a control variable. In these problems the minimum cost depends both on the state and control variable. The control may be restricted to take values on a general set. The freedom to specify the set of possible controls combined with possibility of dealing with general cost functions covers a wide range of control engineering problems.

The mathematical formulation of a OCP appears in three forms: Bolza, Lagrange and Mayer problems.

We start by introducing the Bolza problem, as following:

$$(OCP_B) \quad \begin{cases} \text{Minimize} & g(x(t_0), x(t_1)) + \displaystyle\int_{t_0}^{t_1} L(t, x(t), u(t))dt \\ \text{subject to} & \dot{x}(t) = f\,(t, x(t), u(t)) \qquad \text{a.e.} t \in [t_0, t_1] \\ & (x(t_0), x(t_1)) \in C \\ & u(t) \in \Omega(t). \end{cases}$$

The data for this problem comprise functions $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $L : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $f : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, a closed set $C \subset \mathbb{R}^n \times \mathbb{R}^n$ and a multifunction $\Omega : [t_0, t_1] \rightsquigarrow \mathbb{R}^m$.

The function to minimize

$$g(x(t_0), x(t_1)) + \int_{t_0}^{t_1} L(t, x(t), u(t))dt \tag{2.1}$$

is known as cost function.

The variable x is called the state. The function describing state time evolution, $x(t)$, $t_0 \leq t \leq t_1$ is called *state trajectory*.

The set of *control functions* for (OCP_B), denoted U, is the set of measurable functions $u : [t_0, t_1] \to \mathbb{R}^m$ such that $u(t) \in \Omega(t)$ a.e. $t \in [t_0, t_1]$.

The domain of the above optimization problem is the set of *admissible processes*, namely pairs (x, u) comprising a control function u and a corresponding state trajectory x which satisfy the constraints of (OCP_B).

If the cost function (2.1) is simply

- $\displaystyle\int_{t_0}^{t_1} L(t, x(t), u(t))dt$, then the problem is known as *Lagrange problem*;
- $g(x(t_0), x(t_1))$, then the problem is known as *Mayer problem*.

The Bolza problem can be transformed in these two special problems, Lagrange and Mayer problem, by adding a new state variable, (see for example: [PF62], [Tor02]).

Example of optimal control problems in three forms can be found in [Ber95].

According to [Vin00], the importance of Mayer formulation is that it embraces a wide range of significant optimization problems which are beyond the reach of traditional variational techniques and it is very well suited to the derivation of general necessary conditions of optimality. In next chapters, we consider OCP in Mayer form.

Additional constraints can be added to the problem. For example:

- *equality state constraint:*
  $k(t, x(t)) = 0$ for $t \in [t_0, t_1]$, for a given function $k : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}$;
- *inequality state constraint*$^{(a)}$*:*
  $h(t, x(t)) \leq 0$ for $t \in [t_0, t_1]$, for a given function $h : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}$;
- *implicit state constraint*$^{(b)}$*:*
  $x(t) \in X(t)$ for $t \in [t_0, t_1]$, in which $X : [t_0, t_1] \rightsquigarrow \mathbb{R}^n$ is given multifunction;
- *mixed state constraint :*
  $g(t, x(t), u(t)) \leq 0$ for a.e. $t \in [t_0, t_1]$, in which $g : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ is given.

State constraints$^{(a),(b)}$ are object of study in this thesis. Problems with the mixed state constraint are considered in [Aru00], [dP03], and [MdRdPZ01].

Here, we consider fix-time problems. However, free-time problem could be considered, where the problem is defined on an interval $[t_0, t_0 + T]$ and it is desired minimize time T (see for example [PF62],[Ber95]). These problems are known as *minimal time problems.*

2.2.2 Maximum Principle

The NCO for OCP appear in the form of Maximum Principle (MP). It is usually accepted that the MP was introduced by Pontryagin and his collaborates in the paper [PBGM62].

The original formulation of the MP applied to problems with very basic restrictions and with smoothness hypotheses.

Assuming that, the (pseudo-) Hamiltonian function [1] is defined as follows:

$$H(t, x, p, u) = p \cdot f(t, x, u) - \lambda L(t, x, u).$$

The MP under smoothness hypotheses, states that if $(\bar{x}, \bar{u})$ is a minimizer of (OCP_B), then there exists an absolutely continuous function p and $\lambda \geq 0$, not both zero, such that the following conditions are satisfied:

The Adjoint Condition:

$$-\dot{p}(t) = H_x(t, \bar{x}(t), p(t), \bar{u}(t)) \quad \text{a.e.;}$$

The Maximum Principle: $\bar{u}(t)$ maximizes over $\Omega(t)$ the function

$$u \to H(t, \bar{x}(t), p(t), u) \quad \text{a.e.;}$$

The Transversality Condition:

$$(p(t_0), -p(t_1)) - \lambda g_x(\bar{x}(t_0), \bar{x}(t_1)) \text{ is normal to } C \text{ at } (x(t_0), x(t_1)).$$

A brief historical survey of NCO for optimal control and calculus of variations problems can be found in [Sar00].

2.3 Nonsmooth NCO

Optimization problems in which the cost function to minimize is not differentiable appear frequently. Two simple examples of nondifferentiable functions are:

[1] What Hamilton really defined was the "maximized" hamiltonian $H(t, x, p) = p \cdot v(x, p) - \lambda L(t, x, v(x, p))$, where v is not treated as independent variable.

- functions that state lengths and distances;
- function defined as the max or min of a collection of differentiable functions.

Others examples of problems with nonsmooth data can be found, for example, in [Cla83].

Nowadays, there exists a great interest in developing necessary conditions for problems with nonsmooth data.

2.3.1 Nonsmooth NCO for Calculus of Variations Problems

An extension of the Euler-Lagrange (see for example [Cla89]), allowing nonsmooth data is

$$(\dot{p}(t), p(t)) \in \tilde{\partial} L(t, \bar{x}(t), \dot{\bar{x}}\ (t)) \text{ a.e..}$$

Here, $\tilde{\partial} L$ denotes the *Clarke's subdifferential* with respect to (x, u).

If function $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous on a neighborhood of a point $x \in \mathbb{R}^n$, the *Clarke's subdifferential* is given by

$$\tilde{\partial} f(x) = co\ \{\eta \in \mathbb{R}^n : \exists x_i \to x, x_i \notin \Omega, f_x(x_i) \text{ exist and } f_x(x_i) \to \eta\},$$

where $\Omega \subset \mathbb{R}^n$ having Lebesgue measure zero.

We have defined this subdifferential only for Lipschitz continuous function, however Clarke provided an extension to lower semicontinuous functions, see [Cla89].

If L is continuously differentiable, then $\tilde{\partial} L(t, \bar{x}(t), \dot{\bar{x}}\ (t))$ reduces to the singleton set $\{L_{x,u}(t, \bar{x}(t), \dot{\bar{x}}\ (t))\}$.

2.3.2 Nonsmooth NCO for Optimal Control Problems

In this section, we introduce NCO for OCP in Mayer form with endpoint state constraints. Without loss of generality, we consider the interval $[0, 1]$ as the "time" domain of our problem. The problem of interest is:

$$(OCP_{M1}) \quad \begin{cases} \text{Minimize} & g(x(0), x(1)) \\ \text{subject to} & \dot{x}(t) = f\,(t, x(t), u(t)) \qquad \text{a.e.} t \in [0,1] \\ & x(0) \in C_0 \\ & x(1) \in C_1 \\ & u(t) \in \Omega(t) \qquad \text{a.e.} t \in [0,1]. \end{cases}$$

The data for this problem comprise functions $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $f : [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, the sets C_0 and C_1 and a multifunction $\Omega : [0,1] \rightsquigarrow \mathbb{R}^m$.

Remark 2.3.1 ***(On Differential Inclusions)*** *The control system*

$$\begin{cases} \dot{x}(t) = f\,(t, x(t), u(t)) \\ u(t) \in \Omega(t) \end{cases} \tag{2.2}$$

can be interpreted as

$$\dot{x}(t) \in F(t, x(t)) \quad a.e., \tag{2.3}$$

in which, for each (t, x), $F(t, x)$ is a given subset of $\mathbb{R}^n$.

If $F(t, x(t)) = f(t, x(t), \Omega(t))$, the set of solutions to (2.2) coincides with set of solutions to differential inclusion (2.3), under the following mild hypotheses on the data for (OCP_{M1}), (see [Vin00], pag.73):

(i) *$f(\cdot, x, \cdot)$ is $\mathcal{L} \times \mathcal{B}^m$ measurable and $f(t, \cdot, u)$ is continuous;*

(ii) *$Gr\,\Omega$ is $\mathcal{L} \times \mathcal{B}^m$ measurable.*

Remark 2.3.2 ***(On minimizer)*** *When we seek a solution of an optimal control problem, we must specify if we are looking for a local or a global minimizer. The meaning of local needs to be clarified. Different choices of topology on the set of admissible processes give rise to different notions of local minimizer.*

Throughout this thesis, we say that an admissible process $(\bar{x}, \bar{u})$ is a local minimizer *if there exists $\delta > 0$ such that*

$$g(\bar{x}(0), \bar{x}(1)) \leq g(x(0), x(1)),$$

for all admissible processes (x, u) *satisfying*

$$\|x(t) - \bar{x}(t)\|_{L^\infty} \leq \delta.$$

For (OCP_{M1}), the Hamiltonian function is

$$H(t, x, p, u) = p \cdot f(t, x, u).$$

We provide here a version of MP under minimum hypotheses in which it makes sense to talk about a OCP, as Clarke mentions in the paper [Cla76a]. They are denoted here and throughout as the Basic Hypotheses.

Theorem 2.3.3 *Let* $(\bar{x}, \bar{u})$ *be a local minimizer for* (OCP_{M1})*. Assume that, for some* $\delta' > 0$*, the following* ***Basic Hypotheses*** *are satisfied.*

$\mathbf{H1_b}$ *The function* $(t, u) \to f(t, x, u)$ *is* $\mathcal{L} \times \mathcal{B}^m$ *measurable for each* x*. (*$\mathcal{L} \times \mathcal{B}^m$ *denotes the product* σ*-algebra generated by the Lebesgue subsets* $\mathcal{L}$ *of* $[0, 1]$ *and the Borel subsets of* $\mathbb{R}^m$*.)*

$\mathbf{H2_b}$ *There exists a* $\mathcal{L} \times \mathcal{B}^m$ *measurable function* $k(t, u)$ *such that* $t \mapsto k(t, \bar{u}(t))$ *is integrable and*

$$\|f(t, x, u) - f(t, x', u)\| \leq k(t, u)\|x - x'\|$$

for $x, x' \in \bar{x}(t) + \delta'\mathbb{B}$*,* $u \in \Omega(t)$ *a.e.*$t \in [0, 1]$*.*

$\mathbf{H3_b}$ *The function* g *is Lipschitz continuous on* $\bar{x}(1) + \delta'\mathbb{B}$*.*

$\mathbf{H4_b}$ *The graph of* Ω *is* $\mathcal{L} \times \mathcal{B}^m$ *measurable.*

$\mathbf{H5_b}$ *The sets* C_0 *and* C_1 *are closed.*

Then there exist $p \in W^{1,1}([0, 1] : \mathbb{R}^n)$ *and* $\lambda \geq 0$ *such that*

$$\|p\|_{L^\infty} + \lambda > 0,$$

$$-\dot{p}(t) \in \operatorname{co} \partial_x^L H(t, \bar{x}(t), p(t), \bar{u}(t)) \quad a.e. t \in [0,1],$$

$$(p(0), -q(1)) \in N_{C_0}^L(\bar{x}(0)) \times N_{C_1}^L(\bar{x}(1)) + \lambda \partial^L g(\bar{x}(0), \bar{x}(1)), \tag{2.4}$$

and for almost every $t \in [0,1]$, $\bar{u}(t)$ maximizes over $\Omega(t)$

$$u \to H(t, \bar{x}(t), p(t), u).$$

Remark 2.3.4 *Here, co C is the convex hull of a set $C \subset \mathbb{R}^n$. The set $N_C^L(x)$ is the limiting normal cone to the closed set $C \subset \mathbb{R}^n$ at $x \in C$ defined as*

$$\begin{aligned} N_C^L(x) = \quad & \{\eta \in \mathbb{R}^n : \exists \text{ sequences } \{M_i\} \in \mathbb{R}^+, x_i \to x,\ \eta_i \to \eta \text{ such that} \\ & x_i \in C \text{ and } \eta_i \cdot (y - x_i) \leq M_i \|y - x_i\|^2 \text{ for all } y \in \mathbb{R}^n, i = 1, 2, ...\}. \end{aligned}$$

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function and $x \in$ dom f. Then the set $\partial^L f(x)$ defined as

$$\partial^L f(x) = \{\eta \in \mathbb{R}^n : (\eta, -1) \in N_{epi\ f}^L(x, f(x))\},$$

where epi $f = \{(x, \alpha) \in \mathbb{R}^{n+1} : \alpha \geq f(x)\}$, is the limiting subdifferential of f at x.

Further details of nonsmooth analysis involved are provided in the appendix.

2.3.3 Nonsmooth NCO for Optimal Control Problems with State Constraints

In this section, we introduce the MP for an OCP with inequality state constraints, as the following:

$$(OCP_{M2}) \quad \begin{cases} \text{Minimize} & g(x(0),x(1)) & \\ \text{subject to} & \dot{x}(t) = f\left(t, x(t), u(t)\right) & \text{a.e.} t \in [0,1] \\ & x(0) \in C_0 & \\ & x(1) \in C_1 & \\ & u(t) \in \Omega(t) & \text{a.e.} t \in [0,1] \\ & h(t,x(t)) \leq 0 & \text{for all } t \in [0,1]. \end{cases}$$

In [HSV95], we can find a survey of MP for problems with state constraints. There references to the direct adjoint approach, indirect adjoint approach, and methods that use transformations converting problems with state constraints into problems without state constraints are made. However, problems with nonsmooth data are not addressed.

The NCO for nonsmooth and state constraints OCP were introduced in [VP82]. This result generalized MP introduce by [Cla76a], by allowing state constraints in the form

$$h(t,x(t)) \leq 0, \text{ for all } t \in [0,1]. \tag{2.5}$$

They show that Clarke's methodology can be adapted to permit such constraints. The underlying idea is to replace constraints (2.5) by a penalty term added to the cost

$$g(x(0),x(1)) + k \int_0^1 \max\{0, h(t,x(t))\} dt,$$

for some $k > 0$.

Nonsmooth MP for state constrained problems are also proved in [Cla83] and [VZ98].

Next, we introduce the results from [Vin00], which is a refinement of Clarke's necessary conditions.

Assume that, in addition to $\mathbf{H1_b}$-$\mathbf{H5_b}$, the following hypothesis are imposed on (OCP_{M2}):

$\mathbf{H6_b}$ The function h is upper semicontinuous in t and there exists a scalar $K_h > 0$ such that the function $x \to h(t,x)$ is Lipschitz of rank K_h for all $t \in [0,1]$.

Then the MP is stated in the following form:

Theorem 2.3.5 *([Vin00]) If $(\bar{x}, \bar{u})$ is an local minimizer, then there exist $p \in W^{1,1}([0,1]:\mathbb{R}^n)$, measurable function γ, a nonnegative Radon measure $\mu \in C^*([0,1],\mathbb{R})$ and a scalar $\lambda \geq 0$ such that*

$$\mu\{[0,1]\} + \|p\|_{L^\infty} + \lambda > 0,$$

$$-\dot{p}(t) \in \operatorname{co}\partial_x^L H(t,\bar{x}(t),q(t),\bar{u}(t)) \quad a.e. t \in [0,1],$$

$$(p(0),-q(1)) \in N_{C_0}^L(\bar{x}(0)) \times N_{C_1}^L(\bar{x}(1)) + \lambda\partial^L g(\bar{x}(0),\bar{x}(1)),$$

$$\gamma(t) \in \partial_x^> h(t,\bar{x}(t)) \quad \mu - a.e.,$$

$$supp\ \{\mu\} \subset \{t \in [0,1] : h(t,\bar{x}(t)) = 0\}, \tag{2.6}$$

and, for almost every $t \in [0,1]$, $\bar{u}(t)$ maximizes over $\Omega(t)$,

$$u \to H(t,\bar{x}(t),q(t),u).$$

where

$$q(t) = \begin{cases} p(t) + \int_{[0,t)} \gamma(s)\mu(ds) & t \in [0,1) \\ p(t) + \int_{[0,1]} \gamma(s)\mu(ds) & t = 1. \end{cases}$$

Here, $\partial_x^> h(t,x)$, denotes the hybrid partial subdifferential of h in the x-variable defined as

$$\partial_x^> h(t,x) = \operatorname{co}\{\xi : \text{ there exist } (t_i,x_i) \to (t,x) \text{ s.t. } h(t_i,x_i) > 0, \\ h(t_i,x_i) \to h(t,x), \text{ and } h_x(t_i,x_i) \to \xi\}.$$

The condition (2.6) is denoted by "*Complementary Slackness Condition*"; it states that μ is equal to zero if the state constraint is inactive at x on t (i.e. $h(t, x(t)) < 0$).

Note that if the state constraint is inactive at x, then the statement of the theorem simplifies due to the fact that all mention to μ (and the corresponding integrals) may be removed.

It is worth mentioning that introduction of chapter 9 in the book of [Vin00], we can find the key ideas behind the derivation of NCO for problem with state constraints.

2.4 Existence and Regularity

Application of NCO to identify a set of candidates to the optimal solution only make sense if the optimal solution exists. Therefore, there is great interest in studding the existence of optimal solutions.

It was Tonelli (1915) who introduced the first theorem of the existence of solution for CVP. Even today, the Tonelli's theorem remains the central existence theorem for CVP, although the hypotheses of the theorem can be relaxed, see for example [Vin00]. For OCP, results that guarantee the existence of solution can be found in [Cla83], for example.

The hypotheses under which existence of an optimal solution may not coincide with those under which NCO are valid.

A simple example of that occurs in calculus of variations: the Tonelli's theorem guarantee the existence of minimizers in the class of absolutely continuous functions, whereas the Euler-Lagrange condition is applied for arcs with essentially bounded derivatives.

Regularity analysis helps us to identify classes of problems, for which all minimizers satisfy known NCO. This analysis seeks information about regularity of miniminizers, for example when the minimizers arcs are Lipschitz continuous (we call Lipschitz regularity), minimizers arcs with higher-order derivatives or the optimal control that are Lipschitz continuous.

In recent years many authors got interested on the study of the Lipschitz regularity of the optimal trajectory, because of its important implications. In Control Engineering, this regularity condition allows to compute the true optimal trajectory by numerical methods. On other hand this condition ensures the non-occurrence of the Lavrentiev phenomenon - the infimum cost over the space of absolutely continuous functions is strictly less than the infimum cost over the space of Lipschitz continuous functions. A simple example in which this phenomenon occurs was given by Maniá, (see for example [Cla89]).

Many authors contributed to the investigation of Lipschitzianity of optimal trajectories for CVP, see for example [CV85] and [Vin00]. Less is known for OCP. In this respect we refer the reads to [ST00] and [GV03] for OCP, (where the controlled differential equation is linear in the control variable), the result of [DK95] and [CLV97] for linear quadratic problems with state constraints. However, Lipschitz regularity of the optimal trajectory for nonlinear OCP with state constraints is still an open question.

Chapter 3

The Degeneracy Phenomenon of Necessary Conditions of Optimality

In this chapter, we discuss the degeneracy phenomenon in optimization problems with inequality constraints. We start by describing this phenomenon in the context of mathematical programming problems, recalling the Fritz-John and Kuhn-Tucker conditions. Later, we address the degeneracy phenomenon in the context of optimal control problems. We review and discuss nondegenerate necessary conditions of optimality for optimal control problems with state constraints. An overview of the main literature in this area is made, including a comparison with some recent results from the authors.

3.1 Degeneracy in Mathematical Programming

The general mathematical programming problem (MPP) consists in minimizing a given function $f(x)$ subject to three types of constraints: inequality constraints, equality constraints and implicit state constraints. Here, we consider the MPP with inequality constraints:

$$(MPP_1)\quad \begin{cases} \text{Minimize} & f(x) \\ \text{subject to} & g^i(x) \leq 0,\ i = 1, 2, ..., n. \end{cases}$$

Throughout this section, we assume that the functions f and g^i for each $i = 1, 2, ..., n$ are continuously differentiable.

If $\bar{x}$ is a solution to the problem (MPP_1), then the NCO in the form of Fritz-John conditions [Joh48] in [BSS93] guarantee the existence of nonnegative multipliers λ and μ_i, with $i = 0, 1, 2, ..., n$ such that

$$(\lambda, \mu_1, ..., \mu_n) \neq 0 \tag{3.1}$$

$$\lambda f_x(\bar{x}) + \sum_{i=1}^{n} \mu_i g_x^i(\bar{x}) = 0 \tag{3.2}$$

$$\mu_i g^i(\bar{x}) = 0, \quad \text{for } i = 1, ..., n. \tag{3.3}$$

If the second condition is satisfied with $\lambda = 0$, the cost function is not involved in the choice of candidates to minimizers. So, the NCO does not give any information about the candidate to minimizers and the NCO are merely a relation between the constraints. When this happens, we say that the NCO are degenerated.

A way of forcing the cost function to be involved in the NCO is to assume that $\lambda = 1$ on the conditions (3.1)-(3.3), known as normal form of the NCO. However, we have to guarantee that the NCO are still satisfied at local minimum. If it is not the case, the NCO are not valid. So additional hypotheses, known as Constraint Qualification (CQ), are considered to identify the problems under which the normal form is ensured.

Some of the best known examples of a CQ are:

Linear Independence CQ: for every local minimizer $\bar{x}$, the gradients of the active constraints are linearly independent;

Mangasarian-Fromovitz CQ: for every local minimizer $\bar{x}$ there exists a vector

$v \in \mathbb{R}^n$ such that

$$g^i{}_x(\bar{x}) \cdot v < 0 \quad \text{if } g^i(\bar{x}) = 0, \quad i = 1, 2, ..., n.$$

Another type of CQ, called "calmness", was introduced in [Cla73], see [Cla83].

Assume that $\bar{x}$ is minimizer to (MPP_1) and $P(p)$ is the problem of minimizing $f(x)$ over points $x \in \mathbb{R}^n$ which satisfy the constraints $g(x) + p \leq 0$. The (MPP_1) is calm at $\bar{x}$ provided that there exist positive ε and M such that, for all $p \in \varepsilon\mathbb{B}$, for all $x' \in \bar{x} + \varepsilon\mathbb{B}$ which are feasible for $P(p)$, one has

$$f(x') - f(\bar{x}) + M\|p\| \geq 0.$$

The calmness of MPP at $\bar{x}$ allows to write the NCO (3.1)-(3.3) with $\lambda = 1$.

The Kuhn-Tucker conditions [KT51] are precisely a normal version of the Fritz John conditions valid under a suitable CQ. They state that the conditions (3.1)-(3.3) can be written with $\lambda = 1$ for all problems complying with the CQ.

The work of Kunh and Tucker, probably one the most cited results in optimization, is in fact a strengthened and nondegenerate form, of the Fritz John conditions. This fact justifies the importance of studying nondegenerate versions of NCO for constrained optimization problems. This problem is well-studied in the context of mathematical programming for along time. However, the degeneracy phenomenon in the OCP context has witness many important advances in the very recent years.

3.2 Degeneracy in Optimal Control Problems

In this section, we discuss strengthened forms of MP for OCP, like (OCP_{M2}), which guarantee nondegeneracy and/or normality.

The term "degeneracy" has been used in optimal control literature to describe a particular type of degeneracy occurring due to the presence of pathwise state constraints which are active at the initial time. Assuming that the pathwise state constraint is active in the initial instant of time, i.e.

$$h(0, x_0) = 0, \tag{3.4}$$

the set of multipliers (degenerate multipliers)[1]

$$\lambda = 0,\ \mu = \delta_{t=0}, p = -h_x(0, x_0) \tag{3.5}$$

satisfies the NCO for all admissible process (x, u). This can be easily seen by noting that the quantity $p(t) + \int_{[0,t)} h_x(s, \bar{x}(s))\mu(ds)$ vanishes almost everywhere and all conditions of the MP, (Theorem 2.3.5), are satisfied independently of the value of $\bar{x}$ or $\bar{u}$. In this case, the NCO are said to be degenerate.

In this thesis we will be concentrated in this kind of degeneracy. However other type of degeneracy can occur, namely "the q-degeneracy" (see for example [Fon99]).

The case (3.4) is encountered in certain applications of interest, namely Model Predictive Control. A further discussion of this point can be seen in [FV94, Fon99].

In order to avoid the degeneracy, the MP can be strengthened with additional conditions, typically a strengthened form of the nontriviality condition.

The term normality is used when the MP for OCP can be written with the multiplier associated with the objective function λ not zero.

Definition 3.2.1 *(Normality) An optimal control problem is said normal if the conditions of Theorem 2.3.5 are satisfied with $\lambda = 1$.*

.

The normality and regularity [2] are closely connected.

In [Fer06], it is proved that the conditions imposed to get the Lipschitz continuity of the optimal control may also contribute to guarantee the normality of MP.

Results where Lipschitz regularity is ensured as a consequence of normal NCO, can be found in [FM06].

[1] Here $\delta_{\{0\}}$ denotes the unit measure concentrated at $\{0\}$.
[2] The term regularity as the same meaning as in section 2.4.

Normal necessary conditions have been developed for problems with nonsmooth as well as smooth data, problems in which the dynamic constraints involves a differential inclusion, or a differential equation, and in which the state constraint is formulated as a set inclusion as well as a functional inequality.

In next section, we make an overview of the main literature in these area.

3.2.1 Avoiding the Degeneracy Phenomenon

Calmness

As in mathematical programming, the new type of CQ introduce by Clarke "calmness" allow to strength the MP with $\lambda = 1$.

For the problem

$$(OCP_{M3}) \quad \begin{cases} \text{Minimize} & g(x(0), x(1)) \\ \text{subject to} & \dot{x}(t) \in F(t, x(t)) \quad a.e., \end{cases}$$

calmness is defined as follows:

Definition 3.2.2 *Let $\phi^i : \mathbb{R}^n \to [-\infty, \infty]$, $i = 0, 1$, be defined as*

$$\phi^0(s) = \inf\{g(x(0) + s, x(1)) : \dot{x} \in F(t, x(t)) \;\; a.e. \;\},$$
$$\phi^1(s) = \inf\{g(x(0), x(1) + s) : \dot{x} \in F(t, x(t)) \;\; a.e. \;\}.$$

Then, problem (OCP_{M3}) is said to be calm if, for $i = 0$ or 1,

$$\phi^i = \liminf_{s \to 0} \left[\phi^i(s) - \phi^i(0)\right] / |s| > -\infty.$$

As shown in [Cla76b], calmness allows to write the MP with $\lambda = 1$, when $F(t, x)$ is measurable in t and Lipschitz in x near $\bar{x}$ and $g : \mathbb{R}^n \times \mathbb{R}^n \to (-\infty, \infty]$ is lower semicontinuous. However, pathwise state constraints are not considered.

In the remaining of these sections, we consider OCP with pathwise state constraints: inequality constraints or implicit constraints.

Nondegenerate Result from [AA97]

In [AA97], a new MP is developed to avoid the degeneracy, for Lipschitz continuous trajectories where the problem is:

$$(OCP_{M3}) \quad \begin{cases} \text{Minimize} & g(x(0), x(1)) \\ \text{subject to} & \dot{x}(t) \in F(t, x(t)) & \text{a.e.} t \in [0,1] \\ & (x(0), x(1)) \in C_0 \times C_1 \\ & x(t) \in X & \forall t \in [0,1], \end{cases}$$

The MP contains additional information about the behavior of the Hamiltonian at the endtimes:

$$\tilde{H}\left(t, \bar{x}(t), p(t) + \int_{[0,t)} d\mu\right) = \tilde{H}\left(t, \bar{x}(t), p(t) + \int_{[0,t)\cup\{t\}} d\mu\right) \tag{3.6}$$

for all $t \in [0,1]$ where:

- $\tilde{H}(t, x, q) := \max_{f \in F(t,x)} q \cdot f$ is the true (maximized) Hamiltonian;
- supp $\mu \subset \{t \in [0,1] : x(t) \in bdy\,(X)\}$; $\mu(t) \in N_X(\bar{x}(t)) \, \forall t \in [0,1]$.

(N_X is Clarke normal cone)

The condition (3.6), combined with the following constraint:

CQ$_{\mathbf{AA97}}$

$$\tilde{H}(0, x_0, -g) > 0,$$

$$\forall g \in N_X(\bar{x}(0)) \cap N^L_{(C_0 \cap X)}(\bar{x}(0)).$$

eliminates the degenerate multipliers.

Loosely speaking, **CQ$_{\mathbf{AA97}}$** requires the existence of a control function pulling the state away from the state constraint boundary at the initial time.

For the results in [AA97] to be valid, it is required that the multivalued mapping F is locally Lipschitz with nonempty convex compact values.

Nondegenerate Result from [FV94]

Another result to avoid degeneracy is developed in [FV94]. It also required $f(t, x, \Omega(t))$ to be convex but data are merely required to be measurable in time. For a problem like (OCP_{M2}) (see section 2.3.3) whith initial state fixed and free final state, the nondegeneracy NCO are strengthened with the nontriaviality condition

$$\int_{(0,1]} \mu(ds) + \lambda > 0,$$

if one of the following CQ are satisfied:

$\mathbf{CQ_{1FV94}}$: there exists a control $\tilde{u}$ such that

$$h_x(t, x_0) \cdot [f(t, x_0, \tilde{u}) - f(t, x_0, \bar{u}(t)] < 0,$$

for t near 0 (that means, there exits control function pulling the state away from the boundary of the state constraint set faster than the optimal control);

$\mathbf{CQ_{2FV94}}$: : there exists $\bar{t} \in (0, 1]$ such that $h(t, \bar{x}(t)) < 0, \forall t \in (0, \bar{t}]$, (that means, the minimizing trajectory itself leaves the boundary immediately).

Conditions to ensure normality are described in terms of the dynamic equations, linearized with respect to the state variables. The constraints qualifications $\mathbf{CQ_{1FV94}}$ and $\mathbf{CQ_{2FV94}}$ are strengthened with the following condition:

$\mathbf{CQ_{3FV94}}$:

$$h_x(t, \bar{x}(t)) \cdot y_u(t) < 0 \ \forall t \in (0, 1] \cap \{t \in [0, 1] : h(t, \bar{x}(t)) = 0\},$$

where y_u is the unique absolutely continuous function satisfying:

$$\begin{aligned} &\dot{y}_u(t) = f_x(t, \bar{x}(t), \bar{u}(t)) \cdot y_u(t) + f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t)) \quad \text{a.e. } t \in [0, 1] \\ &y_u(0) = 0, \end{aligned} \tag{3.7}$$

given a control u.

Nondegenerate Result from [FFV99]

The result in [FFV99] generalizes the nondegenerate result in [FV94] with $\mathbf{CQ_{1FV94}}$ by allowing the final state to belong to a given set C_1, the data to be nonsmooth and by not requiring the velocity set $f(t, x, \Omega(t))$ to be convex. In this paper new methods are introduced for proving nondegenerate NCO. The key idea of the proof is to replace the original control problem by one in which the state constraints is eliminated on $[0, \alpha]$, for arbitrary small α.

The multipliers of the MP for this new problem are nondegenerate. Passing to the limit $\alpha \downarrow 0$ we concluded that the limiting multipliers are nondegenerate and the nontriviality condition can be replaced by

$$\mu\{(0, 1]\} + \|q\|_{L^\infty} + \lambda > 0.$$

Normality Result from [Fon00]

Based on nondegeneracy results in [FFV99], [Fon00] ensures the normality of the MP for free final state problem, if there is a control that can pull the trajectory away from the boundary (faster than the optimal control) for every instant that inequality constraints is active.

In the works mentioned above ([FV94], [FFV99], and [Fon00]), the conditions involve the minimizing $\bar{u}$ which we do not know in advance, and consequently the conditions are, in general not easily verifiable, except in special cases, such as CVP. (See next chapter)

Normality Result from [RV99]

Nondegenerate NCO for OCP valid under a CQ that no longer involve the minimizing $\bar{u}$, appear in [RV99]. The MP can be written with $\lambda = 1$, if

$\mathbf{CQ_{RV99}}$**:** there exists a continuous feedback $u = \eta(t, \xi)$ such that

$$h_t(t, \xi) + h_x(t, \xi) \cdot f(t, \xi, \eta(t, \xi)) < -\delta,$$

for some positive δ, whenever (t, ξ) is close to the graph of $\bar{x}(\cdot)$ and ξ is near the state constraint boundary.

The problem considered is (OCP_{M2}), but the functions defining the dynamics is now Lipschitz continuous with respect to time, the final state is free, and the initial state belongs to a given set C_0.

The proof of existence of normal multipliers is based on a main theorem, called neighbouring feasible trajectories theorem. It asserts that for a prespecified process which may violate the state constraint there exists another process that it is suitably close to the first one and satisfies the state constraint.

Nondegenerate Result from [RV00]

Building upon their neighbouring feasible trajectories theorem, [RV00] derived non-degenerate NCO which apply to differential inclusion problems (OCP_{M3}) where the state constraints set X takes the form:

$$X = \bigcup_{j=1}^{m} \{x : h^j(x) \leq 0\}$$

for some functions $h^j : \mathbb{R}^n \to \mathbb{R}$, $j = 1, ..., m$ of class $C^{1,1}$.

Assuming that the velocity set $F(t, x)$ is nonconvex and measurable in time, the NCO are strengthened with the nontriviality condition

$$\lambda + \int_{(0,1]} \sum_j \mu^j(ds) + |p(0) + \sum_j h_x^j(\bar{x}(0))\mu^j(\{0\})| \neq 0,$$

when subject to follow constraint qualification:

CQ$_{\mathbf{RV00}}$: For each $t \in [0, \epsilon]$ and $\xi \in \bar{x}(0) + \delta' \mathbb{B}$

$$min_{v \in F(t,\xi)} h_x^j(\xi) \cdot v < -\delta$$

for all index values j such that $h^j(\bar{x}(0)) = 0$.

Nondegenerate Result from [CF05]

In the paper [CF05], we can find a strengthened MP for (OCP_{M3}) with dynamics given by a nonconvex differential inclusion and fixed initial state.

To derive these results, it was necessary impose the following hypotheses:

Hypothesis 3.1.

i) $F(\cdot,\cdot) : [0,1] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ is a multifunction with nonempty closed values.

ii) $\forall x \in \mathbb{R}^n$, $F(\cdot, x)$ is measurable.

iii) There exists $c > 0$ such that $\forall (t,x) \in [0,1] \times \mathbb{R}^n$, $F(t,x) \subset c(1 + \|x\|)\mathbb{B}$.

iv) There exists $l(\cdot) \in L^1$ such that $F(t,\cdot)$ is $l(t)$- Lipschitz continuous.

v) $g : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz.

Hypothesis 3.2. (*Used to establish the existence of a "linearization" of F along* $(\bar{x}, \dot{\bar{x}})$ *by closed convex processes, which are Lipschitz with respect to the state.*) There exists of a family of closed convex process $A(t,\cdot) : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$, $t \in [0,1]$, that satisfies

i) $A(\cdot, v)$ is measurable $\forall v \in \mathbb{R}^n$.

ii) $A(t,v) \subseteq \bar{d}_x \overline{co} F(t, \bar{x}(t), \dot{\bar{x}}\,(t))v \ \forall v \in \mathbb{R}^n$ for a.e. $t \in [0,1]$.

iii) For some $m \geq 0$, $A(t,\cdot)$ is m-Lipschitz on $\mathbb{R}^n$ for a.e. $t \in [0,1]$.

($\bar{d}_x\, F(\cdot)$ is the adjacent derivative of $\overline{co}F(t,\cdot)$ at $(\bar{x}(t), \dot{\bar{x}}\,(t))$, see appendix.)

Hypothesis 3.3. (*Used to the existence of a convex "linearizations" of constraints along optimal trajectories is also considered.*) X and C_1 are closed subsets of $\mathbb{R}^n$, $Int\,(C_{C_1}(\bar{x}(1))) \neq \emptyset$ and there exists a lower semicontinuous multifunction $G : [0,1] \rightsquigarrow \mathbb{R}^n$ such that for all $t \in [0,1]$, $G(t)$ is a closed convex cone with nonempty interior and for every $v \in Int(G(t))$ we can find $\varepsilon > 0$ such that for all $s \in [t-\varepsilon, t+\varepsilon] \cap [0,1]$, $\bar{x}(s) + [0,\varepsilon](v + \varepsilon\mathbb{B}) \subset X$.

($C_{C_1}(\bar{x}(1))$ denotes the Clarke tangent cone to C_1 at $\bar{x}(1)$, see appendix.)

Theorem 3.2.3 *Let $\bar{x}(\cdot)$ be an optimal solution to (OCP_{M3}) with initial state fixed assume that* ***Hypotheses 3.1-3.3*** *hold true. Further assume that an upper semi-continuous concave positively homogeneous function $\psi : \mathbb{R}^n \to \mathbb{R} \cap \{-\infty\}$ satisfies $Int\,(G(0)) \subset dom(\psi)$ and $\psi \leq D_x^+ V(0, \bar{x}(0))$. Then there exits $\lambda \in \{0,1\}$, $\psi \in NVB([0,1])$ and an absolutely continuous function $p(\cdot) : [0,1] \to \mathbb{R}^n$ such that $\lambda + \|\psi\|_{TV} \neq 0$ and p satisfies the*

$$\dot{p}(t) \in A^*(t, -p(t) - \psi(t)) \quad a.e.\ in\ [0,1]$$

$$p(1) \in -\lambda\tilde{\partial}g(\bar{x}(1)) - \psi(1) - N_{C_1}(\bar{x}(1)),$$

$$(p(t) + \psi(t)) \cdot \dot{\bar{x}}\,(t) = \max_{v \in F(t, \bar{x}(t))} (p(t) + \psi(t)) \cdot v a.e.\ in\ [0,1]$$

$$-p(0) \in \lambda\partial^+\psi(0).$$

Furthermore,

$$\psi(0^+) \in G(0)^-, \quad \psi(t) - \psi(t^-) \in G(t)^-, \quad \psi(t) = \int_{[0,t]} \nu(s)d\mu(s) \quad \forall t \in (0,1]$$

for a positive (scalar) Randon measure μ on $[0,1]$ and a μ-measurable function $\nu(\cdot) : [0,1] \to \mathbb{R}^n$ satisfying

$$\nu(s) \in G(s)^- \cap \mathbb{B} \quad \mu - a.e..$$

If $C_{C_1}(\bar{x}(1)) \cap Int\,(G(1)) \neq \emptyset$, then the following non degeneracy condition holds true

$$\lambda + \sup_{t \in (0,1)} \|p(t) + \psi(t)\| \neq 0 \tag{3.8}$$

and if $\bar{x}(1) \in Int\,(C_1)$, then

$$\lambda + var(\psi, (0,1]) \neq 0, \tag{3.9}$$

where $var(\psi, (0,1])$ denotes the total variation of ψ on $(0,1]$.

Moreover $\lambda = 1$ *if there exists a solution to the constrained differential inclusion*

$$\dot{w}(t) \in \overline{A(t,w) + T_{co(F(t,\bar{x}(t)))}(\dot{\bar{x}}\ (t))}, \tag{3.10}$$

satisfying

$$w(t) \in Int(G(t))\,\forall t \in [0,1], w(1) \in Int(C_{C_1}(\bar{x}(1))). \tag{3.11}$$

Here:

- $NVB([0,1])$ is the space of functions f of bounded variation on $[0,1]$, which are continuous from the right on $(0,1)$ and such that $f(0) = 0$;
- The norm of $f \in NVB([0,1])$ is the total variation of f on $[0,1]$ denoted by $\|f\|_{TV}$;
- G^- is the negative polar cone of G;
- $N_X(x)$ is the Clarke normal cone to the set X at $x \in X$;
- $T_{co(F(t,\bar{x}(t)))}(\dot{\bar{x}}\ (t))$ denotes the tangent cone of convex analysis to $co(F(t,\bar{x}(t)))$ at $\dot{\bar{x}}\ (t)$;
- $V(\cdot,\cdot)$ is the value function

$$V(t_0,y_0) = \inf \left\{ \begin{array}{ll} g(x(1)) : x(\cdot) \text{ is the solution to} & \\ \qquad \left\{ \begin{array}{l} \dot{x}(t) \in F(t,x(t)) \\ (x(0),x(1)) \in C_0 \times C_1 \\ x(t) \in X \\ x(t_0) = y_0 \end{array} \right. & \text{on } [t_0,1], \end{array} \right\}$$

- $D^+V(0,x_0)(\cdot)$ denotes the upper derivative of $V(0,\cdot)$ at x_0;
- $\partial^+\psi(0)$ denotes the superdifferential of ψ at 0;

(Definitions can be found in Appendix.)

Any of the conditions (3.8) and (3.9) eliminates the trivial multipliers for $\bar{x}(0)$.

To prove Theorem 3.2.3, duality of convex analysis is applied. In this way they extend the known relations between the maximum principle and dynamic programming from the unconstrained problems to constrained cases, where the calmness of value function is used to investigate nondegeneracy of MP.

To allow to write the Theorem 3.2.3 with $\lambda = 1$, it was necessary assume the following additional hypothesis:

Hypothesis 3.4. Assume that for some $\eta > 0$ the signed distance

$$h(x) = \begin{cases} -dist(x, \text{bdy } (X)) & \forall x \in X \\ dist(x, \text{bdy } (X)) & otherwise \end{cases}$$

is of class $C^{1,1}_{loc}$ on $bdy\, X + \eta\mathbb{B}$

and the following CQ :

$\mathbf{CQ_{CF05}}$: there exists $\delta > 0$ such that for almost all $t \in [0, 1]$ with $\bar{x}(t) \in bdy\,(X) + \eta\mathbb{B}$ we have

$$\min_{v \in F(t,\bar{x}(t))} h_x(\bar{x}(t)) \cdot v \leq -\delta.$$

Theorem 3.2.4 *Let $\bar{x}(\cdot)$ be an optimal control solution to (OCP_{M3}) with initial state is fixed assume that **Hypotheses 3.1, 3.2** and **3.4** hold true and $\bar{x}(1) \in Int\,(C_1)$. Then all conclusions of Theorem 3.2.3 are valid with $\lambda = 1$ and $G(t) = T_X(\bar{x}(t))$ for every $t \in [0, 1]$.*

The prove is based on ensured the existence of a function like (3.10) satisfying (3.11).

This result is similar to [RV99], however in this one the inward pointing condition is weaker condition, it has to be satisfy just along the optimal trajectory.

Normality Result from [BF07]

For a Bolza problem like (OCP_B) with Lipschitz continuous trajectories, where the initial state belongs to a given set, the final state is free and trajectories are constrained to a closed set $x(t) \in X$, the normality is ensured in [BF07].

This result is valid for problems that satisfy the following constraint qualification:

$\mathbf{CQ_{BF07}}$: $\exists \delta_R > 0$ such that, $\exists u_y \in \Omega(t)$ satisfying

$$\sup_{n \in C_X^-(x) \cap S^{n-1}} n \cdot f(t, y, u_y) \leq -\delta_R,$$

where $S^{n-1} = \{x \in R : \|x\| = 1\}$ and $C_X^-(x)$ is the negative polar of Clarke's tangent cone to the set X at x. (see definition in Appendix)

The advantage of this result is that it allows nonsmooth and nonconvex state constraints.

Normality Result from [Mal03]

The normality for an optimal control problem with mixed control-state and pure state constraints is described in the work of Malanowski [Mal03]. The constraints qualifications involve the gradients of the constraints which are in some sense almost active and involve also the controllability of the linearized state equation. If the data are differentiable and the constraint qualification mention above is satisfied, then there exists an unique normal Lagrange multiplier.

Comments

In summary, the constraint qualifications found in the literature to avoid degeneracy in optimal control problems with state constraints can be divided into two types:

$\mathbf{CQ1_d}$: (from [FV94] and [FFV99]) $\exists \delta, \epsilon > 0$ and $\exists \tilde{u}(t) \in \Omega(t)$:

$$h_x(x_0) \cdot [f(t, x_0, \tilde{u}(t)) - f(t, x_0, \bar{u}(t)] < -\delta \quad \text{a.e.} t \in [0, \epsilon).$$

Loosely speaking, this is the requirement that there exist a control function pulling the state away from the boundary of the state constraint set faster than the optimal control on a neighborhood of the initial time. (see Figure 3.1)

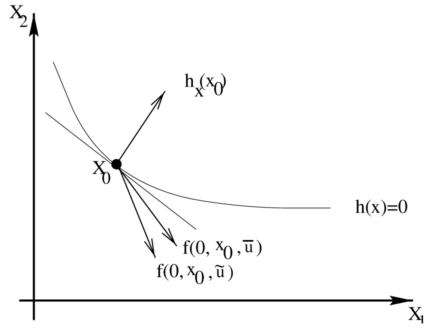

Figure 3.1: Constraint qualification $\mathbf{CQ1_d}$ (adapted from [Fon99]).

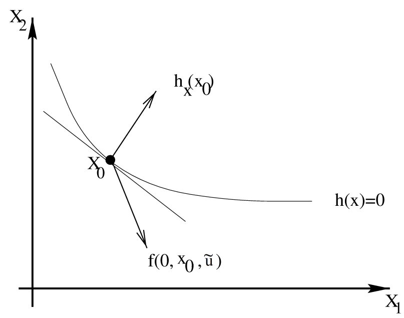

Figure 3.2: Constraint qualification $\mathbf{CQ2_d}$ (adapted from [Fon99]).

$\mathbf{CQ2_d}$**:** (from [AA97] and [RV00]) $\exists \delta, \epsilon > 0$ and $\exists \tilde{u}(t) \in \Omega(t)$:

$$h_x(x_0) \cdot f(t, x_0, \tilde{u}(t)) < -\delta \quad \text{a.e.} t \in [0, \epsilon).$$

Meaning, that there exits a control functions pulling the state away from the state constraint boundary on a neighborhood of the initial time. (see Figure 3.2)

Extending $\mathbf{CQ1_d}$ and $\mathbf{CQ2_d}$, in such way that they are verifiable not only on a neighborhood of the initial time, but also on neighborhood of each instant that the minimizer trajectory touches the boundary, allows to write the MP with $\lambda = 1$. Here, we denote by $\mathbf{CQ1_n}$ and $\mathbf{CQ2_n}$ (respectively), the constraint qualification ensuring the normality of MP. See for example [Fon00], [CF05] and [BF07].

Clearly, a normal form of MP implies a nondegenerate form of MP. However most of these results require some regularity on data. See for example [FV94], [RV99], [Fon00], [CF05] and [BF07].

[RV99], [CF05] and [BF07] use constraints qualifications of the type $\mathbf{CQ2_n}$, while [FV94] and [Fon00] use the constraint qualification of the type $\mathbf{CQ1_n}$. Comparing theses results, we conclude that normal MP using constraint qualification of the type $\mathbf{CQ1_n}$, as in [Fon00], requires less regularity. However, $\mathbf{CQ1_n}$ involves the minimizing $\bar{u}$ which we do not know in advance, and consequently the condition is, in general not easily verifiable.

Notes on Chapter

Part of the contents of this chapter have been presented in [LF07].

Chapter 4

Normality in Calculus of Variations Problems

In this chapter, we show how calculus of variations problems (CVP) can be seen as a particular case of an optimal control problems (OCP) and we study normality of necessary conditions of optimality (NCO) for CVP as a consequence of the normality of NCO for OCP.

4.1 Introduction

Throughout this chapter, we focus the following CVP subject to inequality states constraints:

$$(CVP_2) \quad \begin{cases} \text{Minimize} & J[x] = \displaystyle\int_0^1 L(x(t), \dot{x}(t))dt \\ \text{subject to} & x(0) = x_0 \\ & h(x(t)) \leq 0 \qquad\qquad \forall t \in [0,1]. \end{cases}$$

Observe that the functional defining the state constraints does not depend explicitly on t.

The special structure of CVP permits the derivation of constraint qualifications (CQ) that can be much easier to verify than in the optimal control case. Hence, the interest in exploring dynamic optimization problems with this special structure.

As mentioned before, here we discuss the normality results of OCP, in the particular of CVP. Therefore, we start by seeing the (CVP_2) as a special case of (OCP_{M2}). For that it is enough to consider a new absolutely continuous state variable

$$z(t) = \int_0^t L(x(s), \dot{x}(s))ds$$

and a change of variable $\dot{x}(t) = u$.

The (CVP_2) can then be written as:

$$(OCP_{M4}) \quad \begin{cases} \text{Minimize} & y(1) & \\ \text{subject to} & \dot{z}(t) = f(z(t), u(t)) & \text{a.e.} t \in [0,1] \\ & (x(0), y(0) = (x_0, 0) & \\ & u(t) \in \mathbb{R}^n & \\ & h(x(t)) \leq 0 & \forall t \in [0,1] \end{cases}$$

with $z(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ and $f(z(t), u(t)) = \begin{pmatrix} u(t) \\ L(x(t), u(t)) \end{pmatrix}$.

CQ ensuring normality of OCP with state constraints of the form $h(x(t)) \leq 0$ are of two types:

$\mathbf{CQ1_n}$:$\exists \tilde{u}(t) \in \Omega(t)$:a.e.$t \in [0,1]$

$$\zeta \cdot [f(t, \bar{x}(t), \tilde{u}(t)) - f(t, \bar{x}(t), \bar{u}(t))] < -\delta,$$

for all $\zeta \in \partial^{>} h(\bar{x}(s))$, when $s \in \{t \in [0,1] : h(\bar{x}(t)) = 0\}$, where $\partial^{>} h(x)$ is defined as:

$$\partial^{>} h(x) = \text{co}\ \{\varepsilon : \exists x_i \to x : h(x_i) > 0 \forall i,\ h(x_i) \to h(x) \text{ and } h_x(x_i) \to \varepsilon\}.$$

$\mathbf{CQ2_n}$: $\exists\epsilon > 0$ and $\tilde{u}(t) \in \Omega(t)$:

$$h_x(\bar{x}(t)) \cdot f(t, \bar{x}(t), \tilde{u}(t)) < -\delta,$$

for $t \in (s - \epsilon, s + \epsilon)$ where $s \in \{t \in [0, 1] : h(\bar{x}(t)) = 0\}$. [1]

In the work of [FV94], the normality of NCO for smooth CVP is guaranteed for problems that satisfied the following constraint qualification:

$\mathbf{CQ_{4FV94}}$ $h_x(\bar{x}(t)) \neq 0$ for $t \in \{s \in [0, 1] : h(\bar{x}(s)) = 0\}$.

Two question arises:

- Since the work of [Fon00] allows possibly nonsmooth data for OCP, do we have strengthened NCO for CVP with nonsmooth data applying the normality result in [Fon00]?
- does the CQ of type $\mathbf{CQ2_n}$ give new information when it is applied to CVP?

The answers to these questions are in next sections.

4.2 Normality in CVP Applying the Normal Result of [Fon00]

Applying the normal result of [Fon00] in CVP, we can obtain a strengthened NCO with nonsmooth data. This work was developed in [Lop03] and we mentione the result.

Assume that the following hypotheses are satisfied:

$\mathbf{H1n_{CV}}$ The function $x \to L(x, u)$ is locally Lipschitz continuous for all $u \in \mathbb{R}^n$.

$\mathbf{H2n_{CV}}$ The function $u \to L(x, u)$ is convex and bounded for all $x \in \mathbb{R}^n$.

[1]In [RV99], this CQ have to be satisfy on neighborhood of state constraint boundary, we not consider here to simply the notation.

H3n$_{\mathbf{CV}}$ There exists an increasing function $\theta : [0, \infty) \to [0, \infty)$ such that

$$\lim_{\alpha \to \infty} \frac{\theta(\alpha)}{\alpha} = +\infty,$$

and a constant β such that $L(x, v) > \theta(\|v\|) - \beta\|v\|$ for all $x \in \mathbb{R}^n$, $v \in \mathbb{R}^n$.

H4n$_{\mathbf{CV}}$ There exists a scalar $K_h > 0$ such that the function $x \to h(x)$ is Lipschitz continuous of rank K_h for all $t \in [0, 1]$.

Consider also the following constraint qualifications:

CQ$_{\mathbf{CV1}}$ If $h(x_0) = 0$, then $\exists \varepsilon_0, \delta > 0$ such that

$$\gamma_1 \cdot \gamma_2 > \delta,$$

$\forall \gamma_1, \gamma_2 \in \partial^{>} h(x)$, and $x \in \{x_0\} + \varepsilon_0 \mathbb{B}$.

CQ$_{\mathbf{CV2}}$ $\exists \delta > 0$:

$$\gamma_1 \cdot \gamma_2 > \delta,$$

$\forall \gamma_1, \gamma_2 \in \partial^{>} h(\bar{x}(s))$, and $s \in \{t \in [0, 1] : h(\bar{x}(t)) = 0\}$.

Proposition 4.2.1 *Let $(\bar{x}, \bar{u})$ be a local minimizer for (CVP_2). Assume that hypotheses* **H1n$_{\mathbf{CV}}$ − H4n$_{\mathbf{CV}}$** *and constraint qualifications* **CQ$_{\mathbf{CV1}}$ − CQ$_{\mathbf{CV2}}$** *are satisfied. Then there exist $p \in W^{1,1}([0, 1] : \mathbb{R}^n)$, a measurable function γ and a non-negative Radon measure $\mu \in C^*([0, 1], \mathbb{R})$ such that*

$$\dot{p}(t) \in co\ \partial_x^L (L(\bar{x}(t), \dot{\bar{x}}\ (t)))\ \text{and}\ q(t) \in co\ \partial_u^L (L(\bar{x}(t), \dot{\bar{x}}\ (t))),$$

$$(p(0), -q(1)) \in N^L_{C_0 \times C_1}(\bar{x}(0), \bar{x}(1)),$$

$$\gamma(t) \in \partial_x^{>} h(\bar{x}(t)) \quad \mu - a.e.,$$

$$supp\ \{\mu\} \subset \{t \in [0, 1] : h(\bar{x}(t)) = 0\},$$

where

$$q(t) = \begin{cases} p(t) + \int_{[0,t)} \gamma(s)\mu(ds) & t \in [0,1) \\ p(t) + \int_{[0,1]} \gamma(s)\mu(ds) & t = 1. \end{cases}$$

Remark 4.2.2 *In the case when h is continuously differentiable, the set $\partial^{>} h(\bar{x}(s))$ is a singleton. Therefore, this $\mathbf{CQ_{CV2}}$ reduce to $h_x(\bar{x}(s)) \neq 0$, confirming the CQ in [FV94].*

This proposition generalize the result of [FV94] by allowing nonsmooth data.

4.3 Normality in CVP Applying the Normal Result of [RV99] or [CF05]

To answer the question: "does the CQ of type $\mathbf{CQ2_n}$ give new information when it is applied to CVP?", we apply the constraint qualification $\mathbf{CQ2_n}$ to (OCP_{M4}).

So, we assume $\exists \tilde{u}(t) \in \mathbb{R}^n$ such that

$$h_z(\bar{x}) \cdot f((\bar{x}, y), \tilde{u}) < -\delta,$$

for a constant $\delta > 0$.

Consequently,

$$(h_x(\bar{x}), 0) \cdot \begin{pmatrix} \tilde{u} \\ L(\bar{x}, \tilde{u}) \end{pmatrix} < -\delta.$$

Consider $\tilde{u}(t) = -h_x(\bar{x}(t))$, we have $h_x(\bar{x}) \cdot (-h_x(\bar{x})) = -\|h_x(\bar{x})\|^2$.

It follows that, for CVP, the constraint qualification $\mathbf{CQ2_n}$ reduces to

$$h_x(\bar{x}) \neq 0.$$

Comparing this CQ with the $\mathbf{CQ_{CV1}} - \mathbf{CQ_{CV2}}$, we conclude that the latter is more general; it can be applied to problems with less regularity on the data.

In summary, we can say that in the case of OCP the NCO of [RV99] and [CF05] in comparison with the NCO of [Fon00] have the advantage that they do not involve the control function explicitly, and therefore are easier to verify.

However, in the special case of CVP the $\mathbf{CQ_{CV1}} - \mathbf{CQ_{CV2}}$, (obtained from the results in [Fon00] for OCP) have the advantage that they apply to a wider class of problems.

Notes on Chapter

In [LF06], we can find a more detailed comparison between $\mathbf{CQ_{CV1}} - \mathbf{CQ_{CV2}}$ and CQ obtained by applying the normality result of [RV99].

Part of the contents of section 4.2 have been presented in [Lop03] (see also [LF03]).

Chapter 5

Normality of Optimal Control Problems via Linearization of Control Systems

The main objective of this chapter is to discuss normality of the MP for constrained problems with Lipschitz optimal trajectories. To prove normality, we use J. Yorke type linearization of control systems and show the existence of a solution to a linearized control system satisfying new state constraints. Our main result differers from similar results in the literature since we assume distinct set of hypothesis.

5.1 Introduction

In this chapter we consider the optimal control problem with implicit state constraints:

$$(OCP_1) \quad \begin{cases} \text{Minimize} & g(x(0), x(1)) \\ \text{subject to} & \dot{x}(t) = f\left(t, x(t), u(t)\right) \qquad \text{a.e.} t \in [0,1] \\ & x(0) \in C_0 \\ & x(t) \in X \\ & u(t) \in \Omega(t) \qquad \text{a.e.} t \in [0,1]. \end{cases}$$

Remark 5.1.1 *Define the signed distance by*

$$d(x) = \begin{cases} -dist(x, bdy\ (X)) & \forall x \in X, \\ dist(x, bdy\ (X)) & otherwise. \end{cases}$$

The problem (OCP_1) is equivalent to replacing the state constraint (5.1) by the inequality state constraint

$$d(x(t)) \leq 0 \text{ for all } t \in [0,1].$$

Assume that Basic hypotheses $\mathbf{H1_b}$-$\mathbf{H4_b}$ (see section 2.3.2) and the following hypothesis is satisfied:

$\mathbf{H5_n}$ $Int\, C_X(\bar{x}(t)) \neq \emptyset$ for each $t \in [0,1]$. (This is a sufficient conditions for Vinter's CQ, see [Vin00].)

Here $C_C(x)$ denotes the Clarke's tangent cone,

$$C_C(x) = \{v \in \mathbb{R}^n | \lim_{h\to 0^+ x' \to_C x} \frac{dist(x' + hv, C)}{h} = 0\}.$$

Then the Maximum Principle is the following:

Theorem 5.1.2 ***(The Maximum Principle for Implicit State Constraints)****[Vin00] (Section 9.3) There exists an absolutely continuous function $p : [0,1] \to \mathbb{R}^n$, $\eta \in C^*([0,1] : \mathbb{R}^n)$, and $\lambda \geq 0$ such that*

$$\int_{[0,1]} \zeta(t) \cdot \eta(dt) \leq 0$$

for all $\zeta \in C([0,1] : \mathbb{R}^n)$ satisfying $\zeta(t) \in C_X(\bar{x}(t))$ η a.e.

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

$$supp\{\eta\} \subset \{t \in [0,1] : \bar{x}(t) \in bdy\ (X)\},$$

$$-\dot{p}(t) \in co\, \partial_x^L H(t, \bar{x}(t), q(t), \bar{u}(t)) \text{ a.e.,} \tag{5.1}$$

$$(p(0), -q(1)) \in \lambda \partial^L g(\bar{x}(0), \bar{x}(1)) + N^L_{C_0}(\bar{x}(0)) \times \{0\},$$

$$H(t, \bar{x}(t), q(t), \bar{u}(t)) = \max_{u \in \Omega(t)} H(t, \bar{x}(t), q(t), u) \ \ a.e..$$

Here

$$q(t) = \begin{cases} p(t) + \int_{[0,t)} \eta(ds) & t \in [0,1) \\ p(t) + \int_{[0,1]} \eta(ds) & t = 1. \end{cases}$$

In this chapter we assume a CQ to ensure the normality. This CQ is typically of the kind: there exists a control u and $\epsilon > 0$ such that

$$d_x(x) \cdot f(t, x, u) \leq -\rho \text{ for all } x \in \bar{x}(t) + \epsilon\mathbb{B}, t \in [0,1], \bar{x}(t) \in \text{bdy }(X) \qquad (5.2)$$

for some positive ρ.

Results ensuring normality using CQ of the type mention above can be found in [RV99], [CF05] and [BF07].

In this chapter we improve the result in [RV99], since we assume that the function defining by dynamics is merely measurable with respect to time.

In [RV99], the proof of the main result on normality is based on neighbouring feasible trajectories theorem. In this chapter, and also in [CF05] and [BF07], the proof is based in ensuring the existence of a solution to the problem

$$\begin{cases} \dot{w} = \gamma(t, w) + \varphi(t), \\ \varphi(t) \in T_{\overline{co}(f(t,\bar{x}(t),\Omega(t)))}(\dot{\bar{x}}\ (t)) \end{cases}$$

satisfying

$$w(t) \in Int(T_X(\bar{x}(t))) \ \forall t \in [0,1].$$

Here $T_C(x)$ denotes Contingent Cone,

$$T_C(x) = \{v \in \mathbb{R}^n | \liminf_{h \to 0^+} \frac{dist(x + hv, C)}{h} = 0\}.$$

In the main result of this chapter we considering that $\gamma(t, \cdot)$ is merely $k(\cdot)$-

Lipschitz function, instead of $k \in L^\infty$ as it was proved in [CF05]. The result on normality of [BF07] allows the set X be nonsmooth, however the continuity of $u \to f(t,x,u)$ is assumed.

5.2 Normality Result via Linearization

Assume also the following hypotheses:

$\mathbf{H6_n}$ Let $X \subset \mathbb{R}^n$ be closed and such that for some $\eta > 0$ the signed distance $d(\cdot)$ is of class $C^{1,1}_{loc}$ on bdy $(X) + \eta\mathbb{B}$.

$\mathbf{H7_n}$ $Int\,(T_X(\bar{x}(0))) \cap Int\,(T_{C_0}(\bar{x}(0))) \neq \emptyset$.

$\mathbf{H8_n}$ There exist $t_0,\ t_1,\ ...,\ t_m$ such that $0 = t_0 < t_1 < t_2... < t_m = 1$ and for all $i \in \{0, ..., m-1\}$ either $\bar{x}(t_i, t_{i+1}) \subset Int\,(X)$ and $\bar{x}(t_i)$, $\bar{x}(t_{i+1}) \in$ bdy (X) or $\bar{x}([t_i, t_{i+1}]) \subset$ bdy (X).

$\mathbf{CQ_{n_{VL}}}$ Assume that for all $R > 0$, there exists $r > 0$ and $\overline{\rho} > 0$ such that for a.e. $t \in [0,1]$ the following holds true

$$\begin{aligned} &\forall x \in (\text{bdy }(X) + \overline{\rho}\mathbb{B}) \cap R\mathbb{B} \\ &\qquad \inf\{d_x(x) \cdot f(t,x,u) : u \in \Omega(t), \quad \|f(t,x,u)\| \leq r\} \leq -\overline{\rho}. \end{aligned}$$

We are now in position to state the main result of this chapter.

Theorem 5.2.1 *Let $(\bar{x}, \bar{u})$ be a local minimizer for the problem (OCP_1), where $\bar{x}$ is Lipschitz continuous. Assume that the hypotheses* $\mathbf{H1_b}$-$\mathbf{H4_b}$ *and* $\mathbf{H5_n} - \mathbf{H8_n}$ *and the constraint qualification* $\mathbf{CQ_{n_{VL}}}$ *are satisfied. Then the MP for implicit state constraints theorem 5.1.2 holds true with $\lambda = 1$.*

Remark 5.2.2 *By the regularity hypotheses on the set X, we conclude that $T_X(\bar{x}(t)) = C_X(\bar{x}(t))$, see [Cla83].*

The proof of the main theorem follows directly from the next three Lemmas.

Lemma 5.2.3 *Assume that the assumptions of the theorem and CQ_{nVL} holds. Then there exist $0 < \delta < \eta$, $\rho > 0$ and $v(t) \in f(t, \bar{x}(t), \Omega(t)) \cap r\mathbb{B}$ such that $v(\cdot)$ is measurable and*

$$d_x(\bar{x}(t)) \cdot v(t) \leq -\rho,$$

whenever dist$(\bar{x}(t), bdy\ (X)) \leq \delta$.

Lemma 5.2.4 *Assume that there exist a function $\gamma : [0,1] \times \mathbb{R}^n \to \mathbb{R}^n$ measurable in the first variable and for some $k \in L^1$ and a.e. $t \in [0,1]$, $\gamma(t, \cdot)$ is $k(t)$-Lipschitz. The hypotheses $\mathbf{H6_n}$-$\mathbf{H8_n}$ and the constraint qualification $\mathbf{CQ_{N_{VL}}}$ are satisfied. Additionally assume that $\bar{x} : [0,1] \to X$ is a Lipschitz function, the function whose existence is assumed in Lemma (5.2.3) is essentially bounded and for that $v \in L^\infty(0,1)$, be such that for some $\rho > 0$ and a.e. $t \in [0,1]$ with $\bar{x}(t) \in bdy\ (X) + \eta\mathbb{B}$*

$$d_x(\bar{x}(t)) \cdot v(t) \leq -\rho.$$

Then for every $w_0 \in Int\,(T_X(\bar{x}(0)) \cap Int\,(T_{C_0}(\bar{x}(0)))$ there exists an absolutely continuous solution w to

$$\dot{w} = \gamma(t, w) + \mu(t)(v(t) - \dot{\bar{x}}\ (t)), w(0) = w_0 \tag{5.3}$$

such that

$$w(t) \in Int\, T_X(\bar{x}(t)), \tag{5.4}$$

for all $t \in [0,1]$, for some $\mu \in L^1$ such that $\mu(t) \geq 0$.

Remark 5.2.5 *Define $\Gamma = \{t \in [0,1] : \bar{x}(t) \in bdy\ (X) + \eta\mathbb{B}\}$. By the measurable selection theorem (see for instance 10.2.58 in appendix), there exists a measurable selection $v(t) \in f(t, \bar{x}(t), \Omega(t))$ such that $d_x(\bar{x}(t)) \cdot v(t) \leq -\rho$ for almost all $t \in \Gamma$. We extend v on $[0,1]$ by setting $v(t) = \dot{\bar{x}}\ (t)$ for all $t \notin \Gamma$. Then $\mu(t)(v(t) - \dot{\bar{x}}\ (t)) \in T_{\overline{co}(f(t,\bar{x}(t),\Omega(t)))}(\dot{\bar{x}}\ (t))$.*

Lemma 5.2.6 *If there exists an absolutely continuous solution w to the problem*

$$\begin{cases} \dot{w}(t) &=& A(t)w(t) + \varphi(t), \\ \varphi(t) &\in& T_{\overline{co}(f(t,\bar{x}(t),\Omega(t)))}(\dot{\bar{x}}\,(t)) \\ w(t) &\in& Int\, T_X(\bar{x}(t)),\ \forall t \in [0,1] \\ w(0) &\in& Int\,(T_X(\bar{x}(0))) \cap Int\,(T_{C_0}(\bar{x}(0))) \end{cases}$$

for any $A(t) \in co\ \partial_x^L f(t,\cdot,\bar{u}(t))$, then $\lambda = 1$.

Remark 5.2.7 *If $A(t) \in co\ \partial_x^L f(t,\cdot,\bar{u}(t))$ and $A(\cdot)$ is measurable, then $A(\cdot) \in L^1$. (see Proposition 10.2.82 in appendix).*

5.3 Proof of Lemmas

5.3.1 Proof of Lemma 5.2.3

We start by defining $T = \{t \in [0,1] : \bar{x}(t) \in \text{bdy }(X)\}$. This set is compact. Let $R = \|\bar{x}\|_\infty$. Take $\bar{\rho} = 2\rho$ in $\mathbf{CQ_{nVL}}$, there exists $r > 0$, such that for all $t \in T$, $\forall x \in \bar{x}(t) + 2\rho\mathbb{B}$ we have $\inf\{d_x(x) \cdot f(t,x,u) : u \in \Omega(t), \|f(t,x,u)\| \le r\} \le -2\rho$. Since $\bar{x}(\cdot)$ is continuous, we deduce that for some $\delta > 0$ and a.e. $t \in [0,1]$ satisfying $dist(\bar{x}(t), \text{bdy }(X)) \le \delta$, we have $\inf\{d_x(\bar{x}(t)) \cdot f(t,\bar{x}(t),u) : u \in \Omega(t), \|f(t,\bar{x},u)\| \le r\} \le -\frac{3}{4}\rho$. The measurable selection then yields the result (Proposition 10.2.58 in appendix).

5.3.2 Proof of Lemmas 5.2.4

Note: For all $t \in [0,1]$ such that $\bar{x}(t) \in \text{bdy }(X)$, we have

$$T_X(\bar{x}(t)) = \{w \in \mathbb{R}^n : d_x(\bar{x}(t)) \cdot w \le 0\} \tag{5.5}$$

and

$$Int\, T_X(\bar{x}(t)) = \{w \in \mathbb{R}^n : d_x(\bar{x}(t)) \cdot w < 0\}.$$

Remark:

i) In the proof provided below we construct $\mu \in L^1(0,1)$ and w_0 such that the solution w to

$$\begin{cases} \dot{w} = \gamma(t,w) + \mu(t)(v(t) - \dot{\bar{x}}(t)) \\ \\ w(0) = w_0 \end{cases}$$

with $t \in [0,1]$, satisfies (5.4).

ii) Since $d(\cdot)$ is of class $C^{1,1}_{loc}$ on bdy $(X) + \eta\mathbb{B}$, we know that $d_x(\bar{x}(\cdot))$ is Lipschitz on on bdy $(X) + \eta\mathbb{B}$. Let L denote a Lipschitz constant of $d_x(\bar{x}(\cdot))$. We denote also $\xi(\cdot) = \dfrac{d}{dt} d_x(\bar{x}(\cdot))$.

iii) As v is essentially bounded and $\bar{x}(\cdot)$ is a Lipschitz function, then for some $P > 0$,

$$\|v(t) - \dot{\bar{x}}(t)\| \leq P \text{ a.e. in } [0,1]. \tag{5.6}$$

Note that, if $\bar{x}(\cdot)$ is differentiable at t, then

$$\dot{\bar{x}}(t) = \lim_{h \to 0^+} \frac{\bar{x}(t+h) - \bar{x}(t)}{h} \in T_X(\bar{x}(t))$$

and

$$-\dot{\bar{x}}(t) = \lim_{h \to 0^+} \frac{\bar{x}(t-h) - \bar{x}(t)}{h} \in T_X(\bar{x}(t)).$$

Thus, if $\bar{x}(t) \in$ bdy (X), from (5.5) we obtain

$$\left. \begin{array}{l} d_x(\bar{x}(t)) \cdot \dot{\bar{x}}(t) \leq 0 \\ \\ d_x(\bar{x}(t)) \cdot (-\dot{\bar{x}}(t)) \leq 0 \end{array} \right\} \Rightarrow d_x(\bar{x}(t)) \cdot \dot{\bar{x}}(t) = 0, \ \forall t \in (0,1).$$

Define $\Im := \{t \in [0,1] : t$ is a Lebesgue point of $\dfrac{2(L + k(\cdot))}{\rho} d_x(\bar{x}(\cdot)) \cdot \dot{\bar{x}}(\cdot)\}$. Since $k \in L^1$ and $\dot{\bar{x}} \in L^\infty$, $\Im$ is of full measurable in [0,1].

CLAIM 1: Let $0 \leq t_0 < t_1 \leq 1$ be such that $\bar{x}(t_1) \in$ bdy (X) and $\bar{x}((t_0, t_1)) \in$ $Int\,(X)$. Then for all $\varepsilon > 0$ such that $t_0 \leq t_1 - \varepsilon$ there exists $t \in]t_1 - \varepsilon, t_1[\cap \Im$ such that

$$d(\bar{x}(t)) \cdot \dot{\bar{x}}\,(t) > 0. \tag{5.7}$$

Indeed assume that $\exists \varepsilon > 0$ such that for all $t \in]t_1 - \varepsilon, t_1[\cap \Im$ and $\bar{x}(t_1) \in$ bdy (X),

$$d(\bar{x}(t)) \cdot \dot{\bar{x}}\,(t) \leq 0.$$

Then $\int_{t_1-\varepsilon}^{t_1} d_x(\bar{x}(t)) \cdot \dot{\bar{x}}\,(t) dt = d(\bar{x}(t_1)) - d(\bar{x}(t_1 - \varepsilon)) \leq 0$. On other hand $d(\bar{x}(t_1)) = 0$ and therefore $d(\bar{x}(t_1)) - d(\bar{x}(t_1 - \varepsilon)) > 0$. The obtained contradiction proves our claim.

Step 1: We start the proof of our Lemma, by considering the following case:

$$\bar{x}(t) \in Int\,(X), \forall t \in]0, 1].$$

As $w_0 \in Int\, T_X(\bar{x}(0))$, then any solution to $\dot{w}(t) = \gamma(t, w)$ satisfies $w(t) \in Int\, T_X(\bar{x}(t))$.

Step 2: Next suppose that $\bar{x}([0, 1]) \subset$ bdy (X) and consider the solution w to

$$\begin{cases} \dot{w} = A(t)w(t) + \dfrac{2(L + k(t))}{\rho} \|w(t)\| (v(t) - \dot{\bar{x}}\,(t)) \\ \\ w(0) = w_0 \end{cases}, \; t \in [0, 1].$$

We wish to check that $d_x(\bar{x}(t)) \cdot w(t) < 0$ for all $t \in [0,1]$. Indeed

$$d_x(\bar{x}(t)) \cdot w(t) =$$

$$= d_x(\bar{x}(0)) \cdot w(0) + \int_0^t \xi(s) \cdot w(s) + d_x(\bar{x}(s)) \cdot \dot{w}(s) ds$$

$$\leq d_x(\bar{x}(0)) \cdot w(0)$$
$$+ \int_0^t L\|w(s)\| + d_x(\bar{x}(s)) \cdot (\gamma(s, w(s)) + \frac{2(L+k(s))}{\rho}\|w(s)\|(v(s) - \dot{\bar{x}}(s)))ds.$$

So using the fact that $\|d_x(\bar{x}(s))\| = 1$ and $\|\gamma(s, w(s))\| \leq k(s)\|w(s)\|$, we have

$$d_x(\bar{x}(t)) \cdot w(t) \quad \leq d_x(\bar{x}(0)) \cdot w(0) + \int_0^t -(L+k(s))\|w(s)\| ds$$

$$-\int_0^t \frac{2(L+k(s))}{\rho}\|w(s)\| d_x(\bar{x}(s)) \cdot \dot{\bar{x}}(s) ds.$$

Since $d_x(\bar{x}(s)) \cdot \dot{\bar{x}}(s) = 0$ for a.e. s and $d(\bar{x}(0)) \cdot w(0) < 0$, we get

$$d_x(\bar{x}(t)) \cdot w(t) < -\int_0^t (L+k(s))\|w(s)\| ds, \ \forall t \in [0,1],$$

and the statement of our lemma follows.

Step 3: It remains to consider the case when $\bar{x}([0,1]) \cap \text{bdy } (X) \neq \emptyset$ and $\bar{x}([0,1]) \cap \text{bdy } (X) \neq \bar{x}([0,1])$.
Set $M(t) = \exp\left(\int_0^t k(s) + \frac{2(L+k(s))}{\rho} P ds\right)$ and $C_0 = \|w(0)\|(1 + \frac{P}{\rho}) + 2P$. ($P$ as choose on 5.6)

Fix $w_0 \in Int\, T_{C_0}(\bar{x}(0))$, such that $d_x(\bar{x}(0)) \cdot w(0) \leq -\rho$.

CLAIM 2: We claim that there exist $t_1 > 0$ and a solution $w(\cdot)$ to (5.3) on $[0, t_1]$ such that for all $\tau \in [0, t_1]$, $w(\tau) \in Int\, T_X(\bar{x}(\tau))$, and either $\bar{x}(t_1) \in \text{bdy } (X)$ and $d_x(\bar{x}(t_1)) \cdot w(t_1) \leq -\rho$ or $t_1 = 1$ and $\bar{x}(1) \in Int\,(X)$.

Now, we can have two possible situations:

Case 1 - $\exists t_1 > 0$ such that $\bar{x}([0, t_1]) \subset \text{bdy } (X)$;

Case 2 - $\exists t_1 > 0$ such that $\bar{x}((0, t_1)) \subset Int\,(X)$ and either $t_1 = 1$ and $\bar{x}(t_1) \in Int\,(X)$ or $\bar{x}(t_1) \in \text{bdy } (X)$.

We start by **Case 1**. Then exists an element $t_1 > 0$ such that $t_1 = \max\{t \in (0, 1] : \bar{x}([0, t]) \subset \text{bdy } (X)\}$. We consider the solution w to

$$\begin{cases} \dot{w} = \gamma(t, w) + \dfrac{2(L + k(t))}{\rho} \|w(t)\|(v(t) - \dot{\bar{x}}\,(t)) \\ \\ w(0) = w_0 \end{cases} , \; t \in [0, t_1].$$

Next we prove that for all $t \in [0, t_1]$, $d_x(\bar{x}(t)) \cdot w(t) \leq -\rho$. Therefore

$$\begin{aligned} d_x(\bar{x}(t)) \cdot w(t) &= d_x(\bar{x}(0)) \cdot w(0) + \int_0^t \xi(s) \cdot w(s) + d_x(\bar{x}(s)) \cdot \dot{w}(s) ds \\ &\leq d_x(\bar{x}(0)) \cdot w(0) + \int_0^t \Big(L + k(s) - 2(L + k(s))\Big) \|w(s)\| ds \\ &\quad - \int_0^t \frac{2(L + k(s))}{\rho} \|w(s)\| d_x(\bar{x}(s)) \cdot \dot{\bar{x}}\,(s) ds. \end{aligned}$$

As $d_x(\bar{x}(s)) \cdot \dot{\bar{x}}\,(s) = 0$, for a.e. $s \in (0, t_1)$ and $d_x(\bar{x}(0)) \cdot w(0) \leq -\rho$, we conclude that

$$d_x(\bar{x}(t)) \cdot w(t) \leq -\rho - \int_0^t (L + k(s))\|w(s)\|ds \leq -\rho.$$

So CLAIM 2 is proved in **Case 1**.

In **Case 2**, there exist $t_1 > 0$ such that $t_1 = \sup\{t \in (0, 1] : \bar{x}((0, t)) \subset Int\,(X)\}$.

If $t_1 = 1$ and $\bar{x}(t_1) \in Int\,(X)$, we consider the solution w to $\dot{w} = \gamma(t, w)$ and $w(0) = w_0$, $\forall t \in [0, 1]$. So $w(t) \in Int\,T_X(\bar{x}(t))$ and CLAIM 2 follows.

If this is not the case, by CLAIM 1 there exists a sequence $\{s_i\}$ with $s_i \in \Im$, such that $s_i \to t_1^-$ and $d_x(\bar{x}(s_i))\cdot \dot{\bar{x}}\,(s_i) > 0$. Let $\varepsilon = \frac{\rho}{2M(s_i)C_0}$. Consider h_i such that

$$d(\bar{x}(s_i)) - d(\bar{x}(s_i - h_i)) > 0,\, 0 < h_i < 1. \tag{5.8}$$

Without any loss of generality and using the fact that s_i is a Lebesgue point, we may assume that h_i satisfy

$$\int_{s_i-h_i}^{s_i} \left| \frac{2(L + k(s))}{\rho} d_x(\bar{x}(s))\cdot \dot{\bar{x}}\,(s) - \frac{2(L + k(s_i))}{\rho} d_x(\bar{x}(s_i))\cdot \dot{\bar{x}}\,(s_i) \right| ds \leq \varepsilon h_i. \tag{5.9}$$

Let us define $w(\cdot)$ on the time interval $[0, s_i - h_i]$ by the solution w to $\dot{w}(t) = \gamma(t, w)$, $w(0) = w_0$, then $w(t) \in Int\,T_X(\bar{x}(t))$.

Now, we extend w on time interval $]s_i - h_i, s_i]$ by the solution to

$$\dot{w} = \gamma(t, w) + \frac{2(L + k(t))}{\rho}\|w(t)\|(v(t) - \dot{\bar{x}}\,(t)) + \left(\frac{\|w(s_i - h_i)\|}{\rho h_i} + \frac{2}{h_i}\right)(v(t) - \dot{\bar{x}}\,(t)).$$

Then $w(t) \in Int\,T_X(\bar{x}(t)) = \mathbb{R}^n$ for all $t \in [s_i - h_i, s_i]$.

As $\|v(t)-\dot{\bar{x}}(t)\| \le P$ for a.e. $t \in [0,1]$, we have for all $t \in [s_i - h_i, s_i]$

$$\begin{aligned}
\|w(t)\| &\le \|w(s_i-h_i)\| + \int_{s_i-h_i}^{t} \|\dot{w}(s)\| ds \\
&\le \|w(s_i-h_i)\| + \int_{s_i-h_i}^{t} \left(k(s) + \frac{2(L+k(s))}{\rho} P\right) \|w(s)\| \\
&\quad + \left(\tfrac{\|w(s_i-h_i)\|}{\rho h_i} + \tfrac{2}{h_i}\right) P ds \\
&\le \|w(s_i-h_i)\| + \int_{s_i-h_i}^{t} \left(k(s) + \frac{2(L+k(s))}{\rho} P\right) \|w(s)\| ds \\
&\quad + \left(\frac{\|w(s_i-h_i)\|}{\rho h_i} + \frac{2}{h_i}\right) P(t-(s_i-h_i)).
\end{aligned}$$

Furthermore since $\dfrac{t-(s_i-h_i)}{h_i} \le 1$, we conclude that

$$\begin{aligned}
\|w(t)\| &\le \|w(s_i-h_i)\|(1+\tfrac{P}{\rho}) + 2P + \int_{s_i-h_i}^{t} \left(k(s) + \frac{2(L+k(s))}{\rho} P\right) \|w(s)\| ds \\
&\le \left(\|w(0)\| + \int_0^{s_i-h_i} k(s)\|w(s)\| ds\right)(1+\tfrac{P}{\rho}) + 2P + \\
&\quad + \int_{s_i-h_i}^{t} \left(k(s) + \frac{2(L+k(s))}{\rho} P\right) \|w(s)\| ds \\
&\le \left(\|w(0)\|(1+\tfrac{P}{\rho}) + 2P\right) + \int_0^t \left(k(s) + \frac{2(L+k(s))}{\rho} P\right) \|w(s)\| ds.
\end{aligned}$$

By Gronwall's Lemma, we have

$$\|w(t)\| \le M(t) C_0. \tag{5.10}$$

We next show that $d(\bar{x}(s_i)) \cdot w(s_i) \le -\frac{3}{2}\rho$.

Indeed

$$
\begin{aligned}
& d(\bar{x}(s_i)) \cdot w(s_i) \leq \\
& \quad = \|w(s_i - h_i)\| + \int_{s_i - h_i}^{s_i} \xi(s) \cdot w(s) + d_x(\bar{x}(s)) \cdot \dot{w}(s) ds \\
& \quad \leq \|w(s_i - h_i)\| + \int_{s_i - h_i}^{s_i} (L + k(s) - 2(L + k(s))) \|w(s)\| ds \\
& \quad - \int_{s_i - h_i}^{s_i} \frac{2(L + k(s))}{\rho} d_x(\bar{x}(s)) \cdot \dot{\bar{x}}\ (s) \|w(s)\| ds - \|w(s_i - h_i)\| - 2\rho \\
& \quad - \left(\tfrac{\|w(s_i - h_i)\|}{\rho h_i} + \tfrac{2}{h_i} \right) \int_{s_i - h_i}^{s_i} d_x(\bar{x}(s)) \cdot \dot{\bar{x}}\ (s) ds \\
& \quad \leq \int_{s_i - h_i}^{s_i} - (L + k(s)) \|w(s)\| ds \\
& \quad - \int_{s_i - h_i}^{s_i} \frac{2(L + k(s))}{\rho} d_x(\bar{x}(s)) \cdot \dot{\bar{x}}\ (s) \|w(s)\| ds - 2\rho \\
& \quad - \left(\tfrac{\|w(s_i - h_i)\|}{\rho h_i} + \tfrac{2}{h_i} \right) (d(\bar{x}(s_i)) - d(\bar{x}(s_i - h_i))).
\end{aligned}
\tag{5.11}
$$

The above together with (5.8), imply that

$$
\begin{aligned}
d_x(\bar{x}(s_i)) \cdot w(s_i) \quad & \leq \int_{s_i - h_i}^{s_i} - (L + k(s)) \|w(s)\| ds \\
& - \int_{s_i - h_i}^{s_i} \frac{2(L + k(s))}{\rho} d_x(\bar{x}(s)) \cdot \dot{\bar{x}}\ (s) \|w(s)\| ds - 2\rho.
\end{aligned}
\tag{5.12}
$$

On the other hand by (5.9) and (5.10), we know that

$$-\int_{s_i-h_i}^{s_i} \frac{2(L+k(s))}{\rho} d_x(\bar{x}(s))\cdot \dot{\bar{x}}(s)\|w(s)\|ds =$$

$$-\int_{s_i-h_i}^{s_i} \left(\frac{2(L+k(s))}{\rho} d_x(\bar{x}(s))\cdot \dot{\bar{x}}(s) - \frac{2(L+k(s_i))}{\rho} d_x(\bar{x}(s_i))\cdot \dot{\bar{x}}(s_i)\right)\|w(s)\|ds$$

$$-\int_{s_i-h_i}^{s_i} \frac{2(L+k(s_i))}{\rho} d_x(\bar{x}(s_i))\cdot \dot{\bar{x}}(s_i)\|w(s)\|ds$$

$$\leq \int_{s_i-h_i}^{s_i} \left|\frac{2(L+k(s))}{\rho} d_x(\bar{x}(s))\cdot \dot{\bar{x}}(s) - \frac{2(L+k(s_i))}{\rho} d_x(\bar{x}(s_i))\cdot \dot{\bar{x}}(s_i)\right|\|w(s)\|ds$$

$$\leq \varepsilon M(s_i) C_0 h_i.$$

From (5.12) and the choice of ε, we deduce that

$$\begin{aligned} d(\bar{x}(s_i))\cdot w(s_i) &\leq -\int_{s_i-h_i}^{s_i} (L+k(s))\|w(s)\|ds + \frac{\rho}{2}h_i - 2\rho \\ &= -\int_{s_i-h_i}^{s_i} (L+k(s))\|w(s)\|ds - \frac{3\rho}{2} \\ &\leq -\tfrac{3\rho}{2}. \end{aligned}$$

Again we extend w on time interval $[s_i, t_1]$ by the solution to $\dot{w}(t) = \gamma(t, w)$ then $w(t) \in Int\, T_X(\bar{x}(t)) = \mathbb{R}^n$, for all $t \in [s_i, t_1)$. It remains to check that

$d_x(\bar{x}(t_1)) \cdot w(t_1) \leq -\rho$. Observe that

$$\begin{aligned} d_x(\bar{x}(t_1)) \cdot w(t_1) &= d_x(\bar{x}(t_1)) \cdot (w(t_1) - w(s_i) + w(s_i)) \\ &= d_x(\bar{x}(t_1)) \cdot (w(t_1) - w(s_i)) + (d_x(\bar{x}(t_1)) - d_x(\bar{x}(s_i)) \\ &\quad + d_x(\bar{x}(s_i))) \cdot w(s_i). \qquad (5.13) \\ &\leq \|w(t_1) - w(s_i)\| + L(t_1 - s_i)\|w(s_i)\| \\ &+ d_x(\bar{x}(s_i)) \cdot w(s_i). \end{aligned}$$

Furthermore, as $\dot{w}(t) = \gamma(t, w)$ for all $t \in [s_i, t_1]$,

$$\begin{aligned} \|w(t) - w(s_i)\| &\leq \int_{s_i}^{t} k(\tau)\|w(\tau)\|d\tau \\ &\leq \int_{s_i}^{t} k(\tau)\|w(\tau) - w(s_i)\|d\tau + \int_{s_i}^{t} k(\tau)\|w(s_i)\|d\tau. \end{aligned}$$

By Gronwall's Lemma, we have

$$\|w(t_1) - w(s_i)\| \leq \exp\left(\int_{s_i}^{t_1} k(\tau)d\tau\right) \int_{s_i}^{t_1} k(\tau)d\tau\|w(s_i)\|.$$

Let $\varepsilon_1^i > 0$ be such that

$$\begin{aligned} &\exp(\varepsilon_1^i)\varepsilon_1^i M(t_1)C_0 \leq \tfrac{\rho}{4} \\ \text{and} \qquad & \qquad\qquad\qquad\qquad (5.14) \\ &\varepsilon_2 M(t_1)C_0 L = \tfrac{\rho}{4}. \end{aligned}$$

Since s_i converges to t_1, there exists i_0 such that for all $i \geq i_0$, $|\int_{s_i}^{t_1} k(\tau)d\tau)| \leq \varepsilon_1$. So

$$\|w(t_1) - w(s_i)\| \leq \exp(\varepsilon_1^i)\varepsilon_1^i\|w(s_i)\|. \qquad (5.15)$$

By the fact that $s_i \to t_1$ and all i large enough we have $|t_1 - s_i| \leq \varepsilon_2^i$ and from

(5.13) and (5.15), we conclude

$$d_x(\bar{x}(t_1)) \cdot w(t_1) \leq \exp(\varepsilon_1^i)\varepsilon_1^i \|w(s_i)\| + \varepsilon_2^i L\|w(s_i)\| + d_x(\bar{x}(s_i)) \cdot w(s_i)$$

Since $\|w(s_i)\| \leq M(s_i)C_0$,

$$d_x(\bar{x}(t_1)) \cdot w(t_1) \leq \exp(\varepsilon_1^i)\varepsilon_1^i M(s_i)C_0 + \varepsilon_2^i LM(s_i)C_0 + d_x(\bar{x}(s_i)) \cdot w(s_i).$$

From (5.14) and by the fact of $M(s_i) < M(t_1)$, we deduce that

$$d_x(\bar{x}(t_1)) \cdot w(t_1) \leq \frac{\rho}{2} + d_x(\bar{x}(s_i)) \cdot w(s_i) \leq -\rho.$$

Since $\exists 0 = t_0 < t_1 < ... < t_n = 1$ such that $\bar{x}|_{(t_i,t_{i+1})}$ is either on the boundary or the interior and $\bar{x}(t_i) \in$ bdy (X) for all $i \neq 0, n$, we extend w on $[0,1]$ using the same reasoning as in CLAIM 1 and CLAIM 2.

5.3.3 Proof of Lemma 5.2.6

Define

$$\nu(t) = \begin{cases} \displaystyle\int_{[0,t)} \eta(dr) \text{ for all } t \in [0,1) \\ \displaystyle\int_{[0,1]} \eta(dr) \text{ for } t = 1 \end{cases}$$

where

$$\text{supp}\{\eta\} \subset \{t \in [0,1] : \bar{x}(t) \in \text{bdy } (X)\},$$

and define

$$\begin{aligned} \mathcal{C} &= \{w \in C([0,1]) : w(t) \in Int\, T_X(\bar{x}(t)), \forall t \in [0,1]\}, \\ \mathcal{C}_0 &= \{w \in C([0,1]) : w(0) \in Int\, T_{C_0}(\bar{x}(0))\}, \\ S &= \{w \in W^{1,1}([0,1] : \mathbb{R}^n) : \dot{w}(t) = \gamma(t, w(t)) + \varphi(t), \\ &\quad \varphi(t) \in T_{\overline{co}(f(t,\bar{x}(t),\Omega(t)))}(\dot{x}(t)) \text{ a.e. in } [0,1]\}. \end{aligned}$$

It is well known that ν has bounded variation and so it has right and left limits $\nu(t^+)$ and $\nu(t^-)$ respectively at every $t \in (0,1)$. Furthermore $\nu(0^+)$ and $\nu(1^-)$ do exist (see Lemma 10.2.30 in appendix).

Take $\varphi \in L^1(0,1)$, such that $\varphi(t) \in T_{\overline{co}(f(t,\bar{x}(t),\Omega(t)))}(\dot{\bar{x}}\ (t))$ and the solution to $\dot{w} = \gamma(t, w(t)) + \varphi(t)$ satisfies $w(0) \in Int\, T_{C_0}(\bar{x}(0))$ and $w(t) \in Int\, T_X(\bar{x}(t))$. By the MP $q(t) \cdot \varphi(t) \leq 0$.

We shall need the following result.

Proposition 5.3.1 *Let ν be as defined above. Then $\nu(0^+) \in N_X(\bar{x}(0))$.*

Proof: Fix $t_2 > 0$ and $\delta > 0$ so that $t_2 - \delta > 0$ and ν is continuous at $t_2 - \delta$ and $w_0 \in T_X(\bar{x}(t))$ for all $t \in [0, t_2 - \delta]$. We recall that ν is of bounded variation and it has at most countable number of points of discontinuity.

Let $\bar{w} \in \mathcal{C}$. Fix $\varepsilon > 0$

$$w_\delta(s) = \begin{cases} w_0 & s \in [0, t_2 - \delta] \\ \dfrac{t_2 - s}{\delta} w_0 + \dfrac{-t_2 + \delta + s}{\delta} \varepsilon \bar{w}(s) & s \in (t_2 - \delta, t_2) \\ \varepsilon \bar{w}(s) & s \in [t_2, 1]. \end{cases}$$

Then $w_\delta(s) \in T_X(\bar{x}(t))$ and so

$$\int_0^1 w_\delta(s) d\nu(s) \leq 0$$

$$\Leftrightarrow \quad \int_0^{t_2-\delta} w_0 d\nu(s) + \int_{(t_2-\delta,t_2)} \frac{(t_2 - s)w_0}{\delta} + \frac{-t_2 + \delta + s}{\delta} \varepsilon \bar{w}(s) d\nu(s) + \int_{t_2}^1 \varepsilon \bar{w}(s) d\nu(s) \leq 0$$

$$\Leftrightarrow \quad w_0 \cdot \nu(t_2 - \delta) + \frac{t_2 w_0}{\delta} \cdot \nu(t_2) - \frac{t_2 w_0}{\delta} \cdot \nu(t_2 - \delta^+) - \int_{(t_2-\delta,t_2)} \frac{s w_0}{\delta} d\nu(s) +$$

$$+ \int_{(t_2-\delta,t_2)} \frac{-t_2 + \delta + s}{\delta} \varepsilon \bar{w}(s) d\nu(s) + \int_{t_2}^1 \varepsilon \bar{w}(s) d\nu(s) \leq 0.$$

Integrating by parts we have

$$\int_{(t_2-\delta,t_2)} s d\nu(s) = t_2\nu(t_2) - (t_2-\delta)\nu(t_2-\delta^+) - \int_{(t_2-\delta,t_2)} \nu(s)ds.$$

So

$$w_0\cdot\nu(t_2-\delta) + \tfrac{t_2 w_0}{\delta}\cdot\nu(t_2) - \tfrac{t_2 w_0}{\delta}\cdot\nu(t_2-\delta^+) - \tfrac{t_2 w_0}{\delta}\cdot\nu(t_2) +$$

$$+\tfrac{(t_2-\delta)w_0}{\delta}\cdot\nu(t_2-\delta^+) + \tfrac{w_0}{\delta}\cdot\textstyle\int_{(t_2-\delta,t_2)}\nu(s)ds + \phi(\varepsilon) \le 0$$

$$\Leftrightarrow\quad w_0\cdot\nu(t_2-\delta) - \tfrac{t_2 w_0}{\delta}\cdot\nu(t_2-\delta^+) + \tfrac{t_2 w_0}{\delta}\cdot\nu(t_2-\delta^+) -$$

$$-w_0\cdot\nu(t_2-\delta^+) + \frac{w_0}{\delta}\cdot\int_{(t_2-\delta,t_2)}\nu(s)ds + \phi(\varepsilon)\le 0$$

$$\Leftrightarrow\quad \frac{w_0}{\delta}\cdot\int_{(t_2-\delta,t_2)}\nu(s)ds + \phi(\varepsilon)\le 0.$$

Let $\|\nu(t_2)-\nu(s)\|\le\varepsilon$ when $t_2\to s$. Then

$$\frac{w_0}{\delta}\cdot\nu(t_2)(t_2-(t_2-\delta)) - \frac{w_0}{\delta}\varepsilon(t_2-(t_2-\delta)) + \phi(\varepsilon) \le \frac{w_0}{\delta}\int_{(t_2-\delta,t_2)}\nu(s)ds + \phi(\varepsilon)\le 0.$$

Therefore,

$$w_0\nu(t_2) - \|w_0\|\varepsilon + \phi(\varepsilon)\le 0.$$

Since $\phi(\cdot)$ converge to 0 when $\varepsilon\to 0^+$, then when $t_2\to 0^+$, we have $w_0\cdot\nu(0^+)\le 0$ for all $w_0\in T_X(\bar{x}(0))$.

The proof of Proposition 5.3.1 is complete.

Now we turn back to the proof of Lemma 5.2.6.

Since, for all $t\in[0,1]$, $Int\,T_X(\bar{x}(t))\neq\emptyset$, it follows from [CF05] that $Int\,\mathcal{C}\neq\emptyset$. It is also clear that $Int\,\mathcal{C}_0\neq\emptyset$. Assume for a moment that $\lambda=0$ then $(p(0),-q(1))\in N^L_{C_0}(\bar{x}(0))\times 0$.

We have $N^L_{C_0}(\bar{x}(0)) \subseteq T_{C_0}(\bar{x}(0))^-$. It follows that for every $w \in \mathcal{C} \cap \mathcal{C}_0$,

$$\int_{[0,1]} w(s)d\nu(s) + p(0) \cdot w(0) \leq 0. \tag{5.16}$$

On other hand, for every $w \in S$,

$$\int_0^1 (\dot{p}w + q\dot{w})(s)ds = \int_0^1 -A(s)^* q(s) \cdot w(s) + q(s) \cdot \dot{w}(s)ds.$$

Since $\dot{w}(t) = A(t)w(t) + \varphi(t)$, we have

$$\int_0^1 (\dot{p}w+q\dot{w})(s)ds = \int_0^1 -q(s)\cdot A(s)w(s)+q(s)\cdot(A(s)w(s)+\varphi(s))ds = \int_0^1 q(s)\cdot\varphi(s)ds.$$

Therefore $\int_0^1 (\dot{p}w + q\dot{w})(s)ds \leq 0$. Thus,

$$\begin{aligned} &\int_0^1 \dot{p}(s)w(s) + p(s)\dot{w}(s)ds + \int_0^1 \nu(s)\dot{w}(s)ds \leq 0 \\ \Leftrightarrow \quad & p(1) \cdot w(1) - p(0) \cdot w(0) + \int_0^1 \nu(s)\dot{w}(s)ds \leq 0. \end{aligned}$$

Since

$$\int_0^t \nu(s)\dot{w}(s)ds = \nu(t^-) \cdot w(t) - \int_{[0,t)} w(s)d\nu(s),$$

we have

$$\begin{aligned} & p(1) \cdot w(1) - p(0) \cdot w(0) + \nu(1^-) \cdot w(1) - \int_0^1 w(s)d\nu(s) \leq 0 \\ \Leftrightarrow \quad & q(1) \cdot w(1) - p(0) \cdot w(0) - \int_0^1 w(s)d\nu(s) \leq 0. \end{aligned}$$

In view of the fact that $q(1) = 0$ we deduce that

$$p(0) \cdot w(0) + \int_0^1 w(s)d\nu(s) \geq 0, \tag{5.17}$$

for every $w \in S$.

Since $\overline{S} \cap Int\,(\mathcal{C} \cap \mathcal{C}_0) \neq \emptyset$, there exists $\bar{w} \in S \cap\ Int(\mathcal{C} \cap \mathcal{C}_0)$. Since $\bar{w} \in S$, by inequality (5.17) we have

$$p(0)\cdot\bar{w}(0)+\int_0^1 \bar{w}(s)d\nu(s)\geq 0.$$

On other hand, $\exists\delta>0$, such that $\bar{w}+\delta\mathbb{B}\subset\mathcal{C}\cap\mathcal{C}_0$. Consequently, by inequality (5.16), for all $w\in\bar{w}+\delta\mathbb{B}$

$$p(0)\cdot w(0)+\int_{[0,1]} w(s)d\nu(s)\leq 0.$$

Hence for all $w\in C(0,1)$,

$$p(0)\cdot w(0)+\int_{[0,1]} w(s)d\nu(s)=0.$$

This holds in particular for all absolutely continuous functions on $[0,1]$. Integrating by parts we obtain that for every $w\in W^{1,1}([0,1])$,

$$p(0)\cdot w(0)+\nu(1^-)\cdot w(1)-\int_0^1 \dot{w}(s)\nu(s)ds=0.$$

Using Dubois-Reymond Lemma we deduce that for some $c\in\mathbb{R}^n$, $\nu(s)=c$ a.e. in $[0,1]$.

So

$$\begin{aligned} & p(0)\cdot w(0)+w(1)\cdot c-c\cdot(w(1)-w(0))ds=0\\ \Leftrightarrow\quad & (p(0)+c)\cdot w(0)=0\Leftrightarrow p(0)=-c.\end{aligned}$$

So we have shown that

$$\begin{aligned} c &= \nu(0^+)\in N_X(\bar{x}(0))\\ -c &= p(0)\in N^L_{C_0}(\bar{x}(0)).\end{aligned}$$

Then $\forall w_0\in T_X(\bar{x}(0))\cap T_{C_0}(\bar{x}(0))$, we have

$$\begin{aligned} c\cdot w_0 &\leq 0\\ -c\cdot w_0 &\leq 0,\end{aligned}$$

which implies $c \cdot w_0 = 0$. By our assumptions there exists $w_0 \in \mathbb{R}^n$ and $\delta > 0$ such that $w_0 + \delta\mathbb{B} \in T_X(\bar{x}(0)) \cap T_{C_0}(\bar{x}(0))$. Then, for all $e \in \mathbb{R}^n$ with $\|e\| = 1$, we have $c \cdot (w_0 + \delta e) = 0$. This yields $c \cdot e = 0$ for all $e \in \mathbb{B}$, implying that $c = 0$.

This and adjoint equation yield $p \equiv 0$. Since ν is left continuous we proved that $\nu = 0$ on $(0,1)$. Consider any $w \in W^{1,1}(0,1)$. Then

$$\int_{[0,1]} w(s)d\nu(s) = w(1) \cdot \nu(1^-) - \int_0^1 \nu(s)\dot{w}(s)ds = 0.$$

So $\nu|_{W^{1,1}(0,1)} = 0$. Since $\nu \in C(0,1)^*$ and $W^{1,1}(0,1)$ is dense in $C(0,1)$ we get $\nu = 0$. So $(p, \eta, \lambda) = 0$. The obtained contradiction ends the proof.

Notes on Chapter

This result was developed with Prof. Hélène Frankowska, as fellow in the Control Training Site.

Chapter 6

Nondegeneracy with Integral-type Constraint Qualifications

Strengthened forms of the Maximum Principle (MP), also called nondegenerate MP, are of interest since they permit the identification of classes of problems for which the existence of nondegenerared multipliers is guaranteed.

In this chapter, we propose a nondegenerate MP under constraint qualification (CQ) of an integral type. Such MP, when compared to some of the aforementioned literature, applies to a larger class of problems.

6.1 Introduction

Consider the following OCP, in which the initial state is fixed:

$$(OCP_2) \quad \begin{cases} \text{Minimize} & g(x(1)) & \\ \text{subject to} & \dot{x}(t) = f\left(t, x(t), u(t)\right) & \text{a.e.} t \in [0,1] \\ & x(0) = x_0 & \\ & x(1) \in C & \\ & u(t) \in \Omega(t) & \text{a.e.} t \in [0,1] \\ & h\left(t, x(t)\right) \leq 0 & \text{for all } t \in [0,1]. \end{cases}$$

The strengthened form of the MP introduced in [FFV99], ensures that the non-triviality condition of the MP can be written as

$$\mu\{(0,1]\} + \|q\|_{L^\infty} + \lambda > 0,^{1}$$

when the data of the problem satisfies, besides the usual hypotheses, the following constraint qualification:

CQ$_{\mathbf{FFV99}}$: if $h(0, x_0) = 0$, then there exist positive constants ϵ, ϵ_1, δ, and a control $\tilde{u} \in \Omega(t)$ such that for a.e.$t \in [0, \epsilon)$

$$\|f(t, x_0, \bar{u}(t))\| \leq K_u, \quad \|f(t, x_0, \tilde{u}(t))\| \leq K_u,$$

and

$$\zeta \cdot [f(t, x_0, \tilde{u}(t)) - f(t, x_0, \bar{u}(t))] < -\delta,$$

for all $\zeta \in \partial_x^{>} h(s, x),\ s \in [0, \epsilon),\ x \in \{x_0\} + \epsilon_1\mathbb{B}$.

In this chapter we derive a strengthened MP in the same away of [FFV99] but requiring a different and weaker CQ of an integral-type:

CQ$_{\mathbf{I}}$: if $h(0, x_0) = 0$, then there exist positive constants K_u, ϵ, ϵ_1, δ and a control $\tilde{u} \in \Omega(t)$ such that for a.e. $t \in [0, \epsilon)$

$$\|f(t, x_0, \bar{u}(t))\| \leq K_u, \quad \|f(t, x_0, \tilde{u}(t))\| \leq K_u,$$

and for all $t \in [0, \epsilon)$

$$\int_0^t \zeta \cdot [f(\tau, x_0, \tilde{u}(\tau)) - f(\tau, x_0, \bar{u}(\tau))] d\tau \leq -\delta t,$$

for all $\zeta \in \partial_x^{>} h(s, x),\ s \in [0, \epsilon),\ x \in \{x_0\} + \epsilon_1\mathbb{B}$.

[1]Recall that the nontriviality condition in the more conventional MP is

$$\mu\{[0,1]\} + \|p\|_{L^\infty} + \lambda > 0.$$

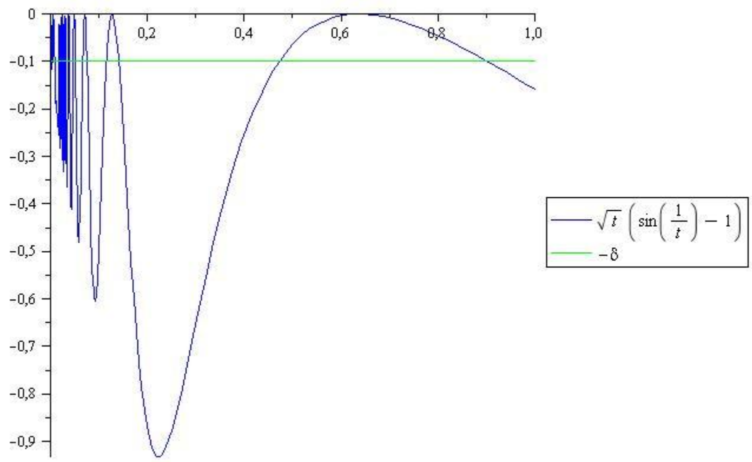

Figure 6.1: Graphic representation of l exceeding any δ we might choose.

It is an easy task to see that $\mathbf{CQ_{FFV99}}$ implies $\mathbf{CQ_I}$. Consequently, the new constraint qualification $\mathbf{CQ_I}$ is applies to a larger class of problems.

To see in more detail $\mathbf{CQ_I}$ as "weaker" condition of $\mathbf{CQ_{FFV99}}$, we reduce $\mathbf{CQ_{FFV99}}$ and $\mathbf{CQ_I}$, respectively, to:

$\exists \delta > 0$ such that

$$l(t) < -\delta \quad \text{a.e.} t \in [0, \epsilon) \tag{6.1}$$

and

$$\int_0^t l(s)ds \leq -\delta t \quad \forall t \in [0, \epsilon), \tag{6.2}$$

Take, for example, the function

$$l(t) = \sqrt{t}(\sin(1/t) - 1).$$

As illustrated in Figure 6.1 and Figure 6.2, this function does not satisfy $\mathbf{CQ_{FFV99}}$ but satisfies $\mathbf{CQ_I}$.

The price to pay for a weaker CQ is the strengthening hypotheses on the data of the problem. In contrast to [FFV99], the NCO given here (valid under $\mathbf{CQ_I}$) require a convex velocity set as an additional hypothesis.

As in [FFV99], we assume that $x \to f(t, x, u)$ is Lipschitz continuous with a

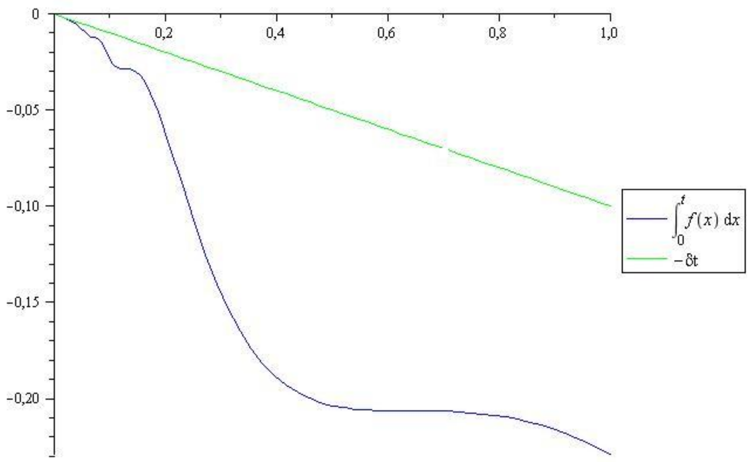

Figure 6.2: Graphic representation of $\int_0^t l(s)ds$ and $-\delta t$ for a particular δ.

constant K_f not depending on t and u, in an initial time interval.

6.2 Nondegenerate Maximum Principle with Integral-type CQ

We impose the basic hypotheses $\mathbf{H1_b}$-$\mathbf{H6_b}$ (see sections 2.3.2 and 2.3.3) and the following two additional hypotheses:

$\mathbf{H2_I}$ There exist scalars $K_f > 0$ and $\epsilon' > 0$ such that

$$\|f(t,x,u) - f(t,x',u)\| \leq K_f \|x - x'\|,$$

for $x, x' \in \bar{x}(0) + \delta'\mathbb{B}$, $u \in \Omega(t)$ a.e.$t \in [0, \epsilon']$.

$\mathbf{H7_I}$ There exists positive constants ϵ and ϵ_1 such that $f(t,x,\Omega(t))$ is convex for all $t \in [0,\epsilon)$ and for all $x \in \{x_0\} + \epsilon_1\mathbb{B}$.

Theorem 6.2.1 *Let $(\bar{x}, \bar{u})$ be a local minimizer for (OCP_2). Assume that hypotheses* $\mathbf{H1_b}$-$\mathbf{H6_b}$*,* $\mathbf{H2_I}$ *and* $\mathbf{H7_I}$*, together with* $\mathbf{CQ_I}$*, are satisfied. Then there exist $p \in W^{1,1}([0,1] : \mathbb{R}^n)$, a measurable function γ, a non-negative measure μ represent-*

ing an element in $C^*([0,1]:\mathbb{R})$ *and* $\lambda \geq 0$ *such that*

$$\mu\{(0,1]\} + \|q\|_{L^\infty} + \lambda > 0, \tag{6.3}$$

$$-\dot{p}(t) \in \operatorname{co}\partial_x^L H(t,\bar{x}(t),q(t),\bar{u}(t)) \quad a.e. t \in [0,1], \tag{6.4}$$

$$-q(1) \in N_C^L(\bar{x}(1)) + \lambda\partial g^L(\bar{x}(1)), \tag{6.5}$$

$$\gamma(t) \in \partial_x^> h(t,\bar{x}(t)) \quad \mu\ a.e., \tag{6.6}$$

$$\operatorname{supp}\{\mu\} \subset \{t \in [0,1] : h\,(t,\bar{x}(t)) = 0\}, \tag{6.7}$$

and, for almost every $t \in [0,1]$, $\bar{u}(t)$ *maximizes over* $\Omega(t)$

$$u \mapsto H(t,\bar{x}(t),q(t),u), \tag{6.8}$$

where

$$q(t) = \begin{cases} p(t) + \displaystyle\int_{[0,t)} \gamma(s)\mu(ds) & t \in [0,1) \\ p(t) + \displaystyle\int_{[0,1]} \gamma(s)\mu(ds) & t = 1. \end{cases}$$

Observe that the set of degenerate multipliers

$$\lambda = 0,\ \mu \equiv \beta\delta_{t=0} \ \text{ and } \ p \equiv -\beta\zeta, \ \text{ with } \ \zeta \in \partial_x^> h(0,x_0) \ \text{ for some } \ \beta > 0,$$

satisfies the traditional nontriviality condition

$$\mu\{[0,1]\} + \|p\|_{L_\infty} + \lambda > 0, \tag{6.9}$$

but not (6.3).

Remark 6.2.2 *When* h *is continuously differentiable,* $\partial_x^> h(0,x_0) = \{h_x(0,x_0)\}$.

The proof of the Theorem above follows the approach in ([FFV99]), i.e., we consider a sequence of approximating problems differing from (OCP_2) insofar as the dynamics near the left endpoint. Modified the standard MP for problems with state constraints applies to each of those problems. Taking limits we obtain the required conclusions.

6.3 Proof of Theorem 6.2.1

We assume that $h(0, x_0) = 0$, since otherwise the MP cannot be satisfied by the trivial multipliers.

Step 1: Consider, for $\alpha \in (0, 1]$, absolutely continuous functions x and y satisfying the system of equations

$$(S) \quad \begin{cases} \dot{x}(t) = f(t, x(t), \bar{u}(t)) + y(t) \cdot \Delta f(t, x(t)) & \text{a.e.} t \in [0, \alpha] \\ x(0) = x_0, & \\ \dot{y}(t) = 0 & \text{a.e.} t \in [0, \alpha] \\ y(0) \in [0, 1] & \end{cases}$$

where

$$\Delta f(t, x) := f(t, x, \tilde{u}(t)) - f(t, x, \bar{u}(t)).$$

Here $\tilde{u}$ is the control function featuring in the constraint qualification $\mathbf{CQ_I}$.

Since $\dot{y} = 0$ and y is absolutely continuous, then y is constant. In what follows, we denote the value of that function by y instead of $y(t)$.

Step 2: By reducing the size of α, we can ensure that

$$h(t, x(t)) \leq 0 \quad \text{for all } t \in [0, \alpha],$$

for all trajectories x solving system (S).

For that, we start by introducing the following lemma, which is a simple consequence of the hypotheses imposed on the data and standard Gronwall-type estimates.

Lemma 6.3.1 *Let x and y be the solution of the system (S) and $\bar{x}$ the minimizer of* *the* (OCP_2)*. There exist positive constants* A *and* B *such that, for* α *small enough,*

$$\begin{aligned} \|x(t) - x_0\| &\leq At \\ \|x(t) - \bar{x}(t)\| &\leq Byt \end{aligned}$$

for all $t \in [0, \alpha]$*.*

Proof. (of Lemma 6.3.1)

Take any $\alpha \in [0, \epsilon)$, where ϵ is defined in $\mathbf{CQ_I}$.

Integrating x we have that

$$\begin{aligned} \|x(t) - x_0\| &\leq \int_0^t \|f(\tau, x(\tau), \bar{u}(\tau)) + y \cdot \Delta f(\tau, x(\tau))\| \, d\tau \\ &= \int_0^t \|f(\tau, x(\tau), \bar{u}(\tau)) - f(\tau, x_0, \bar{u}(\tau)) + y \cdot [f(\tau, x(\tau), \tilde{u}(\tau)) - f(\tau, x_0, \tilde{u}(\tau))] \\ &\quad + y \cdot [-f(\tau, x(\tau), \bar{u}(\tau)) + f(\tau, x_0, \bar{u}(\tau))] + f(\tau, x_0, \bar{u}(\tau)) + \\ &\quad y \cdot [f(\tau, x_0, \tilde{u}(\tau)) - f(\tau, x_0, \bar{u}(\tau))]\| \, d\tau \\ &\leq \int_0^t 3K_f \|x(\tau) - x_0\| \, d\tau + 3K_u t. \end{aligned}$$

Applying Gronwell-Bellman inequality (see e.g. [War72]) yields

$$\begin{aligned} \|x(t) - x_0\| &\leq 3K_u t + e^{3K_f t} \int_0^t 9K_f K_u \tau \, d\tau \\ &= 3K_u t + \frac{9}{2} K_f K_u e^{3K_f t} t^2. \end{aligned}$$

Since $0 \leq t \leq \alpha \leq 1$, we deduce that:

$$\|x(t) - x_0\| \leq 3K_u t + \frac{9}{2} K_f K_u e^{3K_f} t = At,$$

where $A := 3K_u + \frac{9}{2} K_f K_u e^{3K_f}$. The first assertion is proved.

Similarly

$$\begin{aligned}
\|x(t)-\bar{x}(t)\| &\leq \int_0^t \|f(\tau,x(\tau),\bar{u}(\tau))+y\cdot\Delta f(\tau,x(\tau))-f(\tau,\bar{x}(\tau),\bar{u}(\tau))\|\,d\tau \\
&= \int_0^t \|f(\tau,x(\tau),\bar{u}(t))-f(\tau,\bar{x}(\tau),\bar{u}(\tau)) \\
&\quad +y\cdot[f(\tau,x(\tau),\tilde{u}(\tau))-f(\tau,x_0,\tilde{u}(\tau))] \\
&\quad -y\cdot[f(\tau,x(\tau),\bar{u}(\tau))-f(\tau,x_0,\bar{u}(\tau))] \\
&\quad +y\cdot[f(\tau,x_0,\tilde{u}(\tau))-f(\tau,x_0,\bar{u}(\tau))]\|\ d\tau \\
&\leq \int_0^t [K_f\|x(\tau)-\bar{x}(\tau)\|+2yK_f\|x(\tau)-x_0\|]\ d\tau+2yK_ut \\
&\leq \int_0^t K_f\|x(\tau)-\bar{x}(\tau)\|\,d\tau+2yK_f\int_0^t A\tau\,d\tau+2yK_ut \\
&\leq \int_0^t K_f\|x(\tau)-\bar{x}(\tau)\|\,d\tau+yK_fAt^2+2yK_ut.
\end{aligned}$$

Applying Gronwell's Lemma

$$\begin{aligned}
\|x(t)-\bar{x}(t)\| &\leq yK_fAt^2+2yK_ut+e^{K_ft}\int_0^t K_fy[K_fAs^2+2K_us]\,ds \\
&= yK_fAt^2+2yK_ut+yK_fe^{K_ft}\left(\frac{K_fAt^3}{3}+K_ut^2\right).
\end{aligned}$$

As $0\leq t\leq 1$

$$\|x(t)-\bar{x}(t)\| \leq Byt,$$

where $B:=K_fA+2K_u+K_fe^{K_f}\left(K_f\frac{A}{3}+K_u\right)$, proving the second assertion. ■

Choose an α satisfying

$$\alpha<\min\left\{\frac{2\delta}{K_hK_f(2A+B)},\frac{\epsilon_1}{A},\epsilon\right\}. \tag{6.10}$$

Suppose, in contradiction, that for some fixed $t \in [0, \alpha]$

$$h(t, x(t)) > 0. \tag{6.11}$$

Define for $\beta \in [0, 1]$

$$r(\beta) := h(t, \bar{x}(t) + \beta(x(t) - \bar{x}(t))).$$

In view of the properties of h as a function of x, r is continuous. We also have

$$\begin{aligned} r(0) &= h(t, \bar{x}(t)) \leq 0, \\ r(1) &= h(t, x(t)) > 0. \end{aligned}$$

It follows that the set

$$D := \{\beta \in [0, 1] : r(\beta) = 0\}$$

is non-empty, closed and bounded. We can therefore define

$$\beta_m := \max_{\beta \in D} \beta.$$

Since $r(1) > 0$, we have $\beta_m < 1$. Take any $\beta \in (\beta_m, 1]$. Applying the Lebourg Mean-Value Theorem ([Cla83]), we obtain

$$\begin{aligned} h(t, x(t)) - r(\beta) &= \zeta_t \cdot [x(t) - \bar{x}(t) - \beta(x(t) - \bar{x}(t))] \\ &= (1 - \beta)\zeta_t \cdot [x(t) - \bar{x}(t)] \end{aligned} \tag{6.12}$$

for some $\zeta_t \in \operatorname{co} \partial_x^L h(t, \hat{x})$, and $\hat{x}$ in the segment $(x(t), \bar{x}(t) + \beta[x(t) - \bar{x}(t)])$.

Since $r(\beta) > 0$ for all $\beta \in (\beta_m, 1]$, we have that $h(t, \hat{x}) > 0$. Thus, $\operatorname{co}\partial_x^L h(t, \hat{x}) \subset \partial_x^> h(t, \hat{x})$ (see Theorem 10.2.84 and Definition 10.2.90 in appendix).

It follows that $\zeta_t \in \partial_x^> h(t, \hat{x})$.

Expanding the expression (6.12) yields

$$\begin{aligned}
h(t,x(t)) - r(\beta) &= (1-\beta)\,\zeta_t \cdot \int_0^t [f(\tau,x(\tau),\bar{u}(\tau)) + y\Delta f(\tau,x(\tau)) - f(\tau,\bar{x}(\tau),\bar{u}(\tau))]\,d\tau \\
&\leq (1-\beta)\left(\zeta_t \cdot \int_0^t y\Delta f(\tau,x(\tau))\,ds + \|\zeta_t\| K_f \int_0^t \|x(\tau)-\bar{x}(\tau)\|\,d\tau\right) \\
&\leq (1-\beta)\left(\zeta_t \cdot \int_0^t y(\Delta f(\tau,x_0) + \Delta f(\tau,x(\tau)) - \Delta f(\tau,x_0)\,d\tau\right. \\
&\quad \left. + \|\zeta_t\| K_f \int_0^t \|x(\tau)-\bar{x}(\tau)\|\,d\tau\right) \\
&\leq (1-\beta)\left(\int_0^t \zeta_t \cdot y\Delta f(\tau,x_0)\,d\tau + 2K_f\|\zeta_t\| y \int_0^t \|x(\tau)-x_0\|\,d\tau\right. \\
&\quad \left. + K_h K_f \int_0^t \|x(\tau)-\bar{x}(\tau)\|\,d\tau\right) \\
&\leq (1-\beta)\left(\int_0^t \zeta_t \cdot y\Delta f(\tau,x_0)d\tau + 2K_f K_h y \int_0^t \|x(\tau)-x_0\|d\tau\right. \\
&\quad \left. + K_h K_f \int_0^t \|x(\tau)-\bar{x}(\tau)\|d\tau\right) \\
&\leq (1-\beta)\,(-y\delta t + K_h K_f y(A+B/2)t^2) \\
&\leq (1-\beta)\,yt(-\delta + K_h K_f y(A+B/2)t) \\
&\leq 0,
\end{aligned}$$

for all $\beta \in (\beta_m, 1]$.

Here we have used the fact that the norm of every element of the subdifferential is bounded by the Lipschitz rank of the function. In the last two inequalities we have used $\mathbf{CQ_I}$ and (6.10).

Since r is continuous and $r(\beta_m) = 0$, it follows that

$$h(t,x(t)) \leq 0.$$

This contradicts 6.11. The proof is complete.

Step 3: Take a decreasing sequence $\{\alpha_i\}$ on $(0,\alpha)$, converging to zero. Associate with each α_i the following problem (P_i), in which satisfaction of the state constraint is enforced only on the subinterval $[\alpha_i, 1]$:

$$
(P_i) \quad \begin{cases}
\text{Minimize} & g(x(1)) \\
\text{subject to} & \dot{x}(t) = f(t, x(t), \bar{u}(t)) + y(t) \cdot \Delta f(t, x(t)) \\
& \qquad\qquad\qquad\qquad\qquad \text{a.e.} t \in [0, \alpha_i) \\
& \dot{x}(t) = f\,(t, x(t), u(t)) \qquad \text{a.e.} t \in [\alpha_i, 1] \\
& \dot{y}(t) = 0 \qquad \text{a.e.} t \in [0, \alpha_i) \\
& x(0) = x_0 \\
& x(1) \in C \\
& y(0) \in [0, 1] \\
& u(t) \in \Omega(t) \qquad \text{a.e.} t \in [\alpha_i, 1] \\
& h\,(t, x(t)) \leq 0 \qquad \forall t \in [\alpha_i, 1].
\end{cases}
$$

We start by proving the following Lemma.

Lemma 6.3.2 *The trajectory $y \equiv 0$ and $x \equiv \bar{x}$ solves all problems (P_i).*

Proof. (of Lemma 6.3.2) By contradiction assume that there exist $(\hat{y}, \hat{x}) \neq (0, \bar{x})$ that solve (P_i). Hence $g(\hat{x}(1)) < g(\bar{x}(1))$ and $\hat{x}(t) = f(t, \hat{x}(t), \bar{u}(t)) + \hat{y} \cdot \Delta f(t, \hat{x}(t))$ a.e.$t \in [0, \alpha_i)$.

By convexity hypotheses $(\mathbf{H7_I})$

$$\hat{y} f(t, \hat{x}, \hat{u}) + (1 - \hat{y}) f(t, \hat{x}, \bar{u}) \in f(t, x(t), \Omega(t)).$$

Then $\exists \hat{u}(\cdot) : [0, 1] \to \mathbb{R}^m$:

$$\hat{x}(t) = f(t, \hat{x}, \hat{u}(t)) \text{ a.e.} t \in [0, 1].$$

We conclude that $\hat{x}$ is an admissible trajectory for (OCP_2) with $g(\hat{x}(1)) < g(\bar{x}(1))$. ∎

The Maximum Principle (Theorem 2.3.5) for the problem (P_i) asserts the existence of an arc $(p_i, c_i) : [0,1] \mapsto \mathbb{R}^n \times \mathbb{R}$, a measurable function γ_i, a nonnegative Radon measure $\mu_i \in C^*([\alpha_i, 1], \mathbb{R})$, and a scalar $\lambda_i \geq 0$ such that

$$\mu_i\{[\alpha_i, 1]\} + \|(p_i, c_i)\|_{L^\infty} + \lambda_i > 0, \tag{6.13}$$

$$-\dot{p}_i(t) \in \begin{cases} p_i(t) \cdot co\partial_x^L f(t, \bar{x}(t), \bar{u}(t)), \text{a.e. } t \in [0, \alpha_i), \\ \left(p_i(t) + \displaystyle\int_{[\alpha_i, t)} \gamma_i(s)\mu_i(ds)\right) \cdot co\partial_x^L f(t, \bar{x}(t), \bar{u}(t)), \text{a.e. } t \in [\alpha_i, 1], \end{cases}$$

$$-\dot{c}_i(t) = \begin{cases} p_i(t) \cdot \Delta f(t, \bar{x}(t)), & \text{a.e. } t \in [0, \alpha_i), \\ 0, & \text{a.e. } t \in [\alpha_i, 1], \end{cases} \tag{6.14}$$

for almost every $t \in [\alpha_i, 1]$, $\bar{u}(t)$ maximizes over $\Omega(t)$

$$u \mapsto \left(p_i(t) + \int_{[\alpha_i, t)} \gamma_i(s)\mu_i(ds)\right) \cdot f\left(t, \bar{x}(t), u\right), \tag{6.15}$$

$$\text{supp}\{\mu_i\} \subset \left\{t \in [\alpha_i, 1] : h\left(t, \bar{x}(t)\right) = 0\right\},$$

$$\gamma_i(t) \in \partial_x^> h(t, \bar{x}(t)) \quad \mu \, \text{a.e.},$$

for some $\xi_i \in \partial_x^L g(\bar{x}(1))$,

$$-\left(p_i(1) + \int_{[\alpha_i, 1]} \gamma_i(s)\mu_i(ds) + \lambda_i \xi_i\right) \in N_C^L(\bar{x}(1)),$$

$$-c_i(1) = 0,$$

$$c_i(0) \in N_{[0,1]}^L(0).$$

It remains to pass to the limit as $i \to \infty$ and thereby, obtain a set of nondegenerate multipliers for the original problem.

Without changing the notation, we extend μ_i as a regular Borel measure on $[0,1]$

$$\mu_i(\mathcal{B}) = \mu_i(\mathcal{B} \cap [\alpha_i, 1]) \text{ for all Borel set } \mathcal{B} \subset [0,1].$$

Extend also γ_i, originally defined on $[\alpha_i, 1]$, arbitrarily to the interval $[0,1]$ as a Borel measurable function. With theses extensions and noting that $\mu([0,\alpha_i)) = 0$ we can write

$$-\dot{p}_i(t) \in \left(p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds) \right) \cdot co\partial_x^L f(t, \bar{x}(t), \bar{u}(t)) \text{a.e.} t \in [0,1].$$

It is easy to see that c_i can be omitted from (6.13), since $p_i \equiv 0$ implies $c_i \equiv 0$. By scaling the multipliers we ensure that

$$\|\mu_i\{[\alpha_i, 1]\}\| + \|p_i\|_{L^\infty} + \lambda_i = 1. \tag{6.16}$$

The multifunction $\partial_x^> h$ is uniformly bounded, compact, convex, and has a closed graph. Since $\{p_i\}$ is uniformly bounded and $\{\dot{p}_i\}$ is uniformly integrally bounded, we can arrange by means of subsequence extraction (Proposition 10.2.65 and Proposition 10.2.67 in appendix) that

$$p_i \to p \text{ uniformly}, \quad \gamma_i d\mu_i \to \gamma d\mu \text{ weak}^*, \quad \lambda_i \to \lambda, \xi_i \to \xi,$$

where μ is the weak* limit of μ_i in the space of nonnegative-valued functions in $C^*([0,1], \mathbb{R})$, γ is a measurable selection of $\partial_x^> h(t, \bar{x}(t))$ μ a.e., and $\xi \in \partial^L g(\bar{x}(1))$. To obtain ξ we have used the fact that $\partial^L g(\bar{x}(1))$ is a compact set.

It follows that the conditions (6.7), (6.6), (6.4) for problem (OCP_2) are satisfied and since $N_C^L(\bar{x}(1))$ is closed, (6.5) also holds. Moreover from (6.16) we deduce

$$\mu\{[0,1]\} + \|p\|_{L^\infty} + \lambda = 1. \tag{6.17}$$

Consider the set $S_i = [\alpha_i, 1]\backslash\Omega_i$ where Ω_i is a null Lebesgue measure set in $[\alpha_i, 1]$ containing all times where the maximization of (6.15) is not achieved at $\bar{u}$. We can then write

$$\left(p_i(t) + \int_{[\alpha_i,t)} \gamma_i(s)\mu_i(ds)\right)\cdot f\left(t, \bar{x}(t), u\right) \leq \left(p_i(t) + \int_{[\alpha_i,t)} \gamma_i(s)\mu_i(ds)\right)\cdot f\left(t, \bar{x}(t), \bar{u}(t)\right),$$

for all $t \in S_i$ and for all $u \in \Omega(t)$.

Now consider the full measure set $S = (0, 1]\backslash \bigcup_i \Omega_i$. Fix some t in S. Then for all $i > N$, where N is such that $\alpha_N \leq t$, we have

$$\left(p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)\right)\cdot f\left(t, \bar{x}(t), u\right) \leq \left(p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)\right)\cdot f\left(t, \bar{x}(t), \bar{u}(t)\right).$$

for all $u \in \Omega(t)$. Applying limits to both sides of this inequality we obtain (6.8).

We have established that the set of multipliers (p, μ, λ), obtained as limit of (p_i, μ_i, λ_i) satisfy the conditions (6.4)- (6.8) for the original problem (OCP_2) together with (6.17).

Step 4: It remains to verify

$$\mu\{(0, 1]\} + ||q||_{L^\infty} + \lambda > 0.$$

We start by proving the following lemma:

Lemma 6.3.3 *The adjoint vector p_i in the necessary conditions of optimality for problem (P_i) satisfies*

$$\int_0^{\alpha_i} p_i(t) \cdot \Delta f(t, \bar{x}(t))dt \leq 0.$$

Proof. (of Lemma 6.3.3)

Since the cost function g does not depend on y, we have $c_i(1) = 0$. The set $N^L_{[0,1]}(0)$ is $(-\infty, 0]$, so $c_i(0) \leq 0$. Now, by integrating the differential equation involving c_i (6.14) we get

$$c_i(1) = c_i(0) + \int_0^{\alpha_i} -p_i(t) \cdot \Delta f(t, \bar{x}(t)) = 0.$$

The result easily follows.

■

In view of the constraint qualification, there exists positive constants ϵ and δ such that for all $t \in [0, \epsilon)$

$$\int_0^t \zeta \cdot [f(\tau, x_0, \tilde{u}(\tau)) - f(\tau, x_0, \bar{u}(\tau))] d\tau \leq -\delta t$$

for all $\zeta \in \partial_x^> h(s, x),\ s \in [0, \epsilon),\ x \in \{x_0\} + \epsilon_1 \mathbb{B}$.

Suppose, in contradiction, that

$$\mu\{(0, 1]\} + ||q||_{L^\infty} + \lambda = 0.$$

Since (6.17), we must have

$$\begin{aligned} &\lambda = 0, \\ &\mu = \beta \delta_{\{0\}}, \\ &p(t) = -\beta\zeta \quad \text{for some } \beta > 0 \text{ and } \zeta \in \partial_x^> h(0, x_0). \end{aligned}$$

The constraint qualification ($\mathbf{CQ_I}$) implies

$$\int_0^t -p(s) \cdot \Delta f(s, x_0) ds = \int_0^t \beta\zeta \cdot \Delta f(s, x_0) ds \leq -\delta\beta t.$$

On other hand

$$\begin{aligned} &\int_0^{\alpha_i} p_i(t) \cdot \Delta f(t, \bar{x}(t)) dt \\ &= \int_0^{\alpha_i} p(t) \cdot \Delta f(t, x_0) + (p_i(t) - p(t)) \Delta f(t, x_0) + p_i(t)[\Delta f(t, \bar{x}(t)) - \Delta f(t, x_0)] dt \\ &\geq \delta\beta\alpha_i - \int_0^{\alpha_i} 2K_u \|p_i(t) - p(t)\| + 2K_f \|\bar{x}(t) - x_0\| \|p_i(t)\| dt \\ &\geq \delta\beta\alpha_i - \int_0^{\alpha_i} 2K_u \|p_i(t) - p(t)\| + 2K_f At \|p_i(t)\| dt. \end{aligned}$$

By the uniform convergence of p_i, we can make $\|p_i - p\| < \bar{\epsilon}$ for any $\bar{\epsilon} > 0$ of our choice provided we choose a sufficient large i. Moreover $\|p_i\| \leq 1$.

It follows that

$$\int_0^{\alpha_i} p_i(t) \cdot \Delta f(t, \bar{x}(t)) dt \geq \delta\beta\alpha_i - (2K_u\bar{\epsilon}\alpha_i + K_f A\alpha_i^2) > \delta\beta/2\alpha_i > 0,$$

if $\bar{\epsilon} < \dfrac{\delta\beta}{8K_u}$ and $\alpha_i < \dfrac{\delta\beta}{4K_f A}$.

So, we would have $\displaystyle\int_0^{\alpha_i} p_i(t) \cdot \Delta f(t, \bar{x}(t)) dt > 0$ contradicting Lemma 6.3.3. We deduce (6.3).

Notes on Chapter

The contents of this chapter were published in [LFdP07].

Chapter 7

Nondegeneracy with easier verifiable Constraint Qualification

In the literature strengthened forms of the MP to avoid degeneracy are validated under mainly two types of constraints qualifications:

Type 1: assume the existence of a control pulling the state away from the state constraint boundary faster then the optimal control on a neighborhood of the initial time.

Type 2: assume the existence of a control pulling the state away from the state constraint boundary in a neighborhood of the initial time.

Results involving CQ of type 1 are tipically valid on weaker conditions on the data of the problem. The main setback of this type of CQ is that it involves the optimal control which we do not known in advance, and consequently, this condition is, in general not easy to verify, except in special cases, such as CVP. In this chapter, we discuss the hypotheses under which the first type of CQ can be reduced to the second one.

7.1 Introduction

We focus on the following problem:

$$(OCP_3) \begin{cases} \text{Minimize} & g(x(1)) & \\ \text{subject to} & \dot{x}(t) = f\left(t, x(t), u(t)\right) & \text{a.e.} t \in [0,1] \\ & x(0) = x_0 & \\ & u(t) \in \Omega(t) & \text{a.e.} t \in [0,1] \\ & h\left(x(t)\right) \leq 0 & \text{for all } t \in [0,1]. \end{cases}$$

Observe that the functional defining the state constraint does not depend explicitly on t.

As we mentioned before the CQ to avoid degeneracy of type 1 mainly impose:

$\mathbf{CQ1_d}$: $\exists \delta, \epsilon > 0$ and $\exists \tilde{u}(t) \in \Omega(t)$:

$$h_x(x_0) \cdot [f(t, x_0, \tilde{u}(t)) - f(t, x_0, \bar{u}(t)] < -\delta \quad \text{a.e.} t \in [0, \epsilon).$$

Whereas CQ of type 2 imposed:

$\mathbf{CQ2_d}$: $\exists \delta, \epsilon > 0$ and $\exists \tilde{u}(t) \in \Omega(t)$:

$$h_x(x_0) \cdot f(t, x_0, \tilde{u}(t)) < -\delta \quad \text{a.e.} t \in [0, \epsilon).$$

Nondegenerate results involving a CQ of the type $\mathbf{CQ1_d}$ can be found in [FV94] and [FFV99]. The result in [FFV99] generalizes the nondegenerate result in [FV94], by allowing the final state to belong a given set C_1, the data to be nonsmooth and by not requiring the velocity set $f(t, x, \Omega(t))$ to be convex and the data is merely measurable with respect to the time variable. In [AA97] and [RV00], the nondegenerate results involve a constraint qualification of the type $\mathbf{CQ2_d}$. In [AA97], it is required that the velocity set $f(t, x, \Omega(t))$ is convex and the data is Lipschitiz continuous with respect to the time variable. On the other hand, in [RV00], the nondegenerate NCO are derived for OCP involving differential inclusion conditions with general endpoint constraints. Moreover, the data is measurable with respect to the time variable, velocity sets are nonconvex, and $h : \mathbb{R}^n \to \mathbb{R}$ are functions of class $C^{1,1}$ (functions which are continuously differentiable with locally Lipschitz continuous derivatives).

As stated before results involving $\mathbf{CQ1_d}$ type conditions require less regularity on data. However, $\mathbf{CQ1_d}$ involves the minimizing $\bar{u}$ which we do not know in advance, and consequently the condition is, in general not easily verifiable, except in special cases, such as calculus of variations problems. (see chapter 4.)

In this chapter we developed a strengthened MP with CQ of type $\mathbf{CQ2_d}$.

To prove this result, we consider three cases:

Case 1: the minimizing state trajectory leaves the boundary immediately;

Case 2: the minimizing state trajectory remains on the boundary on a neighborhood of the initial time;

Case 3: Case 1 or Case 2 occurs a infinite numbers of times on neighborhood of the initial time.

Case 3 will be clarified shortly.

Remark 7.1.1 *We note that the case 2 can occurs a infinite numbers of times on neighborhood of the initial time, one example of that it is consider*

$$h(x) = \min\{0, \sin(\frac{1}{x})\}.$$

In case 1, we apply the nondegenerate result developed in [FV94] under weaker hypotheses.

In case 2 and case 3, we show that $\mathbf{CQ2_d}$ implies $\mathbf{CQ1_d}$ and consequently we are in conditions to apply the nondegenerate result developed in [FFV99].

7.2 Easier Verifiable Nondegenerate Result

Assuming that, there exists a $\delta' > 0$, such that

$\mathbf{H1_{EV}}$ The function $(t, u) \to f(t, x, u)$ is continuous for each x;

$\mathbf{H2_{EV}}$ There exists a $\mathcal{L} \times \mathcal{B}$ measurable function $k(t, u)$ such that $t \mapsto k(t, \bar{u}(t))$ is integrable and

$$\|f(t, x, u) - f(t, x', u)\| \leq k(t, u)\|x - x'\|$$

for $x, x' \in \bar{x}(t) + \delta'\mathbb{B}$, $u \in \Omega(t)$ a.e.$t \in [0, 1]$. Furthermore there exist scalars $K_f > 0$ and $\epsilon' > 0$ such that

$$\|f(t, x, u) - f(t, x', u)\| \leq K_f\|x - x'\|$$

for $x, x' \in \bar{x}(0) + \delta'\mathbb{B}$, $u \in \Omega(t)$ a.e.$t \in [0, \epsilon']$.

$\mathbf{H3_{EV}}$ The function g is locally Lipschitz continuous;

$\mathbf{H4_{EV}}$ The $Gr\ \Omega$ is a Borel set;

$\mathbf{H5_{EV}}$ The function $x \to h(x)$ is continuously differentiable.

Additionally, assume that

$\mathbf{CQ_{EV}}$ If $h(x_0) = 0$, then there exist positive constants K_u, ϵ, δ, and a control $\tilde{u} \in \mathcal{U}$ such that for a.e.$t \in [0, \epsilon)$

$$\|f(t, x_0, \bar{u}(t))\| \leq K_u, \quad \|f(t, x_0, \tilde{u}(t))\| \leq K_u,$$

and

$$h_x(x_0) \cdot f(t, x_0, \tilde{u}(t)) < -\delta.$$

The constraint qualification $\mathbf{CQ_{EV}}$ is of type $\mathbf{CQ2_d}$.

Theorem 7.2.1 *Let $(\bar{x}, \bar{u})$ be a local minimizer for (OCP_3), where the optimal control is a piecewise continuous function to the left. Assume that hypotheses* $\mathbf{H1_{EV}}$*-*$\mathbf{H5_{EV}}$ *together with* $\mathbf{CQ_{EV}}$ *are satisfied. Then there exist $p \in W^{1,1}([0, 1] : \mathbb{R}^n)$,* *a measurable function γ, a non-negative measure μ representing an element in $C^*([0,1]:\mathbb{R})$ and $\lambda \geq 0$ such that*

$$\mu\{(0,1]\} + \|q\|_{L^\infty} + \lambda > 0. \tag{7.1}$$

$$-\dot{p}(t) \in \mathrm{co}\partial_x^L \left(q(t) \cdot f(t, \bar{x}(t), \bar{u}(t))\right), \text{ a.e. } t \in [0,1],$$

$$q(1) \in \lambda \partial^L g(\bar{x}(1)),$$

$$\mathrm{supp}\{\mu\} \subset \{t \in [0,1] : h(\bar{x}(t)) = 0\},$$

and for almost every $t \in [0,1]$, $\bar{u}(t)$ maximizes over $\Omega(t)$

$$u \mapsto q(t) \cdot f(t, \bar{x}(t), u)$$

where

$$q(t) = \begin{cases} p(t) + \displaystyle\int_{[0,t)} h_x(\bar{x}(s))\mu(ds) & t \in [0,1) \\ p(1) + \displaystyle\int_{[0,1]} h_x(\bar{x}(s))\mu(ds) & t = 1. \end{cases}$$

Remark 7.2.2 *Notice that the* $\mathbf{CQ_{EV}}$ *allow to replace the more traditional non-triviality condition of the MP*

$$\mu\{[0,1]\} + \|p\|_{L_\infty} + \lambda > 0,$$

by the stronger condition (7.1).

Remark 7.2.3 *In the Theorem, we assume that the optimal control is piecewise continuous to the left (not merely measurable), which is a strong condition but it is satisfied in many applications.*

Remark 7.2.4 *Comparing this result with [RV00], the result of [RV00] have the advantage of be applied for problems in which the initial and final states belong to given sets and it is not required the continuity of* $(t, u) \to f(t, x, u)$ *and piecewise continuity to the left of* $\bar{u}$. *However, here we have*

- *weaker hypotheses on state constraint,* $h(\cdot)$ *have to be continuously differentiable and not of class* $C^{1,1}$;
- *CQ has to be satisfy just along the optimal trajectory.*

7.3 Proof of Theorem 7.2.1

We assume that $h(x_0) = 0$, since otherwise the conditions of MP cannot be satisfied by the degenerate multipliers.

So, we can consider three cases:

Case 1: the minimizing state trajectory leaves the boundary immediately, i.e. there exists $r \in (0, 1)$ such that $h(\bar{x}(t)) < 0, \forall t \in (0, r]$;

Case 2: the minimizing state trajectory remains on the boundary on a neighborhood of the initial time, i.e. there exists $r \in (0, 1]$ such that $h(\bar{x}(t)) = 0, \forall t \in [0, r]$;

Case 3: Case 1 or Case 2 occurs a infinite numbers of times on neighborhood of the initial time, i.e. there exists a sequence $\{a_j\}$ such that $\{a_j\} \downarrow 0$ and $a_j \in [0, \epsilon]$, $\forall j \in \mathbb{N}$ where

$$\begin{cases} h(\bar{x}(t)) < 0 & \text{for all } t \in (a_{2j-1}, a_{2j}), \text{ all } j \geq 1 \\ h(\bar{x}(t)) = 0 & \text{for all } t \in (a_{2j}, a_{2j+1}), \text{ all } j \geq 1 \\ h(\bar{x}(a_j)) = 0 & \text{for all } j \geq 1. \end{cases}$$

Step 1: We next prove the theorem when case 1 holds.

In the Proposition below, we show that Proposition 2.2 in [FV94] is valid under weaker hypotheses.

Proposition 7.3.1 *Suppose there exists* $r \in (0,1)$ *such that* $h(\bar{x}(t)) < 0, \forall t \in (0,r]$. *Assume also that* (OCP_3) *satisfy the basic hypotheses* $\mathbf{H1_b}$-$\mathbf{H6_b}$ *(see sections 2.3.2 and 2.3.3), then* $(\bar{x},\bar{u})$ *satisfies the conditions of the MP with multipliers* (p,μ,λ) *for which*

$$\mu\{(0,1]\} + \lambda \neq 0.$$

Proof. Take a sequence of points $\{\alpha_i\}$ converging to 0, such that $\alpha_i \in (0,r]$, where r is a point in $(0,1]$ such that $h(\bar{x}(t)) < 0, \forall t \in (0,r]$. Let (P_i), with $i = 1,2,...,$ be a modification of (OCP_3) in which the time interval is $[\alpha_i, 1]$ and the initial condition,

$$x(\alpha_i) = \bar{x}(\alpha_i).$$

That means:

$$(P_i) \quad \begin{cases} \text{Minimize} & g(x(1)) & \\ \text{subject to} & \dot{x}(t) = f\left(t, x(t), u(t)\right) & \text{a.e.} t \in [\alpha_i, 1] \\ & x(\alpha_i) = \bar{x}(\alpha_i) & \\ & u(t) \in \Omega(t) & \text{a.e.} t \in [\alpha_i, 1] \\ & h\left(x(t)\right) \leq 0 & \text{for all } t \in [\alpha_i, 1]. \end{cases}$$

For each i, the process for (P_i) comprising the restrictions of $\bar{u}$ and $\bar{x}$ to $[\alpha_i, 1]$, is a minimizing process for (P_i).

Applying the conventional MP (Theorem 2.3.5) to (P_i), we can ensure the existence of the multipliers $(\tilde{p}_i, \tilde{\mu}_i, \tilde{\lambda}_i)$ such that

$$\tilde{\mu}_i\{[\alpha_i, 1]\} + \tilde{\lambda}_i > 0,$$

$$-\dot{\tilde{p}}_i(t) \in \operatorname{co} \partial_x^L (\tilde{p}_i(t) + \int_{[\alpha_i, t)} h_x(\bar{x}(s))\tilde{\mu}_i(ds)) \cdot f(t, \bar{x}(t), \bar{u}(t)) \quad \text{a.e.} t \in [\alpha_i, 1],$$

$$\tilde{p}_i(1) + \int_{[\alpha_i, 1)} h_x(\bar{x}(s))\tilde{\mu}_i(ds) \in \lambda \partial^L g(\bar{x}(1)),$$

$$\text{supp } \{\tilde{\mu}_i\} \subset \{t \in [\alpha_i, 1] : h(\bar{x}(t)) = 0\}, \tag{7.2}$$

and for almost every $t \in [\alpha_i, 1]$, $\bar{u}(t)$ maximizes over $\Omega(t)$

$$u \to (\tilde{p}_i(t) + \int_{[\alpha_i,t)} h_x(\bar{x}(s))\tilde{\mu}_i(ds)) \cdot f(t, \bar{x}(t), u).$$

Remark 7.3.2 *Since, we are assuming that the final state of* (OCP_3) *is free, we can omitted* $\tilde{p}_i$ *from the nontriviality condition. (*$\tilde{\lambda_i} = 0$ *and* $\tilde{\mu}_i[\alpha_i, 1] \equiv 0$ *implies* $p_i \equiv 0$*).*

Then,

$$supp\{\tilde{\mu}_i\} \subset [r, 1]\,,\ \forall i.$$

For each i, denote by μ_i the extension of $\tilde{\mu}_i$ to the Borel subsets of the interval $[0, 1]$,

$$\mu_i(A) = \tilde{\mu}_i(A \cap [r, 1]), \tag{7.3}$$

and by p_i the extension of $\tilde{p}_i$ to the interval $[0, 1]$,

$$p_i(t) = \begin{cases} \tilde{p}_i(\alpha_i) \text{ for } 0 \le t \le \alpha_i \\ \tilde{p}_i(t) \text{ for } \alpha_i \le t \le 1. \end{cases}$$

By scaling the multipliers, we can ensure that

$$\int_{[0,1]} \mu_i(ds) + \lambda_i = 1. \tag{7.4}$$

By means of subsequence extraction we can arrange that

$$p_i \to p \text{ uniformly, } \mu_i \to \mu \text{ weakly}^* \text{ and } \lambda_i \to \lambda$$

for some (p, μ, λ), which are multipliers for $(\bar{x}, \bar{u})$. (see Proposition 10.2.65 and Proposition 10.2.67 in Appendix)

Consider the continuous function

$$\Phi(t) = \begin{cases} \frac{t}{r} \text{ for } 0 \le t \le r \\ 1 \text{ for } r \le t \le 1. \end{cases}$$

According to (7.3)

$$\int_{[0,1]} \mu_i(ds) = \int_{[0,1]} \Phi(s)\mu_i(ds).$$

By weak* converge and (7.4), we have

$$\int_{[0,1]} \Phi(s)\mu(ds) + \lambda = 1.$$

Since $\Phi(0) = 0$, however

$$\int_{(0,1]} \mu(ds) + \lambda \neq 0.$$

■

Since $x \to h(x)$ is continuous differentiable and $x(t)$ is an absolute continuous function, we have

$$\frac{d}{dt}h(x(t)) \text{ exits for a.e. } t \in [0,1].$$

Step 2: Next we prove the Theorem in **Case 2**. In the **Case 2**, there exists $r \in (0,1]$ such that $h(\bar{x}(t)) = 0, \forall t \in [0,r]$.

For a.e. $t \in [0,r)$, we have

$$\frac{d}{dt}h(\bar{x}(t)) = 0 \Leftrightarrow$$

$$\Leftrightarrow h_x(\bar{x}(t)) \cdot f(t,\bar{x}(t),\bar{u}(t)) = 0.$$

On other hand,

$$\begin{aligned} &|h_x(\bar{x}(t)) \cdot f(t,\bar{x}(t),\bar{u}(t)) - h_x(x_0) \cdot f(t,x_0,\bar{u}(t))| \\ &\quad < |h_x(\bar{x}(t)) \cdot f(t,\bar{x}(t),\bar{u}(t)) - h_x(x_0) \cdot f(0,x_0,\bar{u}(0))| \\ &\quad + |h_x(x_0) \cdot f(0,x_0,\bar{u}(0)) - h_x(x_0) \cdot f(t,x_0,\bar{u}(t))|. \end{aligned}$$

Since $\bar{u}$ is piecewise continuous on the left, $t \to h_x(\bar{x}(t))$ and $t \to f(t, \bar{x}(t), \bar{u}(t))$ are continuous on a neighborhood of the initial time, there exists r_0 sufficiently near of 0 such that for all $t \in [0, r_0]$, we have

$$|h_x(\bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t))| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.$$

Therefore,

$$|h_x(x_0) \cdot f(t, x_0, \bar{u}(t))| < \varepsilon, \text{ a.e. } t \in [0, r_0).$$

Choose $\varepsilon = \frac{\delta}{2}$, where δ is defined as in $\mathbf{CQ_{EV}}$. Since $\mathbf{CQ_{EV}}$ is satisfied, then, for a.e. $t \in [0, \min\{r_0, \epsilon)\}$, we have

$$h_x(x_0) \cdot [f(t, x_0, \tilde{u}(t)) - f(t, x_0, \bar{u}(t))] < -\frac{\delta}{2}.$$

Therefore, we are in conditions to apply the main result in [FFV99] and the result holds immediately.

Step 3: Finally we treat **Case 3**. In the **Case 3**, there exists a sequence $\{a_j\}$ such that $\{a_j\} \downarrow 0$ and $a_j \in [0, \epsilon]$, $\forall j \in \mathbb{N}$ where

$$\begin{cases} h(\bar{x}(t)) < 0 & \text{for all } t \in (a_{2j-1}, a_{2j}), \text{ all } j \geq 1 \\ h(\bar{x}(t)) = 0 & \text{for all } t \in (a_{2j}, a_{2j+1}), \text{ all} j \geq 1 \\ h(\bar{x}(a_j)) = 0 & \text{for all } j \geq 1. \end{cases}$$

We first claim that there exists $\exists s_j \in]a_{2j-1}, a_{2j}[\, \forall j \in \mathbb{N}$:

$$h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) \geq 0. \tag{7.5}$$

Seeking a contradiction assume that for all $t \in]a_{2j-1}, a_{2j}[$

$$h_x(\bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t)) < 0, \; \forall t \in]a_{2j-1}, a_{2j}[.$$

Then

$$\int_t^{a_{2j}} h_x(\bar{x}(s)) \cdot f(s, \bar{x}(s), \bar{u}(s)) ds < 0.$$

So

$$h(\bar{x}(a_{2j})) - h(\bar{x}(t)) < 0.$$

As $h(\bar{x}(a_{2j})) = 0$ and $h(\bar{x}(t)) \leq 0$, the contradiction obtained proves our claim.

By $\mathbf{CQ_{EV}}$ and (7.5) we get

$$h_x(x_0) \cdot f(t, x_0, \tilde{u}(t)) - h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) < -\delta, \ \forall t \in [0, \epsilon).$$

Therefore,

$$\begin{aligned} h_x(x_0) \cdot f(t, x_0, \tilde{u}(t)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t)) < -\delta + \\ h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t)). \end{aligned}$$

To finish our proof, we claim that $\exists \epsilon^* > 0$ such that a.e. $t \in [0, \epsilon^*]$ and $\forall \varepsilon > 0$

$$|h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t))| < \varepsilon.$$

Note that

$$|h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t))|$$

$$\begin{aligned} &= |h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(0, x_0, \bar{u}(0)) + h_x(x_0) \cdot f(0, x_0, \bar{u}(0))| \\ &-h_x(x_0) \cdot f(t, x_0, \bar{u}(t))| \end{aligned}$$

$$\begin{aligned} &\leq |h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(0, x_0, \bar{u}(0))| \\ &+|h_x(x_0) \cdot f(0, x_0, \bar{u}(0)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t))|. \end{aligned}$$

Since $\bar{u}$ is piecewise continuous on the left, $t \to h_x(\bar{x}(t))$ and $t \to f(t, \bar{x}(t), \bar{u}(t))$ are continuous on a neighborhood of the initial time and $s_j \downarrow 0$, there exist j_1

sufficient large such that $\forall j \geq j_1$,

$$|h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(0, x_0, \bar{u}(0))| < \varepsilon, \quad \forall \varepsilon > 0.$$

By continuity of $t \to f(t, x_0, \bar{u}(t))$, we also conclude that there exist r_0 sufficient near of 0 such that for all $t \in [0, r_0)$

$$|h_x(x_0) \cdot f(0, x_0, \bar{u}(0)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t))| < \varepsilon, \quad \forall \varepsilon > 0.$$

Choosing $\varepsilon = \frac{\delta}{4}$, we can apply Theorem 2.1. in [FFV99] for a.e. $t \in [0, \min\{s_{j_1}, r_0, \epsilon\})$ and the result holds immediately.

Notes on Chapter

Chapter 8

Nondegeneracy with easier verifiable Integral-type Constraint Qualification

In this chapter we show that the strengthened Maximum Principle derived in the previous chapter is valid under a different integral-type constraint qualification. In contrast to the constraint qualification used in chapter 6 the constraint qualification we shall focus on is easier to verify since it does not require a priori knowledge of the optimal control $\bar{u}$. We compare the results obtained here with those of the previous chapter.

8.1 Easier Verifiable Nondegenerate Result with Integral-type CQ

Consider the problem OCP_3.

$$(OCP_3) \quad \begin{cases} \text{Minimize} & g(x(1)) & \\ \text{subject to} & \dot{x}(t) = f\left(t, x(t), u(t)\right) & \text{a.e.} t \in [0,1] \\ & x(0) = x_0 & \\ & u(t) \in \Omega(t) & \text{a.e.} t \in [0,1] \\ & h\left(x(t)\right) \leq 0 & \text{for all } t \in [0,1]. \end{cases} .$$

The following hypotheses, involving a parameter $\delta' > 0$, will be of use:

H1$_{\mathbf{EVI}}$ The function $(t,u) \to f(t,x,u)$ is $\mathcal{L}\times\mathcal{B}$ measurable for each x;($\mathcal{L}\times\mathcal{B}$ denotes the product σ-algebra generated by the Lebesgue subsets $\mathcal{L}$ of $[0,1]$ and the Borel subsets of $\mathbb{R}^m$.)

H2$_{\mathbf{EVI}}$ There exists a $\mathcal{L}\times\mathcal{B}$ measurable function $k(t,u)$ such that $t \mapsto k(t,\bar{u}(t))$ is integrable and

$$\|f(t,x,u) - f(t,x',u)\| \leq k(t,u)\|x - x'\|$$

for $x, x' \in \bar{x}(t) + \delta'\mathbb{B}$, $u \in \Omega(t)$ a.e.$t \in [0,1]$. Furthermore there exist scalars $K_f > 0$ and $\epsilon' > 0$ such that

$$\|f(t,x,u) - f(t,x',u)\| \leq K_f\|x - x'\|$$

for $x, x' \in \bar{x}(0) + \delta'\mathbb{B}$, $u \in \Omega(t)$ a.e.$t \in [0,\epsilon']$.

H3$_{\mathbf{EVI}}$ There exists positive constants ϵ and ϵ_1 such that $f(t,x,\Omega(t))$ is convex for all $t \in [0,\epsilon)$ and for all $x \in \{x_0\} + \epsilon_1\mathbb{B}$.

H4$_{\mathbf{EVI}}$ The function g is locally Lipschitz continuous;

H5$_{\mathbf{EVI}}$ The $Gr\ \Omega$ is a Borel set;

H6$_{\mathbf{EVI}}$ The function $x \to h(x)$ is continuously differentiable.

CQ$_{\mathbf{EVI}}$ If $h(x_0) = 0$, then there exist positive constants K_u, ϵ, δ, and a control $\tilde{u} \in \mathcal{U}$ such that for a.e. $t \in [0, \epsilon_2)$

$$\|f(t, x_0, \bar{u}(t))\| \leq K_u, \quad \|f(t, x_0, \tilde{u}(t))\| \leq K_u,$$

and for all $t \in [0, \epsilon_2)$

$$\int_0^t h_x(x_0) \cdot f(s, x_0, \tilde{u}(s))ds < -\delta t.$$

In contrast with CQ in the chapter 6, **CQ$_{\mathbf{EVI}}$** must be satisfied for an admissible control. Thus a priori knowledge of the optimal control is not needed. As we see next, the strengthened as stated in (6.2.1) is still valid under our new and easier to verify CQ.

Another point of interest in exploring easier verifiable integral-type CQ is be the fact that **CQ$_{\mathbf{EV}}$** implies **CQ$_{\mathbf{EVI}}$**, and consequently the nondegenerate result involving **CQ$_{\mathbf{EVI}}$** applies to a larger class of problems.

Theorem 8.1.1 *Let $(\bar{x}, \bar{u})$ be a local minimizer for (OCP_3). Assume that hypotheses* **H1$_{\mathbf{EVI}}$-H6$_{\mathbf{EVI}}$** *together with* **CQ$_{\mathbf{EVI}}$** *are satisfied. Then there exist $p \in W^{1,1}([0,1] : \mathbb{R}^n)$, a measurable function γ, a non-negative measure μ representing an element in $C^*([0,1] : \mathbb{R})$ and $\lambda \geq 0$ such that*

$$\mu\{(0,1]\} + \|q\|_{L^\infty} + \lambda > 0, \tag{8.1}$$

$$-\dot{p}(t) \in \text{co}\partial_x^L \left(q(t) \cdot f(t, \bar{x}(t), \bar{u}(t))\right) \ a.e.\ t \in [0,1],$$

$$q(1) \in \lambda \partial^L g(\bar{x}(1)),$$

$$\text{supp}\{\mu\} \subset \{t \in [0,1] : h(\bar{x}(t)) = 0\},$$

and for almost every $t \in [0,1]$, $\bar{u}(t)$ *maximizes over* $\Omega(t)$

$$u \mapsto q(t) \cdot f(t, \bar{x}(t), u)$$

where

$$q(t) = \begin{cases} p(t) + \int_{[0,t)} h_x(\bar{x}(s))\mu(ds) & t \in [0,1) \\ p(1) + \int_{[0,1]} h_x(\bar{x}(s))\mu(ds) & t = 1. \end{cases}$$

8.2 Proof of Theorem 8.1.1

We assume that $h(x_0) = 0$, since otherwise the conditions of MP cannot be satisfied by the degenerate multipliers.

So, we can consider three cases:

Case 1: the minimizing state trajectory leaves the boundary immediately, i.e. there exists $r \in (0,1)$ such that $h(\bar{x}(t)) < 0, \forall t \in (0,r]$;

Case 2: the minimizing state trajectory remains in the boundary on a neighborhood of the initial time, i.e. there exists $r \in (0,1]$ such that $h(\bar{x}(t)) = 0, \forall t \in [0,r]$;

Case 3: Case 1 or Case 2 occurs a infinite numbers of times on neighborhood of the initial time, i.e. there exists a sequence $\{a_j\}$ such that $\{a_j\} \downarrow 0$ and $a_j \in [0,\epsilon]$, $\forall j \in \mathbb{N}$ where

$$\begin{cases} h(\bar{x}(t)) < 0 & \text{for all } t \in (a_{2j-1}, a_{2j}), \text{ all } j \geq 1 \\ h(\bar{x}(t)) = 0 & \text{for all } t \in (a_{2j}, a_{2j+1}), \text{ all } j \geq 1 \\ h(\bar{x}(a_j)) = 0 & \text{for all } j \geq 1. \end{cases}$$

Step 1: The validation of theorem in case 1 follows from the application of Proposition 7.3.1.

To proof the **Case 2** and **Case 3**, we will apply the Theorem 6.2.1 which is valid under the convexity hypotheses of $f(t, x, \Omega(t))$.

We start by observing that $x \to h(x)$ is continuous differentiable and $x(t)$ is an absolutely continuous function, then

$$\frac{d}{dt}h(x(t)) \text{ exists for a.e. } t \in [0,1].$$

Hence,

$$h_x(x(t)) \cdot f(t, x(t), u(t)) \text{ exists for a.e. } t \in [0,1].$$

Step 2: Next we prove the Theorem in the **Case 2**. Therefore, there exists $r \in (0,1]$ such that $h(\bar{x}(t)) = 0, \forall t \in [0,r]$.

For a.e. $t \in [0,r)$, we have

$$\frac{d}{dt}h(\bar{x}(t)) = 0 \Leftrightarrow$$

$$\Leftrightarrow h_x(\bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t)) = 0.$$

Since $\bar{u}$ is piecewise continuous on the left, $t \to h_x(\bar{x}(t))$ and $t \to f(t, \bar{x}(t), \bar{u}(t))$ are continuous on a neighborhood of the initial time, there exists r_0 sufficiently near of 0 and $r_0 \leq r$ such that for all $t \in [0, r_0]$, we have

$$|h_x(\bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t))| < \varepsilon,$$
$$\text{for a.e. } t \in [0, r_0).$$

Therefore,

$$|h_x(x_0) \cdot f(t, x_0, \bar{u}(t))| < \varepsilon, \text{ a.e. } t \in [0, r_0).$$

Choosing $\varepsilon = \frac{\delta}{2}$ and as $\mathbf{CQ_{EVI}}$ is satisfied, then for a.e $t \in [0, \min\{r_0, \epsilon_2)\}$, we have

$$\int_0^t h_x(x_0) \cdot [f(s, x_0, \tilde{u}(s)) - f(s, x_0, \bar{u}(s))]ds < -\frac{\delta}{2}t.$$

Therefore, we are in the conditions to apply Theorem 6.2.1 and the result holds immediately.

Step 3: Finally we treat case 3. In the last case, there exists sequence a_j such that $a_j \downarrow 0$ and

$$\begin{cases} h(\bar{x}(t)) < 0 & \text{for all } t \in (a_{2j-1}, a_{2j}), \text{ all } j \geq 1 \\ h(\bar{x}(t)) = 0 & \text{for all } t \in (a_{2j}, a_{2j+1}), \text{ all} j \geq 1 \\ h(\bar{x}(a_j)) = 0 & \text{for all } j \geq 1. \end{cases}$$

We first claim that there exists $\exists s_j \in]a_{2j-1}, a_{2j}[\, \forall j \in \mathbb{N}$:

$$h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) \geq 0. \tag{8.2}$$

Seeking a contradiction assume that for all $t \in]a_{2j-1}, a_{2j}[$

$$h_x(\bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t)) < 0, \ \forall t \in]a_{2j-1}, a_{2j}[.$$

Then

$$\int_t^{a_{2j}} h_x(\bar{x}(s)) \cdot f(s, \bar{x}(s), \bar{u}(s)) ds < 0$$

So

$$h(\bar{x}(a_{2j})) - h(\bar{x}(t)) < 0.$$

As $h(\bar{x}(a_{2j})) = 0$ and $h(\bar{x}(t)) \leq 0$, the contradiction obtained proves our claim.

By $\mathbf{CQ_{EVI}}$ and (8.2)

$$\int_0^t h_x(x_0) \cdot f(s, x_0, \tilde{u}(s)) - h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) ds < -\delta t, \ \forall t \in [0, \epsilon).$$

Therefore,

$$\int_0^t h_x(x_0) \cdot f(s, x_0, \tilde{u}(s)) - h_x(x_0) \cdot f(s, x_0, \bar{u}(s)) ds < -\delta t + \int_0^t h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(s, x_0, \bar{u}(s)) ds.$$

To finish our proof, we claim that $\exists \epsilon^* > 0$ such that a.e. $t \in [0, \epsilon^*]$ and $\forall \varepsilon > 0$

$$|h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t))| < \varepsilon.$$

Note that

$$|h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t))|$$

$$= |h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(0, x_0, \bar{u}(0)) + h_x(x_0) \cdot f(0, x_0, \bar{u}(0))|$$
$$-h_x(x_0) \cdot f(t, x_0, \bar{u}(t))|$$

$$\leq |h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(0, x_0, \bar{u}(0))|$$
$$+|h_x(x_0) \cdot f(0, x_0, \bar{u}(0)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t))|.$$

Since $\bar{u}$ is piecewise continuous on the left, $t \to h_x(\bar{x}(t))$ and $t \to f(t, \bar{x}(t), \bar{u}(t))$ are continuous on a neighborhood of the initial time and $s_j \downarrow 0$, there exist j_1 sufficient large such that $\forall j \geq j_1$,

$$|h_x(\bar{x}(s_j)) \cdot f(s_j, \bar{x}(s_j), \bar{u}(s_j)) - h_x(x_0) \cdot f(0, x_0, \bar{u}(0))| < \varepsilon, \quad \forall \varepsilon > 0.$$

By continuity of $t \to f(t, x_0, \bar{u}(t))$, we also conclude that there exist r_0 sufficient near of 0 such that for all $t \in [0, r_0)$

$$|h_x(x_0) \cdot f(0, x_0, \bar{u}(0)) - h_x(x_0) \cdot f(t, x_0, \bar{u}(t))| < \varepsilon, \quad \forall \varepsilon > 0.$$

Choosing $\varepsilon = \frac{\delta}{4}$, we can apply Theorem 6.2.1 for a.e. $t \in [0, \min\{s_{j_1}, r_0, \epsilon\})$ and the result holds immediately.

Chapter 9

Nondegeneracy in Problems with Higher Index State Constraints

In previous chapters, we have studied CQ that allow strengthened terms of the MP to avoid degeneracy. However, for OCP with state constraints that have higher index (i.e. their first derivative with respect to time does not depend on the control), most CQ described in literature are not adequate.

We note that control problems with higher index state constraints arise frequently in practice. An example, explored here, is a common mechanical systems for which there is a constraint on the position (an obstacle in the path, for example) and the control acts as a second derivative of the position (a force or acceleration) which is a typical case arising in the area of mobile robotics.

So, there is a need to develop new constraint qualifications, involving higher derivatives of the state constraint. The results presented here are a generalization of [Fon05], to cover nonlinear problems.

9.1 Introduction

Consider, again, the problem (OCP_3).

$$(OCP_3) \quad \begin{cases} \text{Minimize} & g(x(1)) & \\ \text{subject to} & \dot{x}(t) = f\left(t, x(t), u(t)\right) & \text{a.e.} t \in [0,1] \\ & x(0) = x_0 & \\ & u(t) \in \Omega(t) & \text{a.e.} t \in [0,1] \\ & h\left(x(t)\right) \leq 0 & \text{for all } t \in [0,1]. \end{cases}$$

As we mentioned before, the MP, under the basic hypotheses, for the problem (OCP_3) asserts the existence of an absolutely continuous function $p : [0,1] \to \mathbb{R}^n$, a nonnegative measure $\mu \in C^*([0,1];\mathbb{R})$ and a scalar $\lambda \geq 0$ such that

$$\mu\{[0,1]\} + \lambda > 0, \tag{9.1}$$

$$-\dot{p}(t) \in \text{co}\partial_x^L \left(q(t) \cdot f(t, \bar{x}(t), \bar{u}(t))\right) \text{ a.e. } t \in [0,1], \tag{9.2}$$

$$q(1) \in \lambda \partial^L g(\bar{x}(1)), \tag{9.3}$$

$$\text{supp}\{\mu\} \subset \{t \in [0,1] : h(\bar{x}(t)) = 0\}, \tag{9.4}$$

and for almost every $t \in [0,1]$, $\bar{u}(t)$ maximizes over $\Omega(t)$

$$u \mapsto q(t) \cdot f(t, \bar{x}(t), u) \tag{9.5}$$

where

$$q(t) = \begin{cases} p(t) + \displaystyle\int_{[0,t)} h_x(\bar{x}(s))\mu(ds) & t \in [0,1) \\ p(1) + \displaystyle\int_{[0,1]} h_x(\bar{x}(s))\mu(ds) & t = 1. \end{cases}$$

As we have seen in section 3.2.1, the CQ to avoid the degeneracy are typically of two types:

$\mathbf{CQ1_d}$ $\exists\delta, \epsilon > 0$ and $\exists\tilde{u}(t) \in \Omega(t)$:

$$h_x(x_0) \cdot [f(t, x_0, \tilde{u}(t)) - f(t, x_0, \bar{u}(t)] < -\delta \quad \text{a.e.} t \in [0, \epsilon).$$

$\mathbf{CQ2_d}$ $\exists\delta, \epsilon > 0$ and $\exists\tilde{u}(t) \in \Omega(t)$:

$$h_x(x_0) \cdot f(t, x_0, \tilde{u}(t)) < -\delta \quad \text{a.e.} t \in [0, \epsilon).$$

There are, however, some problems with interest in practice for which the constraint qualifications **$\mathbf{CQ1_d}$** and **$\mathbf{CQ2_d}$** are useless to select a set of problems in which the MP can be strengthened. These problems are known as OCP with *higher index of the state constraint.*

9.2 Higher Index

We define the index of a state constraint as a measure of how many times we have to differentiate the state constraint to have an explicit dependence on the control.

Definition 9.2.1 *(Index of the State Constraint)*

Let $h(x(\cdot))$ be $k+1$ times continuous differentiable and

$$h^{(j)}(x(t)) = \left(\frac{d}{dt}\right)^j h(x(t)).$$

The state constraint is said to have index k, if k is a non-negative integer such that

$$\frac{\partial}{\partial u}(h_x^j(x) \cdot f(t, x, u)) = 0, \quad j = 0, ..., k-1$$

$$\frac{\partial}{\partial u}(h_x^k(x) \cdot f(t, x, u)) \neq 0.$$

If $\dfrac{\partial}{\partial u}(h_x^j(x) \cdot f(t, x, u)) = 0$ for all $j \geq 0$, the state constraint is said to have index $k = \infty$.

In particulary, consider a linear optimal control problems like (OCP_L).

$$
(OCP_L) \quad \begin{cases} \text{Minimize} & \displaystyle\int_0^1 L(x(t), u(t))dt + W(x(1)) & \\ \text{subject to} & \dot{x}(t) = Ax(t) + bu(t) & \text{a.e.} t \in [0,1] \\ & x(0) = x_0 & \\ & u(t) \in \Omega(t) & \text{a.e.} t \in [0,1] \\ & c^T x(t) \leq d & \forall\, t \in [0,1], \end{cases}
$$

Assuming that (OCP_L) have index $k > 3$.

Since $h(x(t)) = c^T x(t) - d$ then

$$h^1(x(t)) = h_x(x) \cdot f(t, x, u) = c^T(Ax(t) + bu(t)).$$

By definition of higher index state constraints, then $c^T b = 0$. And consequently,

$$h^2(x(t)) = h^1_x(x) \cdot f(t, x, u) = c^T A(Ax(t) + bu(t)).$$

By definition of higher index state constraints, then $c^T Ab = 0$. And consequently,

$$h^3(x(t)) = h^2_x(x) \cdot f(t, x, u) = c^T A^2(Ax(t) + bu(t)).$$

Again by definition of higher index state constraints, then $c^T A^2 b = 0$.

By induction, we conclude that for a problem like (OCP_L) the state constraint is said to have index k, if k is a non-negative integer such that

$$
\begin{aligned}
&c^T A^j b = 0, \quad j = 0, 1, ..., k-1 \\
&c^T A^k b \neq 0.
\end{aligned}
$$

As we have said, control problems with higher index state constraints arise frequently in mechanical systems, when there is a constraint on the position and the control acts as a second derivative of the position. This is illustrated in the following example:

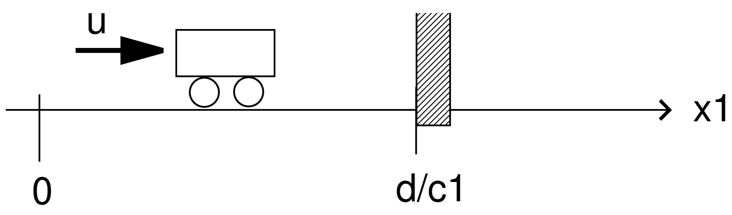

Figure 9.1: A higher index constrained system (from [Fon05]).

Example 9.2.2 *Consider a second order linear system modelling a mass* $(1/b)$ *moving along a line by action of a force* (u) *and in which the position* (x_1) *is constrained to a certain half-space* $(\leq d/c_1)$*. (see Figure 9.1).*

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ b \end{bmatrix} u(t),$$

$$[c_1, 0]x(t) - d \leq 0.$$

We note that the quantity

$$h^{(1)}(x(t)) = h_x(x(t)) \cdot [f\,(t, x(t), u(t))] = [0, c_1]x(t)$$

does not depend explicitly on the control. Therefore, the index is greater than one.

Having introduce the definition of higher index, we now show why the previous CQ's are not adequate for problems with higher index state constraints. We start by showing that $\mathbf{CQ1_d}$ is not satisfied by this type of problems. Assume that the problem has index great then zero, then by definition of index, we have

$$\frac{\partial}{\partial u}(h_x(x) \cdot f(t, x, u)) = 0.$$

That means that the quantity $h_x(x) \cdot f(t, x, u)$ do not depend explicitly of u and therefore,

$$h_x(x) \cdot [f(t, x, \tilde{u}) - f(t, x, \bar{u})] = 0.$$

So,

$$h_x(x_0)\cdot[f(t,x_0,\tilde{u}(t))-f(t,x_0,\bar{u}(t))]$$
$$=h_x(x_0)\cdot[f(t,x_0,\tilde{u}(t))-f(t,x_0,\bar{u}(t))]-h_x(x(t))\cdot[f(t,x,\tilde{u}(t))-f(t,x,\bar{u}(t))]\,.$$

Since $h^1(\cdot)$ is continuous, then for t sufficiently near of 0 we have: $\forall\varepsilon>0$

$$|h_x(x(t))\cdot[f(t,x,\tilde{u}(t))-f(t,x,\bar{u}(t))]-h_x(x_0)\cdot[f(t,x_0,\tilde{u}(t))-f(t,x_0,\bar{u}(t))]\,|<\varepsilon,$$

we conclude that $\mathbf{CQ1_d}$ is never satisfied.

Now, we suppose that $\mathbf{CQ2_d}$ is satisfied. By definition of index, we have

$$h_x(x_0)\cdot f(t,x_0,\tilde{u}(t))=h_x(x_0)\cdot f(t,x_0,\bar{u}(t))<-\delta,$$

for all $t\in[0,\epsilon)$.

On other hand, by continuity of $h^1(\cdot)$, we conclude that there exists ϵ' sufficient near of 0 and $\epsilon'\leq\epsilon$ such that for all $t\in[0,\epsilon']$

$$h_x(\bar{x}(t))\cdot f(t,\bar{x}(t),\bar{u}(t))<-\delta'.$$

Therefore

$$h^{(1)}(\bar{x}(t))<-\delta', \tag{9.6}$$

for all $t\in[0,\epsilon']$.

That means that the initial part of the optimal trajectory leaves the boundary for a period of time.

We can conclude that, if the problem has index great than one, $\mathbf{CQ2_d}$ is satisfied for a particular kind of problems, problems in which the optimal trajectory leaves the boundary for a period of time.

Since we do not know in advance the behavior of the minimizer trajectory, we would have to assume that all admissible trajectories satisfy the inequality (9.6). However, for this kind of problems, the nontriviality condition can be replace by

$\mu\{(0,1]\} + \lambda > 0$, see [FV94]. Therefore, the constraint qualification $\mathbf{CQ2_d}$ looses interest.

In order to remedy this problem, new CQ dependent on the *index of the state constraint* are developed.

Throughout this chapter, we are assuming that the problem have index k.

In [Fon05], linear optimal control problems like (OCP_L) were considered.

The constraint qualification that guarantee the nondegeneracy is the following:

$\mathbf{CQ_{Fon05}}$ $\exists \delta > 0, \epsilon > 0$ and a control $\tilde{u} \in \Omega(t)$ such that

$$c^T A^k b(\tilde{u}(t) - \bar{u}(t)) < -\delta$$

for all $t \in [0, \epsilon)$.

Here, we generalize this result to cover nonlinear OCP.

9.3 Main Results

Assuming that, there exists a $\delta' > 0$, such that

$\mathbf{H1_{HI}}$ The function $(t,u) \to f(t,x,u)$ is $\mathcal{L} \times \mathcal{B}^m$ measurable for each x.

$\mathbf{H2_{HI}}$ The function $x \to f(t,x,u)$ is Lipschitz continuous with a Lipschitz constant K_f, for all $u \in \Omega(t)$ a.e. $t \in [0,1]$;

$\mathbf{H3_{HI}}$ The function g is locally Lipschitz continuous;

$\mathbf{H4_{HI}}$ The $Gr\,\Omega$ is $\mathcal{L} \times \mathcal{B}^m$ measurable.

For technical reasons, the main result must assume that an initial part of the optimal trajectory does not enter and leave the boundary of the state constraint an infinite number of times. That is, the initial part of the optimal trajectory either stays on the boundary of the state constraint for some time or leaves the boundary immediately.

Assumption 1: Either

Case 1: $\exists\tau \in (0,1)$ such that $h(\bar{x}(t)) < 0$ for all $t \in (0,\tau]$,

or

Case 2: $\exists\tau \in (0,1)$ such that $h(\bar{x}(t)) = 0$ for all $t \in [0,\tau]$

Additionally, assume that one of both CQ are satisfied

$\mathbf{CQ_{HI}}$

Let the state constraint have index k, and the function $x \to h^{(k)}(x)$ be continuously differentiable. If $h(x_0) = 0$, then there exist positive constants K_u, ϵ, δ and a control $\tilde{u} \in \Omega(t)$ such that for a.e. $t \in [0,\epsilon)$

$$\|f(t,x_0,\tilde{u}(t))\| \leq K_u, \qquad \|f(t,x_0,\bar{u}(t))\| \leq K_u$$

and

$$h_x^{(k)}(x_0) \cdot [f(t,x_0,\tilde{u}(t)) - f(t,x_0,\bar{u}(t))] < -\delta.$$

$\mathbf{CQ_{EHI}}$ Let the state constraint have index k, and the function $x \to h^{(k)}(x)$ be continuously differentiable. If $h(x_0) = 0$, then there exist positive constants K_u, ϵ, δ and a control $\tilde{u} \in \Omega(t)$ such that for a.e. $t \in [0,\epsilon)$

$$\|f(t,x_0,\tilde{u}(t))\| \leq K_u, \qquad \|f(t,x_0,\bar{u}(t))\| \leq K_u$$

and

$$h_x^{(k)}(x_0) \cdot f(t,x_0,\tilde{u}(t)) < -\delta. \tag{9.7}$$

Theorem 9.3.1 *Let $(\bar{x},\bar{u})$ be a local minimizer for (OCP_3). Assume that hypotheses* $\mathbf{H1_{HI}}$-$\mathbf{H4_{HI}}$*, Assumption 1 together with* $\mathbf{CQ_{HI}}$ *are satisfied. Then, the NCO* *(equations (9.1) to (9.5)) can be strengthened with the condition*

$$\mu\{(0,1]\} + \|q\|_{L^\infty} + \lambda > 0.$$

Theorem 9.3.2 *Let $(\bar{x}, \bar{u})$ be a local minimizer for (OCP_3). Assume that hypotheses* $\mathbf{H1_{HI}}$-$\mathbf{H4_{HI}}$*, Assumption 1 together with* $\mathbf{CQ_{EHI}}$ *are satisfied. Then, the NCO (equations (9.1) to (9.5)) can be strengthened with the condition*

$$\mu\{(0,1]\} + \|q\|_{L^\infty} + \lambda > 0.$$

Remark 9.3.3 *In Theorem 9.3.1 we generalize the result of [Fon05] to cover non-linear OCP. In Theorem 9.3.2 we strengthen the MP, by means of a CQ that do not involves the minimizing $\bar{u}$, and therefore is easier to verify.*

9.4 Proof of Main Results

We will consider separately the **cases 1** and **2** in Assumption 1.

In **Case 1**, we are in the condition to apply directly Proposition 2.2 of [FV94], under weaker hypotheses and the result holds.

In **Case 2**, we by observe that $h^{(i)}(x)$ can be determined recursively by

$$\begin{cases} h^{(i)}(x) = h_x^{(i-1)}(x) \cdot f(t, x, u), \\ h^{(0)}(x) = h(x). \end{cases}$$

Note that: $h^{(i)}(x) = \dfrac{d}{dt} h^{(i-1)}(x(t)) = h_x^{(i-1)}(x) \dfrac{d}{dt} x(t) = h_x^{(i-1)}(x) \cdot f(t, x, u)$.

Step 1: We prove the following lemma.

Lemma 9.4.1 *If* $\mathbf{CQ_{EHI}}$ *is satisfied and the initial part of the optimal trajectory stays on the boundary of the state constraint for some time, then* $\mathbf{CQ_{HI}}$ *is satisfied.*

Proof.

Since the initial part of the optimal trajectory stays on the boundary of the state constraint for some time, then exists a positive scalar τ such that $h(\bar{x}(t)) = 0, \forall t \in [0, \tau]$. Therefore

$$h^{(k+1)}(\bar{x}(t)) = 0, \text{ for all } t \in [0, \tau).$$

Recursively, we conclude that

$$h_x^{(k)}(\bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t)) = 0, \text{ for all } t \in [0, \tau).$$

On other hand, we have

$$\begin{aligned}
&|h_x^{(k)}(\bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t)) - h_x^{(k)}(x_0) \cdot f(t, x_0, \bar{u}(t))| \\
&\quad = |h_x^{(k)}(\bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t)) - h_x^{(k)}(x_0) \cdot f(t, \bar{x}(t), \bar{u}(t))| \\
&\quad +|h_x^{(k)}(x_0) \cdot f(t, \bar{x}(t), \bar{u}(t)) - h_x(x_0)^{(k)} \cdot f(t, x_0, \bar{u}(t))| \\
&\quad \leq \|h_x^{(k)}(\bar{x}(t)) - h_x^{(k)}(x_0)\| \|f(t, \bar{x}(t), \bar{u}(t))\| + \|h_x(x_0)\| \|f(t, \bar{x}(t), \bar{u}(t)) - f(t, x_0, \bar{u}(t))\|.
\end{aligned}$$

Since $h_x^{(k)}(\cdot)$ and $\bar{x}(\cdot)$ are continuous functions, then for any $\varepsilon_1 > 0$, there exists r_1 sufficient near of 0 and $r_1 \leq r$ such that $\|h_x^{(k)}(\bar{x}(t)) - h_x^{(k)}(x_0)\| \leq \varepsilon_1$. Therefore

$$\begin{aligned}
&\|h_x^{(k)}(\bar{x}(t)) - h_x^{(k)}(x_0)\| \|f(t, \bar{x}(t), \bar{u}(t))\| + \|h_x^{(k)}(x_0)\| \|f(t, \bar{x}(t), \bar{u}(t)) - f(t, x_0, \bar{u}(t))\| \\
&\leq \varepsilon_1(\|f(t, \bar{x}(t), \bar{u}(t)) - f(t, x_0, \bar{u}(t))\| + \|f(t, x_0, \bar{u}(t))\|) + \|h_x^{(k)}(x_0)\| K_f \|\bar{x}(t) - x_0\| \\
&\leq \varepsilon_1(K_f \|\bar{x}(t) - x_0\| + K_u) + \|h_x^{(k)}(x_0)\| K_f \|\bar{x}(t) - x_0\|.
\end{aligned}$$

Again by continuity of $\bar{x}(\cdot)$, for any $\varepsilon_2 > 0$, there exists r_2 sufficient near of 0 and $r_2 \leq r$ such that $\|\bar{x}(t) - x_0\| \leq \varepsilon_2$. Therefore

$$\begin{aligned}
&\varepsilon_1(K_f \|\bar{x}(t) - x_0\| + K_u) + \|h_x^{(k)}(x_0)\| K_f \|\bar{x}(t) - x_0\| \\
&\leq \varepsilon_1(K_f \varepsilon_2 + K_u) + \|h_x^{(k)}(x_0)\| K_f \varepsilon_2.
\end{aligned}$$

Choosing an appropriate ε_1 and ε_2, we can conclude that for any $\varepsilon > 0$

$$\begin{aligned}
&|h_x^{(k)}(\bar{x}(t)) \cdot f(t, \bar{x}(t), \bar{u}(t)) - h_x^{(k)}(x_0) \cdot f(t, x_0, \bar{u}(t))| < \varepsilon, \\
&\qquad\qquad \text{for a.e. } t \in [0, r_0 = \min\{r_1, r_2\}).
\end{aligned}$$

Therefore,

$$|h_x^{(k)}(x_0) \cdot f(t, x_0, \bar{u}(t))| < \varepsilon, \text{ for a.e. } t \in [0, \epsilon'].$$

Choosing $\varepsilon = \frac{\delta}{2}$ and since (9.7) is satisfied the results holds. ■

Step 2: We distinguish the cases when **k=0**, when **k=∞**, and when **k is positive and finite**.

If **k=0**, then the state constraint is not of higher index, by the lemma above and Theorem 2.1 in ([FFV99]) the results holds.

If **k=∞**, the process minimizer $(\bar{x}, \bar{u})$ remains a minimizer when the state constraint is dropped from the problem specification.

To see this, we can write

$$\begin{aligned} &h(x(t)) - h(\bar{x}(t)) = \\ &h(x(0)) - h(\bar{x}(0)) + \sum_{i=1}^{+\infty} \frac{t^i}{i!} \left[h^{(i)}(x(t)) - h^{(i)}(\bar{x}(t)) \right]_{t=0} \end{aligned}$$

and

$$h^{(i)}(x(t)) = h_x^{(i-1)}(x(t)) \cdot f(t, x(t), u(t)).$$

We conclude that

$$\begin{aligned} &h(x(t)) - h(\bar{x}(t)) = \\ &\sum_{i=1}^{+\infty} \frac{t^i}{i!} h_x^{(i-1)}(x_0) \cdot \left[f(0, x_0, u(0)) - f(0, x_0, \bar{u}(0)) \right]. \end{aligned}$$

By the fact of $k = \infty$, then

$$h(x(t)) = h(\bar{x}(t)), \text{ for all absolutly continuous function } x.$$

So the state constraint does not depend on the trajectory, and therefore the state constraint can be ignored.

Suppose that k **is positive and finite**.

Since exists a positive scalar τ such that $h(\bar{x}(t)) = 0, \forall t \in [0, \tau]$, we have

$$h^{(k)}(\bar{x}(t)) = 0, \text{ for all } t \in [0, \tau).$$

Therefore, the minimizer $(\bar{x}, \bar{u})$ for problem (OCP_4) is also a minimizer for the same problem with the additional constraint

$$h^k(\bar{x}(0)) = 0.$$

We can rewrite the new state constraint(s) of the problem as

$$\tilde{h}(t,x) = \begin{cases} \max\{h(x), h^{(k)}(x)\} & \text{if } t = 0 \\ h(x) & \text{if } t > 0. \end{cases}$$

This function is upper semi-continuous and the nondegenerate NCO in [FFV99] apply to this problem provided the following CQ is satisfied:

If $\tilde{h}(0, x_0) = 0$, then there exists positive constants δ and ϵ, and a control value $\tilde{u}$ such that

$$\xi \cdot [f(t, x_0, \tilde{u}(t)) - f(t, x_0, \bar{u}(t))] < -\delta$$

for all $\xi \in \partial_x^{>} \tilde{h}(s, x), s \in (0, \epsilon)$.

Knowing that (see [Cla83])

$$\xi \in \{(\alpha h_x(x_0) + (1 - \alpha) h_x^{(k)}(x_0)) : \alpha \in [0, 1]\}.$$

We have

$$\begin{aligned} &\alpha h_x(x_0) \cdot [f(t, x_0, \tilde{u}(t)) - f(t, x_0, \bar{u}(t))] \\ &+ (1 - \alpha) h_x^{(k)}(x_0) \cdot [f(t, x_0, \tilde{u}(t)) - f(t, x_0, \bar{u}(t))] < -\delta, \end{aligned}$$

provided that

$$h_x^{(k)}(x_0) \cdot [f(t, x_0, \tilde{u}) - f(t, x_0, \bar{u}(t))] < -\delta'. \tag{9.8}$$

Therefore, if $\mathbf{CQ_{HI}}$ or $\mathbf{CQ_{EHI}}$ is satisfied, then the CQ in [FFV99] is satisfied with $\tilde{h}$ and the corresponding NCO can be applied, yielding the result.

Notes on Chapter

Part of the contents have been introduced in [LF08a] and [LF08b].

Chapter 10

Conclusion

In this chapter, we summarize the main contributions of this thesis and we also suggests some future developments.

10.1 Contributions

The main contribution of this thesis is the development of nondegenerate necessary conditions of optimality for a Mayer problems in order to avoid a particular kind of degeneracy (when the initial state belongs to the boundary) or ensure the normality.

The results here developed improve on the existent literature in the sense that they address problems with less restrictions on its data and they are valid under constraints qualifications that are verified for more problems or are easier to verify whether they are satisfied, as it is described below.

The normality result developed in Chapter 5 improves on the results existent in the literature by the fact that it is valid under weaker nonsmooth hypotheses: the velocity set is merely required to be $k(\cdot)$-Lipschitz with respect to x (where $k(\cdot)$ is an integrable function) and the continuity on u is not required.

In Chapter 6, a strengthened Maximum Principle to avoid the degeneracy phenomenon was developed under a new type of constraint qualification that we have called "integral-type of constraint qualification". This constraint qualification applies to a larger class of problems than the constraints qualifications introduced in

[FFV99]. However, these constraint qualifications involve the optimal control which we do not know in advance and consequently are not directly verify, excepted in special case as calculus of variations problems.

In order to remedy this problem a nondegenerate result, valid under constraint qualifications that do not involve the optimal control was developed in Chapter 7 and Chapter 8.

Nondegenerate results involving constraint qualification as the ones introduced in Chapter 7 already exits in literature. The novelty is that $h(\cdot)$ has just to be continuously differentiable and not of class $C^{1,1}$. Also, the constraints qualification just has to be satisfy along the optimal trajectory.

The constraint qualification, that was developed in chapter 8, is a of integral-type. This type of constraint qualification has the advantage of it applies to a larger class of problems.

Since most constraints qualifications described in the literature are not adequate for optimal control problems with state constraints that have higher index, a nondegenerate result valid under constraint qualifications involving higher derivatives of the state constraint was developed in Chapter 9. This result generalizes the result in [Fon05], by allowing nonlinear problems.

The results of chapter 6 to chapter 8 together with the result developed in [FFV99] are strictly connected, as we can seen in Figure 10.1. (In the Figure 10.1 the symbol $\Rightarrow$ means "imply" and the hypotheses under the constraint qualification are additionally hypotheses concerning the nondegenerate result in [FFV99].)

10.2 Future works

The research described here naturally leads on to several open questions and suggests some future developments.

There is a perspective of development of nondegenerate necessary conditions of optimality subject to weaker hypotheses. In particular, it is desirable to remove the convexity hypothesis in Chapter 6.

CQ-EV (*Chapter* 7) $\Longrightarrow$ **CQ-EVI** (*Chapter* 8)

$\Big\Downarrow$ no terminal constraints; $h(\cdot)$ continuously differentiable; $\bar{u}$ is piecewise continuous on the left; $f(\cdot, x, \cdot)$ is continuous (left column)

$\Big\Downarrow$ no terminal constraints; $h(\cdot)$ continuously differentiable; $\bar{u}$ is piecewise continuous on the left; $f(\cdot, x, \cdot)$ is continuous (right column)

CQ-FFV99 $\Longrightarrow$ **CQ-I** (*Chapter* 6)

$\Big\Downarrow$ (left column)

$\Big\Downarrow$ convexity of $f(t, x, \Omega(t))$ near x_0 (right column)

Nondegenerate Result **Nondegenerate Result**

Figure 10.1: The connection between the results of Chapters 6 to 8.

More generally, constraints can be considered, for example the initial and final state belonging to a given set and/or mixed state constraints.

Another perspective is to strengthen the nondegenerate results to avoid other type of degeneracy (not only for the left endpoint), or to ensure normality. Also, the developed higher order conditions for the case that the nondegenerate first order necessary conditions do not provided enough information, is also a field that can be explored.

As it was shown in this thesis, normality and regularity are strictly connected. Therefore, development of regularity results is a suggestion to future work.

In particulary, the nondegenerate result involving problems with higher index constraints is valid for optimal trajectories that leave the boundary immediately or belongs to the boundary for a period of time. We wish consider the case where the optimal trajectory touches the boundary on an infinite number of times or proof that for problems with higher index constraints this case does not occurs.

Appendix

Algebra

Definition 10.2.1 *Let $\mathcal{M}$ be a collection of subsets of a set Ω. Then $\mathcal{M}$ is called a **algebra** (the term field is also used) iff $\Omega \in \mathcal{M}$ and $\mathcal{M}$ is closed under complementation and finite union, that is*

(i) $\Omega \in \mathcal{M}$;

(ii) *If $A \in \mathcal{M}$, then $A^c \in \mathcal{M}$;*

(iii) *If A_1, A_2, ..., $A_n \in \mathcal{M}$, then $\bigcup_{i=1}^{n} A_i \in \mathcal{M}$.*

De Morgan's Laws immediately show that an algebra must satisfy other properties: for A_1, A_2, ..., $A_n \in \mathcal{M}$, then

$$\bigcap_{i=1}^{n} A_i = (\bigcup_{i=1}^{n} A_i^c)^c \in \mathcal{M}$$

and for all i, j $\in \{1, 2, ..n\}$

$$A_i - A_j \in \mathcal{M}.$$

Definition 10.2.2 *Let $\mathcal{M}$ be an algebra then $\mathcal{M}$ is a σ**-algebra** if also*

$$A_1, A_2, A_3... \in \mathcal{M}, \text{ then } \bigcup_{i=1}^{\infty} A_i \in \mathcal{M}.$$

Again, it follows immediately from De Morgan's Laws that $\bigcap_{i=1}^{\infty} A_i \in \mathcal{M}$, as well. The prefix σ is used to signify that "countable sums" of sets in $\mathcal{M}$ are also in

$\mathcal{M}$. Thus, in a σ-algebra all the standard set operations can be performed countably many times on sets in $\mathcal{M}$.

Functional Analysis

Topological Space

Definition 10.2.3 *A collection τ of subsets of a set Ω is said to be a* ***topology*** *in Ω if τ has the following three properties:*

(i) $\emptyset \in \tau$ *and* $\Omega \in \tau$;

(ii) *If* $V_i \in \tau$ *for* $i = 1, ..., n$, *then* $V_1 \cap V_2 \cap ... \cap V_n \in \tau$;

(iii) *If* $\{V_\alpha\}$ *is an arbitrary collection of members of τ (finite, countable or uncountable), then* $\bigcup_\alpha V_\alpha \in \tau$.

The pair (Ω, τ) is called a ***topological space****, but if τ is understood, we refer to Ω as a topological space. The sets in τ are called the open sets of (Ω, τ).*

Definition 10.2.4 *If Ω and Γ are topological spaces and if f is a mapping of Ω into Γ, then f is said to be* ***continuous*** *provided that $f^{-1}(\Gamma)$ is an open set in Ω for every open set V in Γ.*

Normal Space

Definition 10.2.5 *A topological space Ω is a* ***normal space*** *if it has the following properties:*

- *Sets consisting of single points are closed;*
- *For every pair of disjoint closed sets A and B, there are disjoints neighborhoods of A and B.*

Metric Space

Definition 10.2.6 *A* ***metric*** *on a set* X *is a function* $d : X \times X \to \mathbb{R}$ *that satisfy the following conditions:*

(i) $d(x, y) \geq 0$ *for all* $x, y \in X$*;*

(ii) $d(x, y) = 0$ *if and only if* $x = y$*;*

(iii) $d(x, y) = d(y, x)$ *for all* $x, y \in X$*;*

(iv) $d(x, z) = d(x, y) + d(y, z)$ *for all* $x, y, z \in X$*.*

A ***metric space*** *is a 2-tuple* (X, d) *where* X *is a set and* d *is a metric on* X*.*

Theorem 10.2.7 *[DS88] A metric space is normal.*

Definition 10.2.8 *A sequence* $\{a_n\}$ *in a topological space is said to* ***converge*** *to a point* a *in the space if every neighborhood of* a *contains all but a finite number of points* a_n*. This notation is written symbolically* $a_n \to a$*, or* $\lim_{n\to\infty} a_n = a$*. A sequence* $\{a_n\}$ *is said to be* ***convergent*** *if* $a_n \to a$ *for some* a*. A sequence* $\{a_n\}$ *in a metric space is a* ***Cauchy sequence*** *if* $\lim_{n,m} d(a_m, a_n) = 0$*. If every Cauchy sequence is convergent, a metric space is said to be* ***complete****.*

Lemma 10.2.9 *[DS88] In a metric space, a convergent sequence is a Cauchy sequence. A Cauchy sequence converges if and only if it has a convergent sequence.*

Examples: The real numbers with the function $d(x, y) = |y - x|$ given by the absolute value, and more generally Euclidean n-space with the Euclidean distance, are complete metric spaces.

Linear Space

Definition 10.2.10 *A set* X *is a* ***linear space*** *if the operations of addition and scalar multiplication are defined and if* X *is closed under these operations, that is for any pair of elements* $x, y \in X$*, and for any of scalars* α*,* β*, the element* $\alpha x + \beta y$ *is again in* X*.*

Definition 10.2.11 *A transformation T mapping a vector space X into a vector space Y is said to be* ***linear*** *if for every x_1, $x_2 \in X$ and all scalars α_1, α_2 we have $T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2)$.*

Definition 10.2.12 *A transformation from a vector space X into the space of real (or complex) scalars is said to be a* ***functional*** *on X.*

Normed Space

Definition 10.2.13 *A linear space X is a* ***normed linear space****, or a* ***normed space****, if to each $x \in X$ corresponds a real number $\|x\|$ called the* ***normed*** *of x which satisfies the conditions:*

(i) $\|0\| = 0$; $\|x\| > 0$, $x \neq 0$;

(ii) $\|x + y\| \leq \|x\| + \|y\|$, $x, y \in X$;

(iii) $\|\alpha x\| = |\alpha|\|x\|$, $x \in X$.

The properties (i), (ii), and (iii) show that d, defined by $d(x, y) = \|x - y\|$, is a metric in X. The metric topology in a normed linear space is sometimes called its **norm** or **strong** topology.

Definition 10.2.14 *A transformation T mapping a normed space X into a normed space Y is* ***continuous*** *at $x_0 \in X$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that $\|x - x_0\| < \delta$ implies that $\|T(x) - T(x_0)\| < \epsilon$. If T is continuous at each point $x_0 \in X$, we say that T is continuous.*

Definition 10.2.15 *A transformation T mapping a normed space X into a normed space Y is* ***uniformly continuous*** *on $X' \subset X$, if for every $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x, y \in X'$ with $\|x - y\|_X < \delta$ we have $\|f(x) - f(y)\| < \epsilon$.*

Proposition 10.2.16 *A transformation T mapping a normed space X into a normed space Y is continuous at the point $x_0 \in X$ if and only if $x_n \to x_0$ implies $T(x_n) \to T(x_0)$.*

Proposition 10.2.17 *[Lue69] If a linear functional on a normed space X is continuous at a single point, it is continuous throughout X.*

Definition 10.2.18 *A linear functional l on a normed space is **bounded** if there is a constant M such that $|f(x)| \leq M\|x\|$ for all $x \in X$.*

$$\|f\| = \inf\{M : |f(x)| \leq M\|x\|, \text{ for all } x \in X\}.$$

Proposition 10.2.19 *[Lue69] A linear functional on a normed space is bounded if and only if it is continuous.*

Theorem 10.2.20 *[Lue69] **Riesz Representation Theorem** Let f be a bounded linear functional on $X = C[a, b]$. Then there is a function v of bounded variation on $[a, b]$ such that for all $x \in X$*

$$f(x) = \int_a^b x(t)d\nu(t)$$

and such that for the norm of f is the total variation of ν on $[a, b]$. Conversely every function of bounded variation on $[a, b]$ defines a bounded linear functional on X in this way.

Definition 10.2.21 *Let X be a normed linear vector space. The space of all bounded linear functionals on X is called **normed dual** of X and is denoted X^*. The norm of an element $f \in X^*$ is*

$$\|f\| = \sup_{\|x\| \leq 1} |f(x)|.$$

Let $x^* \in X^*$. We often employ the notation $< x, x^* >$ for the value function x^* at the point $x \in X$.

Definition 10.2.22 *A sequence $\{x_n\}$ in a normed linear vector space X is said to **converge weakly** to $x \in X$, if for every $x^* \in X^*$ we have $< x_n, x^* > \to < x, x^* >$. In this case we write $x_n \to x$ weakly.*

Proposition 10.2.23 *[Lue69] If $x_n \to x$ strongly, then $x_n \to x$ weakly.*

Definition 10.2.24 *A sequence $\{x_n^*\}$ in X^* is said* ***converge weak**** *to the element x^* if for every $x \in X$, $< x, x_n^* > \to < x, x^* >$. In this case we write $x_n^* \to x^*$ weak*.*

Note: Strong implies weak and weak implies weak*.

Banach Space

A **Banach Space** is a normed linear space which is complete in its norm topology.

Example:

- Euclidean spaces $\mathbb{R}^n$ with Euclidean norm $\|x\| = (\sum_{i=1}^{n} x_i^2)^{1/2}$;
- $C^k([0,1] : \mathbb{R}^n)$ are a Banach space. ($C^k([0,1] : \mathbb{R}^n)$ denote the space of all k-times continuous differentiable functions from $[0,1]$ to $\mathbb{R}^n$, where $k = 0, 1, ..., \infty$).

Continuity in $\mathbb{R}^m$

Definition 10.2.25 *A function $f : X \to \mathbb{R}^m$ is* ***Lipschitz Continuous*** *on $A \subset \mathbb{R}^n$ if there is some nonnegative scalar K satisfying*

$$\|f(x) - f(y)\| \leq K\|x - y\| \text{ for all } x, y \in A.$$

Rademacher's theorem states that a Lipschitz continuous map $f : I \to \mathbb{R}$, where I is an interval in $\mathbb{R}$, is almost everywhere differentiable (that is, it is differentiable everywhere except on a set of Lebesgue measure 0). If K is the Lipschitz constant of f, then $|\dot{f}(x)| \leq K$ whenever the derivative exists. Conversely, if $f : I \to R$ is a differentiable map with bounded derivative, $|\dot{f}(x)| \leq L$ for all $x \in I$, then f is Lipschitz continuous with Lipschitz constant $K = L$.

Let C be a nonempty closed subset of $\mathbb{R}^n$. A Lipschitz continuous function related to C it is **distance function** d_C, defined by

$$d_C(x) = \min\{\|x - c\| : c \in C\}.$$

In [Cla83], we can find the proof of the following inequality

$$|d_C(x) - d_C(y)| \le \|x - y\|.$$

Definition 10.2.26 *A real-valued function f defined on a real interval $I = [a, b]$ is said to be* ***absolutely continuous*** *if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$\sum_{i=1}^{n} |f(b_i) - f(a_i)| < \varepsilon$$

for every finite collection of disjoints (a_i, b_i) subintervals of I with $\sum_{i=1}^{n} |b_i - a_i| < \delta$. The space $W^{1,1}(I : \mathbb{R})$ is defined for an intervals I and consists of all absolutely continuous functions on I.

Namely if f is absolutely continuous, it is continuous, it is a function of bounded variation and it is differentiable almost everywhere.

A Lipschitz continuous function is absolutely continuous, but the inverse is not necessarily true.

Definition 10.2.27 *A function $f : A \to \mathbb{R}$ is* ***lower semicontinuous*** *at a point x of $A \subset \mathbb{R}^n$ if*

$$f(x) \le \lim_{i\to\infty} f(x_i)$$

for every sequence x_1, x_2, ..., in A such that x_i converges to x and the limit of $f(x_1)$, $f(x_2)$, ..., exist.

This condition may be expressed as

$$f(x) \le \liminf_{y\to x} f(y) = \lim_{\varepsilon\downarrow 0}(\inf\{f(y) : \|y - x\| \le \varepsilon\}).$$

Similarly, f is said to be **upper semicontinuous** at x if

$$f(x) \ge \limsup_{y\to x} f(y) = \lim_{\varepsilon\downarrow 0}(\sup\{f(y) : \|y - x\| \le \varepsilon\}).$$

The combination of lower and upper semi continuity at x is ordinary continuity at x.

Theorem 10.2.28 *[Roc70] A function $f : \mathbb{R}^n \to \mathbb{R}$ is lower semicontinuous if and only if the epigraph of f is a closed set.*

Definition 10.2.29 *Consider now any function $f : [0,1] \to \mathbb{R}$. Set*

$$F(x) = \sup \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|$$

in which the supremum is taken over all n and over all t_i such that $0 = t_0 < t_1 ... < t_n = x$, for all $x \in [0,1]$. F is called the ***total variation*** *of f. If $F(b) < \infty$, then f is said to be of bounded variation on $[0,1]$.*

If f is of bounded variation, then f is differentiable almost everywhere and $\dot{f} \in L^1([0,1];\mathbb{R})$.

Lemma 10.2.30 *[DS88] Let f be a function of bounded variation in the interval (a,b). Then $f(a^+)$ and $f(b^-)$ exist.*

Measurable Space

An ordered pair $(\Omega, \mathcal{M})$ consisting of a set Ω and a σ-algebra $\mathcal{M}$ of subsets of Ω is called a **measurable space**. Any set in $\mathcal{M}$ is called an $\mathcal{M}$ - measurable set, but when the σ-algebra $\mathcal{M}$ is fixed (as is generally the case), the set will usually be said to be measurable.

Measurable Functions

Definition 10.2.31 *Assume that Ω is any set and $\mathcal{M}$ is any σ-algebra of subsets of Ω. Suppose $f : \Omega \to [-\infty,\infty]$ then f is* ***measurable function*** *if for all $t \in [-\infty,\infty]$, the set $f^{-1}([-\infty,t])$ belongs to $\mathcal{M}$. In other words*

$$\{x \in \Omega : f(x) \leq t\} \in \mathcal{M}.$$

Theorem 10.2.32 *[Bar95] Let f and g be measurable functions and let c be real number. The functions*

$$cf, f^2, f+g, fg, |f|,$$

are also measurable.

Theorem 10.2.33 *[HM61] If $f(y)$ is a continuous function and $y = g(x)$ is a measurable function, then the composite function $f(g(x))$ is measurable.*

Theorem 10.2.34 *[Bar95] Let $\{f_n(x)\}$ be a sequence of measurable functions and define the functions*

$$f(x) = \inf f_n(x), \qquad F(x) = \sup f_n(x),$$
$$f^*(x) = \liminf f_n(x), \quad F^*(x) = \limsup f_n(x).$$

Then f, F, f^* and F^* are measurable functions.

Definition 10.2.35 *The smallest σ-algebra $\mathcal{B}$ containing all the closed sets of a given topological space Ω is called the **Borel algebra** of Ω, and the set in $\mathcal{B}$ are called the **Borel sets**.*

Borel Functions

A mapping $f : \Omega \to \Gamma$, where Ω and Γ are metric spaces, is Borel measurable if $f^{-1}(U)$ is a Borel subset of Ω for every open set $U \subset \Gamma$.

Note: We denote $\mathcal{L}^n \times \mathcal{B}^k$ the σ- algebra of subsets of $\mathbb{R}^n \times \mathbb{R}^k$ generated by products of sets in the Lebesgue σ-algebra of $\mathbb{R}^n$ and the Borel σ-algebra of $\mathbb{R}^k$.

Proposition 10.2.36 *[Vin00] Consider a function $f : [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ satisfying the following hypotheses:*

(i) *$f(t,.,u)$ is continuous for each $(t,u) \in [a,b] \times \mathbb{R}^m$;*

(ii) *$f(.,x,.)$ is $\mathcal{L} \times \mathcal{B}^n$ measurable for each $x \in \mathbb{R}^n$.*

Then for any Lebesgue measurable function $x : [a,b] \to \mathbb{R}^n$, the mapping $(t,u) \to f(t,x(t),u)$ is $\mathcal{L} \times \mathcal{B}^n$ measurable.

Measure Space

Definition 10.2.37 *A* ***set function*** *is a function defined on a family of sets, and having values either in a Banach Space, which may be the set of real numbers or in the extended real number system, in which case its range contains at most one of the improper values $+\infty$ and $-\infty$. A positive set function is a real valued or extend real valued set function which has no negative values.*

Definition 10.2.38 *Let μ be a vector valued or extended real valued additive set function defined on a algebra $\mathcal{M}$ of subsets of a set Ω. Then μ is said to be* ***countably additive*** *if*

$$\mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$$

whenever E_1, E_2, ... are disjoints sets in $\mathcal{M}$ whose union also belongs to $\mathcal{M}$.

Definition 10.2.39 *A measure space is a triple $(\Omega, \mathcal{M}, \mu)$ consisting of a set Ω, a σ-algebra of $\mathcal{M}$ of Ω, and a countable additive μ defined on $\mathcal{M}$. The* ***measure space*** *is said to be finite if μ does not take on either of the values $+\infty$ or $-\infty$, and to be positive if μ never takes on a negative value.*

Definition 10.2.40 *A measure μ on $\mathbb{R}^n$ is called* ***Borel regular*** *if for each $\mathcal{A} \subset \mathbb{R}^n$ there exists a Borel set $\mathcal{B}$ such that $\mathcal{A} \subset \mathcal{B}$ and $\mu(\mathcal{A}) = \mu(\mathcal{B})$.*

Definition 10.2.41 *The* ***support of a measure*** *$\mu \in C^*([a,b] : \mathbb{R}^n)$, written supp$\{\mu\}$, is the smallest closed subset $A \subset [a,b]$ with the property that for all relatively open subsets $B \subset [a,b] \backslash A$ we have $\mu(B) = 0$.*

Integral

Definition 10.2.42 *A function φ on a measurable space Ω whose range consists of only finitely many points will be called a simple function. Among these are the nonnegative simple functions, whose the range is a finite subset of $[0, \infty)$. Note that* *we explicitly exclude* ∞ *from values of a simple function. If* $\alpha_1, ..., \alpha_n$ *are distinct values of a simple function* φ*, and if we set* $A_i = \{x : \varphi(x) = \alpha_i\}$*, then*

$$\varphi = \sum_{i=1}^{n} \alpha_i \mathcal{X}_{A_i}, \tag{10.1}$$

where $\mathcal{X}_A$*, is the characteristic function of* A_i*.*

Note that φ is measurable if and only if each of sets A_i is measurable.

Definition 10.2.43 *Let* $\varphi : \Omega \to [0, \infty)$ *be a measurable simple function, we define*

$$\int \varphi d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i),$$

where α_i*,* A_i*,* $i = 1, ..., n$ *are as in (10.1).*

The convention $0 \cdot \infty = 0$ is used here; it may happen that $\alpha_i = 0$ for some i and that $\mu(A_i) = \infty$.

Theorem 10.2.44 *[Rud87] Let* $f : \Omega \to [0, \infty]$ *be measurable. There exist simple measurable functions* φ_n *on* Ω *such that*

(i) $0 \leq \varphi_1 \leq \varphi_2 \leq ... \leq f$.

(ii) $\varphi_n(x) \to f(x)$ *as* $n \to \infty$*, for every* $x \in \Omega$*.*

Definition 10.2.45 *If* $f : \Omega \to [0, \infty]$ *is measurable function, the integral of* f *is defined by*

$$\int f d\mu = \sup \int \varphi d\mu,$$

where the supremum is taken over all simple measurable functions φ *such that* $0 \leq \varphi \leq f$*.*

Definition 10.2.46 *For any measurable set* E*, and nonnegative measurable function* f*,* $\int_E f dx = \int f \mathcal{X}_E d\mu$ *is the integral of* f *over* E*.*

Theorem 10.2.47 *(Lebesgue's Monotone Convergence Theorem) [Rud87] Let* $\{f_n\}$ *be a sequence of measurable functions on* Ω*, and suppose that*

(i) $0 \le f_1(x) \le f_2(x) \le ... \le \infty$ *for every* $x \in \Omega$*,*

(ii) $f_n(x) \to f(x)$ *as* $n \to \infty$*, for every* $x \in \Omega$*.*

Then f *is measurable, and*

$$\int_\Omega f_n d\mu \to \int_\Omega f d\mu \text{ as } n \to \infty.$$

Lemma 10.2.48 *(Fatou's Lemma) [Rud87] If* $f_n : \Omega \to [0, \infty]$ *is measurable, for each positive integer* n*, then*

$$\int_\Omega (\liminf_{n\to\infty} f_n) d\mu \le \liminf_{n\to\infty} \int_\Omega f_n d\mu.$$

Theorem 10.2.49 *[Bar95] If* $f : \Omega \to [0, \infty]$ *is measurable and if* λ *is defined on* $\mathcal{M}$ *by*

$$\lambda(E) = \int_E f d\mu,$$

then λ *is a measure.*

Definition 10.2.50 *Let* $f : \Omega \to [0, \infty]$ *be a measurable function, we defined*

$$f^+(x) = \max(f(x), 0),\ f^-(x) = \max(-f(x), 0)$$

as positive and negative parts of f *respectively.*

Theorem 10.2.51 *[Bar74]*

(i) $f = f^+ - f^-;\ |f| = f^+ + f^-;\ f^+;\ f^- \ge 0.$

(ii) f *is measurable iff* f^+ *and* f^- *are measurable.*

We now proceed to some properties of the integral. In the following result, all function are assumed measurable from Ω to $[-\infty, \infty]$.

Theorem 10.2.52 *[Ash00]*

(i) *If* $\int f d\mu$ *exists and* $c \in \mathbb{R}$, *then* $\int c f d\mu$ *exists and equals* $c \int f d\mu$.

(ii) *If* $g(x) \geq f(x)$ *for all* x, *then* $\int g d\mu \geq \int f d\mu$ *in the sense that if* $\int f d\mu$ *exists and is greater that* $-\infty$, *then* $\int g d\mu$ *exists and* $\int g d\mu \geq \int h d\mu$.

(iii) *If* $\int f d\mu$ *exists, then* $|\int f d\mu| \leq \int |f| d\mu$.

(iv) *If* $f \geq 0$ *and* $B \in \mathcal{M}$, *then* $\int_B f d\mu = \sup\{\int_B \varphi d\mu : 0 \leq \varphi \leq f, \varphi$ *is a simple function*$\}$.

(v) *If* $\int f d\mu$ *exists, so does* $\int_A f d\mu$ *for each* $A \in \mathcal{M}$; *if* $\int f d\mu$ *is finite, then* $\int_A f d\mu$ *is also finite for each* $A \in \mathcal{M}$.

Definition 10.2.53 *We define* $L^1(\mu)$ *to be collection of all measurable functions* f *on* Ω *for which*

$$\int_\Omega |f| d\mu < \infty.$$

Theorem 10.2.54 *Suppose* f *and* $g \in L^1(\mu)$ *and* α *and* β *are real numbers. Then* $\alpha f + \beta g \in L^1(\mu)$ *and*

$$\int_\Omega (\alpha f + \beta g) d\mu = \alpha \int_\Omega f d\mu + \beta \int_\Omega g d\mu.$$

Theorem 10.2.55 *(Lebesgue's Dominated Convergence Theorem)[Rud87] Suppose* $\{f_n\}$ *is a sequence of measurable functions on* Ω *such that*

$$f(x) = \lim_{n\to\infty} f_n(x)$$

exists for every $x \in \Omega$. *If there is a function* $g \in L^1(\mu)$ *such that*

$$|f_n(x)| \leq g(x) \; n = 1, 2, 3, ...; x \in \Omega,$$

then $f \in L^1(\mu)$,

$$\lim_{n\to\infty} \int_\Omega |f_n - f| d\mu = 0,$$

and

$$\lim_{n\to\infty}\int_\Omega f_n d\mu = \int_\Omega f d\mu.$$

Convergence of Measures

The set of elements $\mu \in C^*([a,b];\mathbb{R})$ taking nonnegative values on nonnegative-valued function in $C([a,b];\mathbb{R})$ is denoted $C^\oplus(a,b)$. The norm on $C^\oplus(a,b)$, written $\|\mu\|_{T.V.}$, is the total variation of μ, $\int_{[a,b]}\mu(ds)$.

Given $\mu \in C^\oplus(a,b)$, μ- continuity set is a Borel subset $\mathcal{B} \subset [a,b]$ for which $\mu(bdy\ \mathcal{B}) = 0$. Take $\mu \in C^\oplus(a,b)$. Then there is a countable set $S \subset (a,b)$, such that all sets of the form $[s,t], [s,t), (s,t]$ with $s,t \in ([a,b]\backslash S)$ are μ-continuity sets.

Take a weak* convergent sequence $\mu_i \to \mu$ in $C^\oplus(a,b)$. Then,

$$\int_B h(t)\mu(dt) = \lim_{i\to\infty}\int_B h(t)\mu_i(dt)$$

for any relatively open subset $B \subset [a,b]$, any $h \in C([a,b];\mathbb{R}^n)$ and any μ- continuity set B.

Multifunction and Trajectories

Definition 10.2.56 *Take a set Ω. A* ***multifunction*** $\Gamma : \Omega \rightsquigarrow \mathbb{R}^n$ *is a mapping from Ω to the subsets of $\mathbb{R}^n$; that means for each $x \in \Omega$, then $\Gamma(x)$ is a subset of $\mathbb{R}^n$.*

A multifunction $\Gamma : \Omega \rightsquigarrow \mathbb{R}^n$ is called closed, compact, convex, or nonempty if for all $x \in \Omega$, $\Gamma(x)$ has the property in question.

Definition 10.2.57 *Let $(\Omega, \mathcal{M})$ be a measurable space. Take a multifunction $\Gamma : \Omega \rightsquigarrow \mathbb{R}^n$. Γ is a* ***measurable*** *when the set*

$$\{x \in \Omega : \Gamma(x) \cap C \neq \emptyset\}$$

is $\mathcal{M}$ for every open set $C \subset \mathbb{R}^n$.

Theorem 10.2.58 *[Cla83]* ***Measurable Selection*** *Let Γ be a measurable, closed, and nonempty on S. Then there exists a measurable function $\gamma : S \to \mathbb{R}^n$ such that $\gamma(x)$ belongs to $\Gamma(x)$ for all $x \in S$.*

Definition 10.2.59 *Take a multifunction $\Gamma : I \rightsquigarrow \mathbb{R}^n$. We say that a function $x : I \to \mathbb{R}^n$ is a* ***measurable selection*** *for Γ if*

(i) *x is Lebesgue measurable, and*

(ii) *$x(t) \in \Gamma(t)$ a.e.*

Theorem 10.2.60 *[Vin00]* ***Aumann's measurable selection theorem*** *Let $\Gamma : I \rightsquigarrow \mathbb{R}^n$ be a nomempty multifunction. Assume that*

$$Gr\,\Gamma \text{ is } \mathcal{L} \times \mathcal{B}^k \text{ measurable,}$$

then Γ has a measurable selection.

Definition 10.2.61 *Consider the case in which $S = [a, b]$, an interval in $\mathbb{R}$. We say that Γ is* ***integrably bounded*** *provided there is an integrable function $\phi(t)$ such that for all $t \in [a, b]$,for all $\gamma \in \Gamma(t)$, $|\gamma| \leq \phi(t)$.*

Definition 10.2.62 *Γ is said to be a* ***Lipschitz Multifunction*** *on S (of rank k) provided that for all x_1, x_2 in S and for all γ_1 in $\Gamma(x_1)$ there exists γ_2 in $\Gamma(x_2)$ such that*

$$\|\gamma_1 - \gamma_2\| \leq k\|x_1 - x_2\|.$$

Definition 10.2.63 *Γ is said to be* ***upper semicontinuous*** *at x if for all $\epsilon > 0$, there exists a $\delta > 0$ such that*

$$\Gamma(x') \subset \Gamma(x) + \epsilon\mathbb{B}$$

for all $x' \in x + \delta\mathbb{B}$. It is ***lower semicontinuous*** *at x if, for all $\gamma \in \Gamma(x)$ and all $\epsilon > 0$, there exists a $\delta > 0$ such that*

$$\Gamma(x') \cap (y + \epsilon\mathbb{B}) \neq \emptyset$$

for all $x' \in x + \delta\mathbb{B}$. It is ***continuous*** *at x if it is simultaneously upper semi-continuous and lower semi-continuous.*

Let S, Ω_t, be the sets defined by

$$S = \{t : (t,x) \in \Omega \text{ for some } x \in \mathbb{R}^n\}$$
$$\Omega_t = \{x : (t,x) \in \Omega\}.$$

Ω is called a tube provided the set S is an interval ($[a,b]$, say) and provided there exist a continuous function $w(t)$ and a continuous positive function ε on $[a,b]$ such that $\Omega_t = w(t) + \varepsilon(t)\mathbb{B}$ for t in $[a,b]$.

Definition 10.2.64 *Let Ω be a tube on $[a,b]$. F is said to be measurably Lipschitz on Ω provided:*

(i) *For each x in $\mathbb{R}^n$, the multifunction $t \rightsquigarrow F(t,x)$ is measurable on $[a,b]$.*

(ii) *There is an integrable function $k(t)$ on $[a,b]$ such that for each t in $[a,b]$, the multifunction $x \rightsquigarrow F(t,x)$ is nonempty and Lipschitz of rank $k(t)$ on Ω_t.*

Proposition 10.2.65 *[Vin00] Take a weak* convergent sequence $\{\mu_i\}$ in $C^{\oplus}(a,b)$, a sequence of Borel measurable functions $\{\gamma_i : [a,b] \to \mathbb{R}^n\}$, and a sequence of closed sets $\{A_i\}$ in $[a,b] \times \mathbb{R}^n$. Take also a closed set A in $[a,b] \times \mathbb{R}^n$, and a measure $\mu \in C^{\oplus}(a,b)$.*

Assume that $A(t)$ is convex for each $t \in domA(\cdot)$ and that the sets A and $A_1, A_2, ...$ are uniformly bounded. Assume further that

$$\lim_{i\to\infty} \sup A_i \subset A,$$

$$\gamma_i(t) \in A_i(t) \quad \mu_i \text{ a.e. for } i = 1, 2, ...$$

and

$$\mu_i \longrightarrow \mu_0 \text{ weakly}^*.$$

Define $\eta_i \in C^*([a,b];\mathbb{R}^k)$

$$\eta_i(dt) = \gamma_i(t)\mu_i(dt).$$

Then, along a subsequence

$$\eta_i \to \eta_0 \text{ weakly}^*,$$

for some $\eta_0 \in C^*([a,b];\mathbb{R}^k)$ *such that*

$$\eta_0(dt) = \gamma_0(t)\mu_0(t),$$

in which γ_0 *is a Borel measurable function that satisfies*

$$\gamma_0(t) \in A(t) \quad \mu_0 \text{ a.e.}.$$

Multifunction in Optimal Control

Definition 10.2.66 *A **trajectory** (for F, or for the differential inclusion) is an arc x such that for almost all $t \in [a,b]$,*

$$\dot{x}(t) \in F(t, x(t)) \text{ a.e.}$$

Theorem 10.2.67 *[Vin00]**Compactness of Trajectories** Take a relatively open subset $\Omega \subset [a,b] \times \mathbb{R}^n$ and a multifunction $F : \Omega \rightsquigarrow \mathbb{R}^n$. Assume that, for some closed multifunction $X : [a,b] \rightsquigarrow \mathbb{R}^n$ such that $GrX \subset \Omega$, the following hypotheses are satisfied:*

(i) *F is a closed, convex, nonempty multifunction.*

(ii) *F is $\mathcal{L} \times \mathcal{B}^n$ measurable.*

(iii) *For each $t \in [a,b]$, the graph of $F(t,.)$ restricted to $X(t)$ is closed.*

Consider a sequence $\{x_i\}$ of $W^{1,1}([a,b];\mathbb{R}^n)$ functions, a sequence $\{r_i\}$ in $L^1([a,b];\mathbb{R})$ such that $\|r_i\|_{L^1} \to 0$ as $i \to \infty$ and a sequence $\{A_i\}$ of measurable subsets of $[a,b]$ such that meas $A_i \to |b-a|$ as $i \to \infty$.

Suppose that:

(iv) $Grx_i \subset GrX$ *for all* i*;*

(v) $\{\dot{x}_i\}$ *is a sequence of uniformly integrally bounded function on* $[a, b]$ *and* $\{x_i(a)\}$ *is a bounded sequence;*

(vi) *there exists* $c \in L^1$ *such that*

$$F(t, x_i(t)) \subset c(t)\mathbb{B}$$

for a.e. $t \in A_i$ *and for i=1,2,... . Suppose further that*

$$\dot{x}_i(t) \in F(t, x_i(t)) + r_i(t)\mathbb{B} \ a. \ e \ t \in A_i.$$

Then along some subsequence

$$x_i \to x \text{ uniformly and } \dot{x}_i \to \dot{x} \text{ weakly in } L^1$$

for some $x \in W^{1,1}([a, b]; \mathbb{R}^n)$ *satisfying*

$$\dot{x}(t) \in F(t, x(t)) \ a.e. \ t \in [a, b].$$

Cones

Definition 10.2.68 *A set* $C \subset \mathbb{R}^n$ *is called a* ***cone*** *if it is nonempty and for all* $\lambda \geq 0$ *and* $v \in C$ *we have* $\lambda v \in C$*.*

Definition 10.2.69 *The* ***Negative Polar cone*** *of a set* $C \subset \mathbb{R}^n$ *is defined by*

$$C^- = \{x^* \in \mathbb{R}^n : x^* \cdot x \leq 0, \forall y \in C\}.$$

Convex Analysis

Definition 10.2.70 *A set $C \subset \mathbb{R}^n$ is **convex** if, for all $x \in C$ and $y \in C$, the line segment $[\{\alpha x + (1-\alpha)\, y \in \mathbb{R}^n,$ with $\alpha \in [0,1]\}]$ belongs to C.*

Definition 10.2.71 *The **convex hull** of a set C, denoted by co C, is the smallest set that containing C. In other words, co C is the intersection of all sets containing C.*

The convex hull can also be defined as:

$$\text{co } C = \{\sum \lambda_i x_i : \sum \lambda_i = 1, \lambda_i \geq 0 \text{ with } i = 1, ..., k \text{ and } k \geq 1,\ x_i \in C\}.$$

Nonsmooth Analysis

Definition 10.2.72 *The **limiting normal cone** of a closed set $C \subset \mathbb{R}^n$ at $x \in C$, denoted by $N_C(x)$, is the set*

$$\begin{aligned} N_C^L(x) = \quad & \{\eta \in \mathbb{R}^n : \exists \textit{ sequences } \{M_i\} \in \mathbb{R}^+, x_i \to x,\ \eta_i \to \eta \textit{ such that} \\ & x_i \in C \textit{ and } \eta_i \cdot (y - x_i) \leq M_i \|y - x_i\|^2 \textit{ for all } y \in \mathbb{R}^n, i = 1, 2, ...\}. \end{aligned}$$

In [Vin00], we can find the proof of some elementary properties of cones:

Proposition 10.2.73 *[Vin00] Take a closed set $C \subset \mathbb{R}^n$ and a point $x \in C$. Then*

(i) *$x \in int\{C\}$ implies $N_C^L(x) = \{0\}$;*

(ii) *$x \in bdy\{C\}$ implies that contains nonzero elements.*

Proposition 10.2.74 *[Vin00] Take closed subsets $C_1 \subset \mathbb{R}^m$ and $C_2 \subset \mathbb{R}^n$, and a point $(x_1, x_2) \in C_1 \times C_2$. Then*

$$N_{C_1 \times C_2}^L(x_1, x_2) = N_{C_1}^L(x_1) \times N_{C_2}^L(x_2)$$

Note: The limiting normal cones and limiting subdifferential are closed sets, but they are not necessarily convex.

Definition 10.2.75 *The* ***Contingent Cone*** *of a closed set $C \subset \mathbb{R}^n$ at $x \in C$, denoted by $T_C(x)$, is the set*

$$T_C(x) = \{v \in \mathbb{R}^n | \liminf_{h \to 0^+} \frac{dist(x + hv, C)}{h} = 0\}.$$

Definition 10.2.76 *The* ***Clarke Tangent Cone*** *of a closed set $C \subset \mathbb{R}^n$ at $x \in C$, denoted by $C_C(x)$, is the set*

$$C_C(x) = \{v \in \mathbb{R}^n | \lim_{h \to 0^+ x' \to_C x} \frac{dist(x' + hv, C)}{h} = 0\},$$

where $\to_C$ denotes the convergence in C.

Note: We can have the following characterization of these cones in terms of sequences:

i) $v \in T_C(x)$ if and only if $\exists h_n \to 0+$ and $\exists v_n \to v$ such that $\forall n$, $x + h_n v_n \in C$. This implies that if $x \in Int(C)$, then $T_C(x) = \mathbb{R}^n$.

ii) $C_C(x)$ comprises vectors ξ such that for any sequences $x_n \to_C x$ and $h_n \downarrow 0$ there exists a sequence k_n in C such that $h_n^{-1}(k_n - x_n) \to \xi$.

Definition 10.2.77 *The* ***Clarke Normal Cone*** *to C at x is defined by*

$$N_C(x) = \{\xi \in \mathbb{R}^n | \xi \cdot v \leq 0, \forall v \in T_C(x)\}.$$

Note that: $N_C(x) = T_C(x)^-$.

Theorem 10.2.78 *[Vin00] Take a closed set $C \subset \mathbb{R}^k$ and $x \in C$. Then the Clark tangent cone $C_c(x)$ and the limiting normal cone $N_C^L(x)^-$ are related according*

$$C_C(x) = N_C^L(x)^-.$$

Definition 10.2.79 *We shall say that a closed subset C is* ***sleek*** *at $x_0 \in C$ if the multifunction,*

$$x \rightsquigarrow T_C(x), \forall x \in C$$

is lower semicontinuous at x_0 and that it is ***sleek*** *if it is sleek at every point of C.*

Theorem 10.2.80 *[AF90] Let C be a closed set of $\mathbb{R}^n$. If C is sleek for all $x \in C$, then $T_C(x) = C_C(x)$, and consequently are convex.*

Definition 10.2.81 *Take a lower semicontinuous function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and a point $x \in domf$. The* ***limiting subdifferential*** *of f at x, written $\partial^L f(x)$, is the set*

$$\partial^L f(x) = \{\eta \in \mathbb{R}^n : (\eta, -1) \in N^L_{epi\ f}(x, f(x)\}. \tag{10.2}$$

where $epi\ f = \{(x, \alpha) \in \mathbb{R}^{n+1} : \alpha \geq f(x)\}$ denotes the epigraph of a function f.

Proposition 10.2.82 *[Vin00] Take a lower semicontinuous function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and a point x on $\mathbb{R}^n$. Assume that f is Lipschitz continuous on a neighborhood of x with Lipschitz constant K. Then $\partial^L f(x)$ is nonempty and $\partial^L f(x) \subset K\mathbb{B}$;*

Definition 10.2.83 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz continuous function on a neighborhood of some point $x \in \mathbb{R}^n$. The* ***Clark's Subdifferential****, denoted by $\tilde{\partial} f$, is defined by $\tilde{\partial} f(x) = co\ \partial f(x)$.*

Consider a function $f : \mathbb{R}^n \to \mathbb{R}$ and a point $x \in \mathbb{R}^n$ such that f is Lipschitz continuous on a neighborhood of x. According to Rademacher's Theorem, f is differential almost everywhere on this neighborhood (with the respect to n-dimensional Lebesgue measure).

Theorem 10.2.84 *[Cla83], [Vin00] Take a function $f : \mathbb{R}^n \to \mathbb{R}$, a point $x \in \mathbb{R}^n$ and any subset $\Omega \subset \mathbb{R}^n$ having Lebesgue measure zero. Assume that f is Lipschitz continuous on a neighborhood of x. Then*

$$co\ \partial^L f(x) = co\ \{\eta \in \mathbb{R}^n : \exists x_i \to x, x_i \notin \Omega, f_x(x_i) \text{ exist and } f_x(x_i) \to \varepsilon\}.$$

Definition 10.2.85 *Take a point $y \in \mathbb{R}^n$ and a function $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ that is Lipschitz continuous on a neighborhood of y. Then the* ***Generalized Jacobian*** *$DL(y)$ of L at y is the set of $m \times n$ matrices:*

$$DL(y) = co\{\eta : \exists y_i \to y \text{ such that } L_y(y_i) \text{ exist } \forall i \text{ and } L_y(y_i) \to \eta\}.$$

Proposition 10.2.86 *[Cla83] For any vector $v \in \mathbb{R}^n$*

$$vDL(y) = \tilde{\partial}(rL)(y).$$

Definition 10.2.87 *Consider a multifunction map $F : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$, Lipschitz around x and let $y \in F(x)$. The* ***adjacent derivative*** *of F at (x, y) is the multifunction map $\bar{d}F(x, y)$ from $\mathbb{R}^n$ into subsets of $\mathbb{R}^n$ defined by*

$$\bar{d}F(x,y)w = \{v \in \mathbb{R}^n : \lim_{s \to 0^+} dist(v, \frac{F(x+sw)-y}{s}) = 0\}.$$

Definition 10.2.88 *Let $x_0 \in dom(h)$. The* ***superdifferential*** *of h at x_0 is the closed convex set*

$$\partial^+ h(x_0) = \{p \in \mathbb{R}^n : \limsup_{x \to_X x_0} \frac{h(x) - h(x_0) - p \cdot (x - x_0)}{\|x - x_0\|} \leq 0\}.$$

Definition 10.2.89 *The* ***upper derivative*** *of $h(\cdot)$ at $x_0 \in dom(h)$ in the direction θ is given by*

$$D^+h(x_0)(\theta) = \limsup_{s \to 0^+,\, \theta' \to \theta,\, x_0 + s\theta' \in X} \frac{h(x_0 + s\theta') - h(x_0)}{s}, \forall \theta \in T_X(x_0)$$

and $D^+h(x_0)(\theta) = -\infty$ for all $\theta \notin T_X(x_0)$.

We also make use of the hybrid partial subdifferential.

Definition 10.2.90 *The* ***hybrid partial subdifferential****, denoted by $\partial_x^> h(t, x)$,* *is defined by*

$$\partial_x^{>} h(t,x) = \; co\ \{\eta : \exists (t_i, x_i) \to (t,x) : h(t_i,x_i) > 0 \forall i,\ h(t_i,x_i) \to h(t,x) \\ and\ h_x(t_i,x_i) \to \eta\}.$$

In the case of the function h do not depend of the time t, we have

$$\partial^{>} h(x) = \; \text{co}\ \{\eta : \exists x_i \to x : h(x_i) > 0 \forall i,\ h(x_i) \to h(x) \\ \text{and } h_x(x_i) \to \eta\}.$$

Now we proceed to derive an assortment that facilitates greatly the calculation of $\tilde{\partial} f$.

Scalar Multiples:

Assume that f is Lipschitz continuous near of a point x. For any scalar α, one has

$$\tilde{\partial}(\alpha f(x)) = \alpha \tilde{\partial} f(x). \tag{10.3}$$

Local Extreme

If f is Lipschitz continuous near of a point x and attains a local minimum or maximum at x, then $0 \in \tilde{\partial} f(x)$.

Sum Rule:

Let f_1 and f_2 be Lipschitz continuous near x, then

$$\tilde{\partial}(f_1 + f_2)(x) \subset \tilde{\partial} f_1(x) + \tilde{\partial} f_2(x). \tag{10.4}$$

Products Rule:

Let f_1 and f_2 be Lipschitz continuous near x. Then $f_1 f_2$ is Lipschitz continuous near x and

$$\tilde{\partial}(f_1 f_2)(x) \subset f_2(x) \tilde{\partial} f_1(x) + f_1(x) \tilde{\partial} f_2(x). \tag{10.5}$$

Quotients Rule:

Let f_1 and f_2 be Lipschitz continuous near x, and suppose $f_2(x) \neq 0$. Then f_1/f_2 is Lipschitz continuous near x, and one has

$$\tilde{\partial}(\frac{f_1}{f_2})(x) \subset \frac{f_2(x)\tilde{\partial}f_1(x) - f_1(x)\tilde{\partial}f_2(x)}{f_2^2(x)}. \tag{10.6}$$

Max Rule

Suppose that f_i is a finite collection of functions ($i = 1, 2, ..., m$) each of which is Lipschitz continuous function on a neighborhood of x The function f defined by $f(x) = \max\{f_i(x) : i = 1, ..., m\}$ is Lipschitz continuous near x and

$$\tilde{\partial}f(x) \subset co\ \{\tilde{\partial}f_i(x) : i \in I(x)\},$$

where $I(x)$ denote the set of indices i for which $f_i(x) = f(x)$, for any x (i.e. the indices at which the maximum defining f is attained).

Mean-Value Theorem

Suppose that f is Lipschitz continuous on a open set containing the line segment $[x, y]$. Then there exist a point $u \in (x, y)$ such that

$$f(y) - f(x) \in \tilde{\partial}f(u) \cdot (y - x). \tag{10.7}$$

Partial Generalized Gradients

Let $f(x_1, x_2)$ be a Lipschitz continuous function on a neighborhood of (x_1, x_2). We denote by $\tilde{\partial}_{x_1}f(x_1, x_2)$ the *partial generalized gradients* of $f(\cdot, x_2)$ at x_1 and by $\tilde{\partial}_{x_2}f(x_1, x_2)$ *partial generalized gradients* of $f(x_1, \cdot)$ at x_2.

If f is convex at $x = (x_1, x_2)$, then

$$\tilde{\partial}_{(x_1,x_2)}f(x_1, x_2) \subset \tilde{\partial}_{x_1}f(x_1, x_2) \times \tilde{\partial}_{x_2}f(x_1, x_2).$$

Some Important Results

Lemma 10.2.91 *[Vin00]* ***Dubois - Reymond Lemma*** *Take a function* $a \in L^2([0,1];\mathbb{R}^n)$*. Suppose that*

$$\int_0^1 a(t) \cdot w(t)dt = 0$$

for every continuous function w *that satisfies*

$$\int_0^1 w(t)dt = 0.$$

Then there exists some vector $c \in \mathbb{R}^n$ *such that*

$$a(t) = d \text{ for a.e. } t \in [0,1].$$

Theorem 10.2.92 *[Vin00]* ***(Exact Penalization Theorem)*** *Take a set* $C \subset \mathbb{R}^n$ *and a Lipschitz function* $f : \mathbb{R}^n \to \mathbb{R}$*, with Lipschitz constant* K*. Let* $\bar{x}$ *be a minimizer for the constrained minimization problem,*

$$\begin{aligned} &\text{Minimize} \;\; f(x) \\ &\text{subject to} \;\; x \in C. \end{aligned}$$

Choose any $\widehat{K} \geq K$*. Then* $\bar{x}$ *is a minimizer also for the unconstrained minimization problem,*

$$\begin{aligned} &\text{Minimize} \;\; f(x) + \widehat{K} d_C(x) \\ &\text{subject to} \;\; x \in \mathbb{R}^n. \end{aligned}$$

Theorem 10.2.93 *[Vin00]**Gronwall's Inequality*** *Take an absolutely continuous function* $z : [S,T] \to \mathbb{R}^n$*. Assume that there exist nonnegative integrable functions* k *and* v *such that*

$$\|\frac{d}{dt} z(t)\| \leq k(t)\|z(t)\| + v(t) \;\; a.e. \;\; t \in [S,T].$$

Then

$$\|z(t)\| \leq \exp\left(\int_S^t k(\sigma)d\sigma\right)\left[\|z(S)\| + \int_S^t \exp\left(-\int_S^\tau k(\sigma)d\sigma\right)v(\tau)d\tau\right]$$

for all $t \in [S, T]$.

Bibliography

[AA97] A. V. Arutyunov and S. M. Aseev, *Investigation of the degeneracy phenomenon of the maximum principle for optimal control problems with state constraints*, SIAM Journal of Control and Optimization **35** (1997), 930–952.

[Aba67] J. Abadie, *On the Kuhn-Tucker theorem*, Nonlinear Programming (J. Abadie, ed.), North Holland, Amsterdam, 1967, pp. 21–56.

[AF90] J. P. Aubin and H. Frankowska, *Set-valued analysis*, Birkhauser, Boston, 1990.

[Aru00] A. V. Arutyunov, *Optimality condition: abnormal and degenerate problems*, Kluwer, 2000.

[Ash00] Robert B. Ash, *Probability and measure theory*, Academic Press, 2000.

[Bar74] G. De Barra, *Introduction to measure theory*, Van Nostrand Reinhold Company, 1974.

[Bar95] Robert G. Bartle, *The elements of integration and lebesgue measure*, A Wiley-Interscience Publication, 1995.

[Ber95] D. P. Bertesekas, *Dynamic programming and optimal control*, Athena Scientific, 1995.

[BF07] P. Bettiol and H. Frankowska, *Normality of the maximum principle for nonconvex constrained bolza problems*, Journal of Differential Equations **243** (2007), 256–269.

[BSS93] A. S. Bazaraa, H.D. Sherali, and C.M. Shetty, *Nonlinear programming*, John Wiley and Sons, Inc., 1993.

[CF05] A. Cernea and H. Frankowska, *A connection between the maximum principle and dynamic programming for constrained control problems*, SIAM Journal of Control and Optimization (2005).

[Cla73] F. H. Clarke, *Necessary conditions for nonsmooth problems in optimal control and calculus of variations*, Ph.D. thesis, University of Washington, 1973.

[Cla76a] ______, *The maximum principle under minimal hypotheses*, SIAM Journal of Control and Optimization **14** (1976), no. 6, 1078–1091.

[Cla76b] ______, *Optimal solutions to differential inclusions*, Journal of Optimization Theory and Applications **19** (1976), no. 3, 469–478.

[Cla83] ______, *Optimization and nonsmooth analysis*, Wiley-Interscience, New York, 1983.

[Cla89] ______, *Methods of dynamic and nonsmooth optimization*, SIAM, 1989.

[CLV97] F. H. Clarke, Y. S. Ledyaev, and R. B. Vinter, *Regularity properties of solutions to linear quadratic optimal control problems with state constraints*, Systems and Control Letters **30** (1997), 265–272.

[CV85] F. H. Clarke and R. B. Vinter, *Regularity properties of solutions to the basic problem in the calculus of variations*, Transactions of the American Mathematical Society **289** (1985), no. 1, 73–98.

[DK95] A. L. Dontchev and I. Kolmanovsky, *On regularity of optimal control*, in Proc. French-German Conf. on Optimization, Eds. R. Durier, C. Michelot, Lecture Notes in Economics and Mathematical Systems, 429, Springer, 1995, pp. 125–135.

[dP03] M. d. R. de Pinho, *Mixed constrained control problems*, Journal of Mathematical Analysis and Applications **278** (2003), 293–307.

[DS88] Nelson Dunford and Jacob T. Schwartz, *Linear operators*, A Wiley-Interscience Publication, 1988.

[Fer06] M. M. A. Ferreira, *On the regularity of optimal controls for a class of problems with state constraints*, International Journal of Systems Science **37** (2006), no. 8, 495–502.

[FFV99] M. M. A. Ferreira, F. A. C. C. Fontes, and R. B. Vinter, *Nondegenerate necessary conditions for nonconvex optimal control problems with state constraints*, Journal of Mathematical Analysis and Applications **233** (1999), 116–129.

[FM06] H. Frankowska and E. M. Marchini, *Lipschitzianity of optimal trajectories for the bolza optimal control problem*, Calculus of Variations and Partial Differential Equations **27** (2006), no. 4, 467–492.

[Fon99] F. A. C. C. Fontes, *Optimisation-based control of constrained nonlinear systems*, Ph.D. thesis, Imperial College of Science Technology and Medicine, University of London, London SW7 2BY, U.K., 1999.

[Fon00] ______, *Normality in the necessary conditions of optimality for control problems with state constraints*, Proceedings of the IASTED Conference on Control and Applications (Cancun, Mexico), 2000.

[Fon05] ______, *Nondegenerate necessary conditions for linear optimal problems with higher index state constraints,*, Proceedings ECC-CDC'05

European Control Conference and Conference on Decision and Control 2005, (Seville, Spain), 2005, pp. 2149–2151.

[FV94] M. M. A. Ferreira and R. B. Vinter, *When is the maximum principle for state constrained problems nondegenerate?*, Journal of Mathematical Analysis and Applications **187** (1994), no. 2, 438–467.

[GV03] G. N. Galbraith and R. B. Vinter, *Lipschitz continuity of optimal controls for state constrained problems*, SIAM Journal of Control and Optimization **42** (2003), no. 5, 1727–1774.

[HM61] S. Hartman and J. Mikusinski, *The theory of Lesbegue Measure and Integration*, Pergamon Press, 1961.

[HSV95] R. F. Hartl, S. P. Sethi, and R. G. Vickson, *A survey of the maximum principles for optimal control problems with state constraints*, SIAM Review **37** (1995), no. 2, 181–218.

[Joh48] Fritz John, *Extremum problems as inequalities as subsidiary conditions*, Studies and Essays: Courant Anniversary Volume (K. O. Friedrichs, O. E. Neugebauer, and J. J. Stoker, eds.), Interscience Publishers, New York, 1948.

[KT51] H. W. Kuhn and A. W. Tucker, *Nonlinear programming*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (J. Neyman, ed.), University of California Press, Berkeley, 1951, pp. 481–492.

[LF03] S. O. Lopes and F. A. C. C. Fontes, *Normality in necessary conditions for dynamic optimization problems*, Third Junior European Meeting on Control, Optimization and Computation (Aveiro, Portugal), 2003.

[LF06] ______, *Comparison of nondegenerate necessary conditions of optimality applied to calculus of variations problems*, Proceedings of Controlo 2006 (Lisboa, Portugal), 2006.

[LF07] ———, *On the degeneracy phenomenon in optimization problems with inequality constraints*, ORP3 conference (Guimaraes, Portugal), 2007.

[LF08a] ———, *Nondegenerate necessary conditions for nonlinear optimal problems with higher-index state constraints*, Proceedings in Applied Mathematics and Mechanics 2008, (ICIAM 2007- 6th International Congress on Industrial and Applied Mathematics) (Zurich, Switzerland), 2008.

[LF08b] ———, *An overview of nondegenerate necessary conditions of optimality applied to optimal control problems with higher index state constraints*, Accepeted to publication in International Journal of Mathematics and Statistics (IJMS) (Vila Real, Portugal), 2008.

[LFdP07] S. O. Lopes, F. A. C. C. Fontes, and M. d. R. de Pinho, *An integral-type constraint qualification for optimal problems with state constraints*, Technical report, Officina Mathematica, University of Minho, Guimaraes, Portugal, http://repositorium.sdum.uminho.pt/, 2007.

[Loe93] P. D. Loewen, *Optimal control via nonsmooth analysis*, American Mathematical Society, Providence, RI, 1993.

[Lop03] S. O. Lopes, *Fenómenos de degeneração em problemas do cálculo das variações com dados não suaves*, Master's thesis, Departamento de Matemática para a Ciência e Tecnologia, Universidade do Minho, Guimarães, Portugal, 2003.

[Lue69] D. G. Luenberguer, *Optimization by vector space methods*, John Wiley, New York, 1969.

[Mal03] K. Malanowski, *On normality of lagrange multipliers for state constrained optimal control problems*, Optimization **52** (2003), no. 1, 75–91.

[Man69] O. L. Mangasarian, *Nonlinear programming*, McGraw-Hill, New York, 1969.

[MdRdPZ01] R. B. Vinter M. d. R. de Pinho and H. Zheng, *A maximum principle for optimal control problems with mixed constraints*, IMA Journal of Mathematical Control and Information **18** (2001), 189–205.

[PBGM62] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The mathematical theory of optimal processes*, Wiley Interscience, New York, 1962.

[PF62] M. R. Pinho and M. M. Ferreira, *Optimal control problems with constraints*, Editura Electus, Bucareste, 1962.

[Roc70] R. Tyrrell Rockafellar, *Convex analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[Rud87] Walter Rudin, *Real and complex analysis*, McGraw-Hill International Editions, 1987.

[RV99] F. Rampazzo and R. B. Vinter, *A theorem on the existence of neighbouring feasible trajectories with aplication to optimal control*, IMA Journal of Mathematical Control and Information **16** (1999), 335–351.

[RV00] ______, *Degenerate optimal control problems with state constraints*, SIAM Journal of Control and Optimization **39** (2000), 989.

[Sar00] R.W.H. Sargent, *Optimal control*, Jounal of Computational and Applied Mathematics **124** (2000), 361–371.

[ST00] A.V. Sarycghev and D.F.M. Torres, *Lipschitzian regularity of minimizers for optimal control problems with control-affine dynamics*, Applied Mathematics and Optimization **41** (2000), no. 2, 237–254.

[Tor02] D. F. M. Torres, *Regularidade dos minimizantes no cálculo das variações e controlo Óptimo*, Ph.D. thesis, Departamento de Matemática, Universidade de Aveiro, Aveiro, Portugal., 2002.

[Vin00] R. Vinter, *Optimal control*, Birkhauser, Boston, 2000.

[VP82] R. B. Vinter and G. Pappas, *A maximum principle for nonsmooth optimal-control problems with state constraints*, Journal of Mathematical Analysis and Applications **89** (1982), 212–232.

[VZ98] R. B. Vinter and H. Zheng, *Necessary conditions for optimal control problems with state constraints*, Transactions of the American Mathematical Society **350** (1998), no. 3, 1181–1204.

[War72] J. Warga, *Optimal control of differential and functional equations*, Academic Press, New York, 1972.