FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

# PHD RESEARCH PLAN

**Md. Haider Ali Biswas**

**PDEEC Student ID. 080535022**

**Academic Year 2009–2010**

Universidade do Porto
FEUP Faculdade de Engenharia

**Faculty of Engineering, University of Porto, Portugal**

**Doctoral Program in Electrical and Computer Engineering**

This report is on the research plan of my PhD thesis prepared to be presented to the *Supervising Committee* of the PDEEC so as to comply with the *PDEEC norms* of academic excellence.

I take the opportunity to thank my supervisor, Maria do Rosario de Pinho, and the Doctoral Program in Electrical and Computer Engineering (PDEEC), of the Department of Electrical and Computer Engineering at Faculty of Engineering, University of Porto, Portugal for the partial grant I received when starting this PhD work. The financial support from the Foundation for Science and Technology (FCT), Portugal with the grant Ref. SFRH/BD/63707/2009 is also greatly acknowledged.

# Contents

# Chapter 1

# Introduction

This report deals with the up to date development of my ongoing PhD research on ***"Necessary Conditions for Optimal Control Problems with State Constraints: Theory and Applications"***. We present here a scenario for my PhD Thesis and the *state-of-art* of our work. In this report we review the very basic notions of optimal control problems with and without state constraints focussing on necessary conditions of optimality for state constrained problems.

## 1.1 Motivation

Necessary Conditions of Optimality (NCO) are elegant method to characterize and find solutions to optimal control problems. The main purpose of necessary conditions of optimality is to identify a *small* set of candidates to local minimizers among the overall set of admissible solutions. It is thus natural the interest to construct NCO *as strong as possible* that further reduce the set of identified candidates to a small set while still identifying all the local minimizers. The first complete and significant version of NCO for the optimal control problems was proposed and proved by Russian mathematician L. Pontryagin and his collaborators in 1950s in the form of *Maximum Principles* (MPs)[53]. This first version of MPs was derived under *smoothness* assumptions for the OCPs. Later on, a large number of modified, extended and generalized versions of MPs have been developed by several authors and in the late 1970s, extensions in the form of *Nonsmooth Maximum Principles* (NMPs) were proposed by F. Clarke [12].

Necessary Conditions of Optimality for Optimal Control Problems with *State Constraints* have attracted attention since probably the early stage of the control research, specially because of the inter-related applications in diverse engineering fields, for examples, aerospace and robotics, process engineering and economics, biology and medicine. Almost in all engineering applications, state constraints are imposed to have the desired output from the optimal control problems in questions. Different versions of MPs and even strengthened versions of MPs for state constrained optimal control problems both for smooth and nonsmooth cases have been developed over the years.

Our purpose in this thesis is firstly to study the up to date developments in this topic for deriving new set of necessary conditions of optimality for optimal control problems with state constraints and secondly implement our newly developed theories to some real life problems. Throughout this thesis, we mainly focus on the fixed time problems. We summarize here our main contributions of this thesis:

- New nonsmooth necessary conditions of optimality for state constrained problems are obtained under some convexity assumptions.
- Previous result is validated and improved when the convexity is removed.
- A new nonsmooth maximum principle is then obtained for problems with additional mixed constraints. Convex case is already studied. Removal of convexity is under study.

In our development, we generalize the work of Clarke and de Pinho [18] to state constraints ( see a discussion on our results in section 3.6 of Chapter 3).

## 1.2 Organizational Overview

This report is organized in the following manner. In **Chapter 2**, we introduce a general review on the literatures of optimal control problems.

**Chapter 3** presents new results on nonsmooth maximum principle for optimal control problems with state constraints.

In **Chapter 4**, we also provide a nonsmooth maximum principles considering both the pure state and mixed constraints [6] under convexity assumption. We hope to extend our results to the nonconvex case later on.

In **Chapter 5**, we discuss another crucial part of this PhD thesis, the *Application* of optimal control to real life problems, specially in epidemiology and medicine. We introduce an existing infectious disease model, namely the HIV model. Our challenge is to reformulate the problem replacing "soft state constraints" by "hard state constraints". To such new model we will study the problem both in analytical and numerical methods comparing the results with existing literatures. This is an example of a problem we aim to study both analytically and numerically. At this stage we cannot be sure that such problem will be the chosen one at the end of this thesis.

Finally in **Chapter 6**, we conclude this report by providing a summary of the *present contributions* of our work and the *future plan* of our research which will be carried out within the rest of the time frame. We also pose some related open questions concerning the state constrained optimal control problems and their applications to the real world problems in the field of epidemiology and medicine to motivate further research.

# Chapter 2

# Optimal Control Problems–Literature Review

It is commonly accepted that optimal control theory was born with the publication of a seminal paper by Pontryagin and collaborates last century, at the end of 50s. Since then optimal control theory has played a relevant role not only in the dynamic optimization but also in the control and system engineering. Another crucial moment in this theory is closely related with the development of nonsmooth analysis during the 70s and 80s. Nonsmooth analysis has triggered a new interest in optimal control problems and thus "brought new solutions to old problems" [56]. Nowadays optimal control theory is essential to different areas like engineering, economics and biology since many problems are modeled as optimal control problems.

Nowadays Optimal Control is an independent field of research. The development of optimal control has gained strength by treating multi-variable, time varying systems, as well as many nonlinear problems arising in control engineering. Several authors contributed to the basic foundation of a very large scale research effort initiated in the end of the 1950s, which continues to the present day. The development of Nonsmooth Analysis (see for example, [12] and [56]) has enhanced a wide scope of research as well as it has opened a new horizon in the optimal control theory. A challenging area of study in this theory remains that of state constraints. Necessary conditions of optimality for optimal control problems with state constraints have been studied since the very beginning of optimal control theory [53]. In spite of all the recent developments, this subject is far from explored. Throughout this report (as with our thesis) we focus on 'fixed time' optimal control problem with state constraints.

## 2.1 Preliminaries and Notations

Before proceeding we need to introduce some definitions and notations that will be used throughout the text.

For $g$ in $\mathbb{R}^m$, inequalities like $g \leq 0$ are interpreted componentwise. Here and throughout this report, $\mathbb{B}$ represents the closed unit ball centered at the origin regardless of the dimension of the underlying space and $|\cdot|$ the Euclidean norm or the induced matrix norm on $\mathbb{R}^{p\times q}$. The *Euclidean distance function* with respect to a given set $A \subset \mathbb{R}^n$ is

$$d_A\colon \mathbb{R}^n \to \mathbb{R}, \qquad y \mapsto d_A(y) = \inf \ \{|y - x| \ : \ x \in A\}\,.$$

A function $h\colon [a, b] \to \mathbb{R}^p$ lies in $W^{1,1}([a, b]; \mathbb{R}^p)$ if and only if it is absolutely continuous; in $L^1([a, b]; \mathbb{R}^p)$ iff it is integrable; and in $L^\infty([a, b]; \mathbb{R}^p)$ iff it is essentially bounded. The norm of $L^\infty([a, b]; \mathbb{R}^p)$ is $\|\cdot\|_\infty$.

The effect of state constraints in the optimal control problems is the introduction of measures as multipliers. These multipliers associated with state constraints are the elements of the topological dual space. The space $C^*([a, b]; \mathbb{R})$ is the topological dual of the space of continuous functions $C([a, b]; \mathbb{R})$. Elements of $C^*([a, b]; \mathbb{R})$ can be identified with finite regular measures on the Borel subsets of [a,b]. The set of elements in $C^*([a, b]; \mathbb{R})$ taking nonnegative values on nonnegative-valued functions in $C([a, b]; \mathbb{R})$ is denoted by $C^\oplus([a, b]; \mathbb{R})$. The norm in $C^\oplus([a, b]; \mathbb{R})$, $|\mu|$, coincides with the total variation of $\mu$, $\int_{[a,b]} \mu(ds)$. The support of a measure $\mu$, written as $\operatorname{supp}\{\mu\}$, is the smallest closed set $A \subset [a, b]$ such that for any relatively open subset $B \subset [a, b]\backslash A$ we have $\mu(B) = 0$. More discussions on measures can be obtained in [56].

## 2.2 Optimal Control Problems

Since the birth of the optimal control, several authors proposed different basic mathematical formulations of OCPs (fixed time problems). For fixed time problems there are three major mathematical formulations of optimal control problems: *Bolza form*, *Lagrange form* and *Mayer form*.

We start with the general form of *Bolza* problem:

$$
(P_B) \quad \begin{cases} \text{Minimize } J(x,u) = l(x(a),x(b)) + \displaystyle\int_a^b L(t,x(t),u(t))dt \\ \text{subject to} \\ \ \dot{x}(t) = f(t,x(t),u(t)) \quad \text{a.e. } t \in [a,b] \\ \ u(t) \in U(t) \quad \text{ a.e. } \ t \in [a,b] \\ \ (x(a),x(b)) \in E. \end{cases}
$$

Here $[a,b]$ is a fixed interval. The function $f\colon [a,b]\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}^n$ describes the system dynamics and $U\colon [a,b] \to \mathbb{R}^m$ is a multifunction. Furthermore, the closed set $E \subset \mathbb{R}^n\times\mathbb{R}^n$ and the functions $l\colon \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ and $L\colon [a,b]\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}$ specify the endpoint constraints and cost. The functional

$$
l(x(a),x(b)) + \int_a^b L(t,x(t),u(t))dt \tag{2.1}
$$

to be minimized is called the *payoff* or *cost* functional. The aim of this problem is to find the pair $(x,u)$ comprising two functions where $u\colon [a,b] \to \mathbb{R}^m$ (the control function) and the corresponding state trajectory $x$ which is an absolutely continuous function $x\colon [a,b] \to \mathbb{R}^n$ (called the state function) satisfying all the constraints of the problem $(P_B)$ and minimizing the cost. A pair $(x,u)$ where $x$ is an absolutely continuous function and $u$ is a function belonging to a certain space $\mathcal{U}$ ($\mathcal{U}$ can be $L^1$, $C$, the space of measurable functions, the space of piecewise continuous functions, etc.) such that $\dot{x}(t) = f(t,x(t),u(t))$ a.e. is called a *process*. A '*process*' satisfying all the constraints of the problem $(P_B)$ is called an *admissible process*. We say that $(x^*,u^*)$ is an *optimal solution* if it minimizes the cost over all admissible processes. For optimal control problems one may speak of local or global minimizers. In this thesis we focus on local minimizers. Local minimizers can be of different types as we will see in section 2.4.

If the function $l\colon \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ is absent from the cost functional (2.1) and all others data remain the same, we obtain the optimal control problem in *Lagrange form*; the cost is simply

$$
J(x,u) = \int_a^b L(t,x(t),u(t))dt \tag{2.2}
$$

On the other hand, if the Lebesgue integrable function $L\colon [a,b]\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}$ is absent from the cost functional (2.1) and all others constraints remain the same, we obtain the *Mayer form* with cost

$$
J(x,u) = l(x(a),x(b)) \tag{2.3}
$$

However, we can reformulate Bolza form (2.1) into Mayer form by *state augmentation*. Let us define,

$$\begin{aligned} &\dot{y}(t) = L(t, x(t), u(t)) \quad \text{a.e.} \\ &y(a) = 0. \end{aligned} \tag{2.4}$$

Then the problem $(P_B)$ can be rewritten as following

$$(P_M) \quad \begin{cases} \text{Minimize } J(x,u) = l(x(a), x(b)) + y(b) \\ \text{subject to} \\ \ \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [a,b] \\ \ \dot{y}(t) = L(t, x(t), u(t)) \quad \text{a.e. } t \in [a,b] \\ \ u(t) \in U(t) \quad \ \text{a.e.} \quad t \in [a,b] \\ \ ((x(a), x(b)), y(a)) \in E \times \{0\}. \end{cases}$$

This new problem $(P_M)$ is now in Mayer form. We refer readers [5, 11, 22, 43, 44] for more extensive studies on the transformations of optimal control problems from the Bolza form to the other two special forms along with their examples.

Different variants of optimal control problems appear in the control system dynamics over the years. The problems we have mentioned here are fixed time problems (since the time interval $[a, b]$ is fixed). The other problems as free time problems, minimum time problems as well as impulsive control problems and others are out of the scope of this thesis.

## 2.3 Constrained Optimal Control Problems

Optimal control problems (OCPs) have become challenging because of imposing a great variety of *physical constraints*. These constraints restrict the range of values of both the control and the state variables. When the pathwise constraints are imposed on the state trajectories of the optimal control problems in question, such types of problems are called the *state constrained optimal control problems*. In most cases, such constraints take the form of *scalar functional inequality constraints* because these kinds of state constraints are frequently encountered in engineering applications. Also such constraints are convenient to derive necessary conditions for other types of state constraints (multiple state constraints, implicit state constraints, etc.) of interest.

Next we briefly review the different constraints usually imposed to optimal control problems in many engineering applications. As we mentioned before, we focus only on constraints for fixed time problems.

### 2.3.1 Endpoint Constraints

The constraints which are usually imposed at the initial point and/or terminal point of a fixed interval $[a, b]$ are called *endpoint constraints*. The most general way of writing this constraint is

$$(x(a), x(b)) \in E \tag{3.5}$$

This includes many types of constraints. Suppose for example,

$$(i) \begin{cases} x(a) = x_a \\ x(b) \in \mathbb{R}^n \end{cases}$$

then $E = \{x_a\} \times \mathbb{R}^n$ and (3.5) $\Leftrightarrow (x(a), x(b)) \in E$.

Again endpoint constraints of the form

$$(ii) \begin{cases} x(a) = x_a \\ x(b) \in E_b \end{cases}$$

where $E_b$ may be a point or nonempty set are common in applications.

Moreover we can have endpoint constraints in the form of equality or inequality or both. i.e., We can have $\phi_1(x(a), x(b)) = 0$ and/or $\phi_2(x(a), x(b)) \leq 0$. These constraints can be written into the inclusion form by the right choice of the set $E$.

### 2.3.2 Control Constraints

The constraints which are imposed on the control variables of an optimal control problem are treated as *control constraints*. In this case the control variables are restricted to take values in a certain set. For example, $u(t) \in U(t)$ is called control constraints, where $u$ takes values from a measurable multifunction $U : [a, b] \to \mathbb{R}^m$.

### 2.3.3 Pathwise Constraints

The pathwise constraints encountered in many optimal control problems may be defined for restricting the range of values taken by mixed functions of both the control and the state variables and such restrictions can be imposed over the entire time interval $[a, b]$ or any (nonempty) time subinterval.

Let us discuss here some common pathwise constraints appearing in the literature. In general, such constraints can be written as

$$(x(t), u(t)) \in C(t) \text{ for all } t \in [a, b]$$

where $C : [a, b] \to \mathbb{R}^n \times \mathbb{R}^m$ is a given multifunction. In the literature, however, we encounter explicit constraints of the form:

***A. Equality state constraint:*** Let $h\colon [a, b] \times \mathbb{R}^n \to \mathbb{R}$ be a given function. Then

$$h(t, x(t)) = 0 \text{ for all } t \in [a, b]$$

is called an *equality state constraint*.

***B. Inequality state constraint:*** Let $h\colon [a, b] \times \mathbb{R}^n \to \mathbb{R}$ be a given function. Then

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

is called an *inequality state constraint*.

***C. Implicit state constraint:*** Let $X\colon [a, b] \to \mathbb{R}^n$ be a given multifunction. Then

$$x(t) \in X(t) \text{ for all } t \in [a, b]$$

is called an *implicit state constraint*.

***D. Mixed inequality state-control constraint:*** Let $g\colon [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ be a given function. Then

$$g(t, x(t), u(t)) \leq 0 \text{ a.e. } t \in [a, b]$$

is called a *mixed inequality state-control constraint*.

***E. Mixed equality state-control constraint:*** Let $b\colon [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ be a given function. Then

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

is called a *mixed equality state-control constraint*.

Usually one refers to constraints of types $A$, $B$ and $C$ as pure state constraints to highlight the difference with those in the form $D$ and $E$ which are mixed constraints. Observe that state constraints are imposed *for all* $t$ in an interval $[a, b]$ or any subinterval of $[a, b]$ while mixed constraints can be imposed simply for *almost every* $t$.

Focussing on the first three types of constraints $A$, $B$ and $C$, it is obvious that type $C$ is the more general (see in this respect the discussion in [56], Chapter 9). Constraints of type $A$ and $B$ can be written as

$$x(t) \in X(t) \ \text{ for all } t \in [a,b],$$

where $X(t) := \{x \in \mathbb{R}^n : h(t,x(t)) = 0\}$ or $X(t) := \{x \in \mathbb{R}^n : h(t,x(t)) \leq 0\}$.

Observe that in fact $C$ can be viewed as a special case of

$$(x(t),u(t)) \in C(t)$$

when $C(t) = X(t) \times \mathbb{R}^m$, if no set constraints are imposed on the control or $C(t) = X(t) \times U(t)$ when we impose $u(t) \in U(t)$. On the other hand, $D$ and $E$ can also be written as

$$(x(t),u(t)) \in C(t)$$

where

$$C(t) = \{(x,u) \in \mathbb{R}^n \times \mathbb{R}^m : g(t,x(t),u(t)) \leq 0\} \ \text{ or } C(t) = \{(x,u) \in \mathbb{R}^n \times \mathbb{R}^m : b(t,x(t),u(t)) = 0\}.$$

or more generally

$$C(t) = \{(x,u) \in \mathbb{R}^n \times \mathbb{R}^m : g(t,x(t),u(t)) \leq 0, \quad b(t,x(t),u(t)) = 0\}.$$

Moreover, control set constraints of the form $u(t) \in U(t)$ can also be incorporated in the multifunction $C(t)$ by defining $C(t)$ to be

$$C(t) = \{(x,u) \in \mathbb{R}^n \times \mathbb{R}^m : g(t,x(t),u(t)) \leq 0, \quad b(t,x(t),u(t)) = 0, \ u \in U(t)\}.$$

See [18] for even more general constraints.

It is customary in the literature to use the inclusion

$$(x(t),u(t)) \in C(t)$$

when we are in the presence of mixed state-control constraints using $x(t) \in X(t)$ for pure state constraints.

From the point of view of optimality conditions state constraints and mixed constraints have different treatments. Necessary conditions for problems with constraints of type $D$ can be obtained when some constraint qualifications (also called in this case regularity conditions) are imposed. Such constraint qualifications involve the control variable. Clearly such constraint qualifications do not make any sense when state constraints are presented since the state constraints exhibit no dependence on the control variable. In some situations pure state constraints of type $B$ and mixed constraints $D$ can be related. That may occur when the function $h$ can be differentiated with respect to $t$ so as to obtain higher order derivatives containing the control variable. See [32] for more discussion on this issue and the relevant recursive procedure of their inter-relations. Before finishing this section we give an example of state constraints appearing in application. Consider the modeling of the temperature of a reactor. Taking the state to be the temperature, $x(t)$ it is natural impose an upper limit $M$ to this variable. This gives rise to the state constraint

$$x(t) \leq M \ \forall t \in [a,b] \text{ and } M \in \mathbb{R}.$$

Thus state constraints are natural features in many practical applications of optimal control problems.

## 2.4 Types of Local Minimizer

Minimizers are the *solutions* of Optimal Control Problems (OCPs)) of interests and thus play the most vital roles in finding the optimal trajectories with respect to control variables of the problems. In the theory of optimal control minimizers are classified into two categories: *global and local.* Suppose for example, we want to find the minimizers of the problem

$$\begin{array}{ll} \text{Minimize} & f(x) \\ \text{subject to} & x \in \mathbb{R}^n \end{array} \tag{4.6}$$

Then $x_G^*$ will be a *global minimizer* of (4.6), if it minimizes the cost over all other $x \in \mathbb{R}^n$, i.e.

$$f(x_G^*) \leq f(x) \quad \forall x \in \mathbb{R}^n,$$

and $x_L^*$ will be a *local minimizer* of (4.6), if it minimizes the cost over all other $x$ in some neighborhood, i.e. there exists $\varepsilon > 0$ such that

$$f(x_L^*) \leq f(x) \quad \forall x \in \mathbb{B}(x_L^*; \varepsilon).$$

It was shown in [31] that in some cases the *local minimizers* are *global minimizers.* Furthermore, all global minimizers are local minimizers. We refer readers to [56, 57] for study on minimizers. Throughout this thesis, we will restrict our discussion to minimizers in the context of optimal control problems to those of local minimizers.

### 2.4.1 Strong Local Minimum

The process $(x^*, u^*)$ is a *strong local minimum* for an optimal control problem if, for some $\varepsilon > 0$, it minimizes the cost over all other admissible processes $(x, u)$ such that

$$\| x - x^* \|_\infty := \max_{t \in [a,b]} |x(t) - x^*(t)| \leq \varepsilon \quad \text{for all} \quad t \in [a, b].$$

**Remark:** We have $\| x \|_\infty := \operatorname{esssup}_{t \in [a,b]} |x(t)|$, so

$$\| x \|_\infty \quad \leq \varepsilon \Rightarrow |x(t)| \leq \varepsilon \text{ a.e. } t \in [a, b].$$

But if $x$ is continuous, then

$$\operatorname{esssup}_{t \in [a,b]} |x(t)| = \max |x(t)|.$$

So

$$|x(t) - x^*(t)| \leq \varepsilon \quad \text{for all} \quad t \in [a, b] \Leftrightarrow \max |x(t) - x^*(t)| \leq \varepsilon.$$

### 2.4.2 Weak Local Minimum

An *admissible process* $(x^*, u^*)$ for an optimal control problem is called a *weak minimizer* if there exists $\epsilon > 0$ such that $l(x^*(a), x^*(b)) \leq l(x(a), x(b))$ holds for all process $(x, u)$ satisfying the following conditions:

$$|x(t) - x^*(t)| \leq \varepsilon \quad \forall t \in [a, b]$$

and

$$|u(t) - u^*(t)| \leq \varepsilon \text{ for almost every } t \in [a, b].$$

### 2.4.3 $W^{1,1}$ Local Minimum

The process $(x^*, u^*)$ is a $W^{1,1}$ *local minimum* for an optimal control problem if, for some $\varepsilon > 0$, it minimizes the cost over all other admissible processes $(x, u)$ such that

$$\| x - x^* \|_\infty \quad \leq \varepsilon, \quad \text{and} \quad \int_a^b |\dot{x}(t) - \dot{x}^*(t)| dt \leq \varepsilon.$$

That is,

$$\| x - x^* \|_{W^{1,1}} = |x(a) - x^*(a)| + \| \dot{x}(t) - \dot{x}^*(t) \|_{L^1}$$
$$= |x(a) - x^*(a)| + \int_a^b |\dot{x}(t) - \dot{x}^*(t)| dt \leq \varepsilon.$$

### 2.4.4 Local Minimum of Radius $R$

Let us define a measurable function $R : [a, b] \to (0, +\infty]$ which is called a *radius function*. The process $(x^*, u^*)$ is a *local minimum of radius* $R$ for an optimal control problem if, for some $\varepsilon > 0$, it minimizes the cost over all other admissible processes $(x, u)$ satisfying

$$\| x - x^* \|_\infty \quad \leq \varepsilon \quad , \quad \int_a^b |\dot{x}(t) - \dot{x}^*(t)| dt \leq \varepsilon,$$

as well as

$$|u(t) - u^*(t)| \leq R(t), \quad \text{a.e.} \quad t \in [a, b].$$

***Remarks:*** A relation between strong and weak local minimizers is that strong local minimizer is always a weak local minimizer but the converse is not necessarily true [57]. It is worth mentioning the fact that throughout this thesis, we will mainly work with *strong minimizers* to derive the optimality conditions and when possible will extend them to $W^{1,1}$ *local minimizers*. The reasons behind choosing this $W^{1,1}$ minimizers is that the class of $W^{1,1}$ local minimizers is larger than the class of $L^\infty$ minimizers. It follows that, by choosing to work with $W^{1,1}$ local minimizers, we are carrying out a sharper analysis of the local nature of the Maximum Principle than would be the case if we chose to derive conditions satisfied by strong local minimizers.

## 2.5 Nonsmooth Analysis and Optimal Control

*Nonsmooth Analysis* deals with the local approximation of non-differentiable functions and of sets with non-differentiable boundaries; consequently it can be treated as the branch of well known *Nonlinear Analysis* [13]. In the last few decades this field has grown up very exceedingly because of the recognition of nondifferentiable phenomena. The field of nonsmooth optimization is significant, not only because of the existence of non-differentiable functions arising directly in applications, but also because several important methods for solving difficult smooth problems lead directly to the need to solve nonsmooth problems, which are either smaller in dimension or simpler in structure.

This has provided a source of much research over the years, the fruits of which have been the birth of *nonsmooth analysis*. See for examples ([12], [13] and [56]) for more details about the background and importance of nonsmooth analysis in optimal control theory.

There has been a sustained and fruitful interaction between Nonsmooth Analysis and Optimal Control. Now-a-days nonsmooth analysis is an important tool in the development of optimal control theory.

### 2.5.1 Fundamentals of Nonsmooth Analysis

In the *classical* sense, derivatives of a function $f$ are related to normal vectors to tangent hyperplanes; for any differentiable function $f$ the vector $(f'(x), -1)$ is a downward normal to the graph of $f$ at $(x, f(x))$. The graph of $f$ is defined by $\text{Gr}f = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} : \alpha = f(x)\}$. This geometric relationship is the key for the development of nonsmooth analysis. Instead of considering derivatives as elements of normal subspaces to smooth sets, generalized derivatives are defined to be elements of normal cones to possibly nonsmooth sets. Let us start with some definitions that will be useful to introduce nonsmooth calculus.

Let $A \subset \mathbb{R}^n$ be a nonempty closed set with $x \in \mathbb{R}^n \backslash A$. We call $y$ the *closest point* in $A$ or $Proj_A(x)$ (i.e. the projection of $x$ onto $A$) (see Figure 2.1) if $y \in \text{Proj}_A(x)$ such that

$$\| x - y' \| \geq \| x - y \|, \ \forall y' \in A$$

which is equivalent to write

$$\langle \omega, y' - y \rangle \leq \sigma \| y' - y \|^2, \ \forall y' \in A \ \text{ and some } \sigma > 0,$$

where the vector $\omega = x - y$ is perpendicular to $A$ at $y$.

Any nonnegative multiple $\zeta = t\omega, \ t > 0$ of $\omega$ is called a *proximal normal* vector (see [56]. That is, a *vector* $\zeta$ is called a *proximal normal* to $A$ at $y$ iff for some $\sigma > 0$ the following *proximal normal inequality* holds:

$$\langle \zeta, y' - y \rangle \leq \sigma \| y' - y \|^2, \ \forall y' \in A.$$

The set of all such vectors, which is a convex cone [18] containing 0 is denoted by $N_A^P(y)$ and is referred to as the *Proximal Normal Cone*.

A vector $\zeta$ is called the *limiting normal to $A$ at $x$* if for each $i \in \mathbb{N}$,

$$\zeta = \lim \zeta_i, \ \forall \zeta_i \in N_A^P(x_i), \ x_i \in A, \ x_i \to x,$$

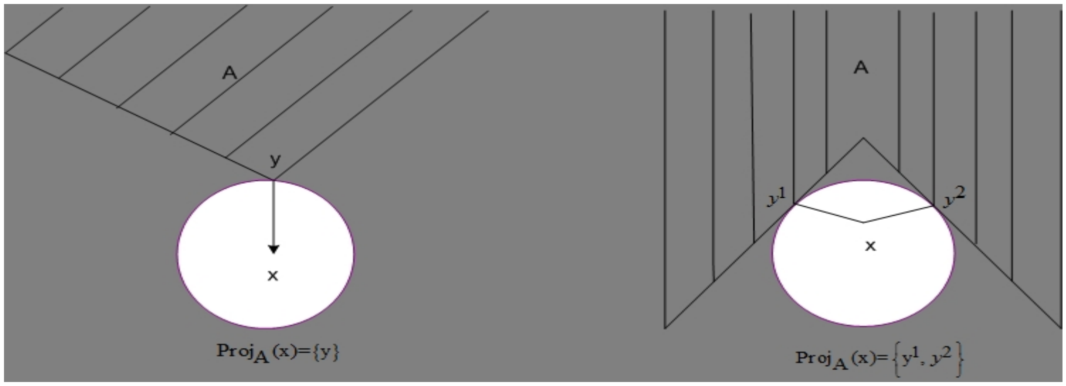


Figure 2.1: Geometrical interpretation of proximal normal and limiting normal cones.

and the set of all such limiting normals, denoted by $N_A^L(x)$ is a cone, called the *Limiting Normal Cone* to $A$ at $x$.

Given a lower semicontinuous function $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and a point $x \in \mathbb{R}^n$ where $f(x) < +\infty$ such that $\mathrm{dom} f = \{x : f(x) < +\infty\}$, then the *proximal subdifferential* (or set of all *proximal subgradients*) of $f$ at $x \in \mathrm{dom} f$ is defined as the set

$$\partial^P f(x) := \{\zeta \in \mathbb{R}^n : (\zeta, -1) \in N^P_{\mathrm{epi} f}(x, f(x))\}.$$

where $\mathrm{epi} f = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} : \alpha \geq f(x)\}$ denotes the epigraph of a function $f$. The *limiting subdifferential* (or set of all *limiting subgradients*) of a function $f$ at $x \in \mathrm{dom} f$ denoted by $\partial^L f(x)$ is obtained by the set

$$\partial^L f(x) := \{\zeta \in \mathbb{R}^n : (\zeta, -1) \in N^L_{\mathrm{epi} f}(x, f(x))\}.$$

However, the nonsmooth calculus can be developed via the theory of *generalized gradients* in the context of *locally Lipschitz function.* If a function $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz near $x$, then the *generalized gradients* $\partial^C f(x)$ coincides with co $\partial^L f(x)$ (*convex hull* of $\partial^L f(x)$); also in similar fashion, the associated normal cone $N_A^C(x)$ to a set $A$ at a point $x$ coincides with $\overline{\mathrm{co}}\, N_A^L(x)$. This *generalized gradients* and its calculus were first defined by Clarke in 1973 [12], so $\partial^C f(x)$ and $N_A^C(x)$ are also called the *Clarke subdifferential* and the *Clarke normal cone* respectively. For more details on such nonsmooth analysis concepts and generalized gradients as well as its basic calculus, we refer readers for example to [12, 13, 47, 56].

## 2.6 The Maximum Principles

The Maximum Principle (MP) is one of the most elegant techniques used to solve the OCPs. It provides a set of necessary conditions which should be satisfied by any optimal solution of optimal control problem. Not surprisingly the idea behind derivation of necessary conditions in the form of Maximum Principles is to obtain MPs that produces the smallest set of candidates to the optimal control problems. It is well known that for some problems the classical Maximum Principle is not only a necessary condition of optimality but also a sufficient condition (for a discussion on this feature in a smooth and nonsmooth context see [22]).

The maximum principle is a milestone in the development of modern optimal control theory. Maximum principle plays significant role not only in solving the smooth problems, but also in problems with nonsmooth functions. When the data of the problem are smooth, we call the corresponding maximum principle smooth, but for the problems with nonsmooth data we call it nonsmooth maximum principle. Here we will present variants of maximum principles both for smooth and nonsmooth optimal control problems with and without state constraints.

### 2.6.1 Smooth Maximum Principles

The 'first version' as well as the widely accepted Pontryagin maximum principle was proved for the optimal control problems with 'smooth' dynamic constraints. This *smooth version* of maximum principles have been the basic foundation of all other extended versions of MPs over the years. To illustrate the smooth maximum principle, we consider the optimal control problem with *state constraints*,

$$(OCP) \quad \begin{cases} \text{Minimize } J(x,u) = l(x(a), x(b)) \\ \text{subject to} \\ \quad \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [a,b] \\ \quad h(t, x(t)) \leq 0 \quad \text{for all } \ t \in [a,b] \\ \quad u(t) \in U(t) \quad \text{a.e. } \ t \in [a,b] \\ \quad (x(a), x(b)) \in E. \end{cases}$$

We assume for the time being that the state constraint $h(t, x(t)) \leq 0$ is *absent* from the problem. The smooth maximum principle, which we present now, is valid under smooth assumptions on the data. Here, and for simplicity, we consider that the functions $f$ , $l$ are all continuously differentiable. Observe that in (OCP) (and again for the sake of simplicity) we assume the multifunction $U$ to be constant (i.e., $U(t) = U$ ) and $U$ is a closed set. We define the *Pseudo-Hamiltonian* (or *Unmaximized*

*Hamiltonian* or *Pontryagin*) function

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

Now the *smooth maximum principle* for the problem (OCP) without state constraints under some appropriate assumptions can be presented in the next Theorem (an adaptation of Theorem 6.2.1 in [56]).

**Theorem 2.1** ***(The Maximum Principle for (OCP) Without State Constraints):*** Let $(x^*, u^*)$ be a strong local minimum for problem $(OCP)$ without state constraints. Then there exist an arc $p \in W^{1,1}([a,b];\mathbb{R}^n)$ and a scalar $\lambda_0 \geq 0$ satisfying
the *Nontriviality Condition* [NT]:

$$(p, \lambda_0) \neq (0,0),$$

the *Euler Adjoint Equation* [AE]:

$$-\dot{p}(t) = \nabla_x \langle p(t), f(t, x^*(t), u^*(t))\rangle \quad \text{a.e.},$$

the global *Weierstrass Condition* [W]:
$\forall \;\; u \in U(t),$

$$\langle p(t), f(t, x^*(t), u)\rangle \leq \langle p(t), f(t, x^*(t), u^*(t))\rangle \quad \text{a.e.},$$

and the *Transversality Condition* [T]:

$$(p(a), -p(b)) = \lambda_0 \nabla l(x^*(a), x^*(b)) + (\eta_1, \eta_2),$$

for some $(\eta_1, \eta_2) \in N_E^L(x^*(a), x^*(b))$.

The function $p$ is called the *costate* (or *adjoint*) function and $\lambda_0$ the *cost multiplier*. The adjoint equation is also called the costate differential equation.

However, we now turn to the more general case of the same problem (OCP) assuming that the state constraint is imposed. We note that the effect of state constraints is the introduction of *measures* as multipliers. The adjoint multiplier $p$ is then replaced by a function $q$ of *bounded variation* defined by,

$$q(t) = \begin{cases} p(t) + \int_{[a,t)} \gamma(s)\mu(ds) & t \in [a,b) \\ p(t) + \int_{[a,b]} \gamma(s)\mu(ds) & t = b. \end{cases} \tag{6.7}$$

Let us assume again that the functions $f$, $l$ and $h$ are all continuously differentiable and as before, that $U$ is a closed set. Then the *smooth maximum principles* for the *state constrained optimal control problems* can be adapted in the following Theorem (an adaptation of Theorem 9.3.1 in [56]).

**Theorem 2.2** ***(The Maximum Principle for (OCP) With State Constraints)****:* Let $(x^*, u^*)$ be a strong local minimum for problem $(OCP)$ with state constraints. Then there exist an arc $p \in W^{1,1}([a,b];\mathbb{R}^n)$, a scalar $\lambda_0 \geq 0$, $\mu \in C^{\oplus}([a,b])$, and a measurable function $\gamma(t) : [a,b] \to \mathbb{R}^n$ such that the following conditions are satisfied:

*(i) The Nontriviality Condition* [NT]:

$$(p, \mu, \lambda_0) \neq (0, 0, 0)$$

*(ii) The Euler Adjoint Equation* [AE]:

$$-\dot{p}(t) = \nabla_x \langle q(t), f(t, x^*(t), u^*(t)) \rangle \ \text{ a.e.,}$$

*(iii) The Weierstrass Condition* [W]:

$\forall \ u \in U(t),$

$$\langle q(t), f(t, x^*(t), u) \rangle \leq \langle q(t), f(t, x^*(t), u^*(t)) \rangle \ \text{ a.e.,}$$

*(iv) The Transversality Condition* [T]:

$$(p(a), -q(b)) = \lambda_0 \nabla l(x^*(a), x^*(b)) + (\eta_1, \eta_2),$$

for some $(\eta_1, \eta_2) \in N_E^L(x^*(a), x^*(b))$,

*(v)* : $\operatorname{supp}\{\mu\} \subset I(x^*)$,

where $I(x^*) := \{t : h(t, x^*(t)) = 0\}$ and $q$ is as in (6.7).

We note that the maximum principle in Theorem 2.2 is of interest only when the state constraint is *nondegenerate*. We briefly discuss this nondegeneracy issue in Section 2.7.

### 2.6.2 Nonsmooth Maximum Principles

The *nonsmooth maximum principles* have been one of the central attractions of the nonsmooth optimization problems for a long time. In the 1970s F. Clarke generalized the convex subdifferentials of Rockafellar to cover Lipschitz continuous functions and to some extent, lower semi-continuous functions (see, for example [12]). He also successfully applied nonsmooth analysis to optimization and optimal control theory. In 1976s Mordukhovich proposed the concept of limiting subdifferential and he showed how transversality conditions in the nonsmooth maximum principle could be weakened. We now discuss here the *nonsmooth maximum principle* for optimal control problems with state constraints. We consider again our problem (OCP) with state constraints, but we will impose the following hypotheses which make reference to an optimal solution $(x^*, u^*)$ and a parameter $\varepsilon > 0$:

**(NH1):** The function $(t,u) \to f(t,x,u)$ is $\mathcal{L} \times \mathcal{B}$ measurable and there exist $\varepsilon > 0$ and an integrable function $k(t)$ such that, for almost every $t \in [a,b]$ the following condition holds:

$$|f(t,x_1,u) - f(t,x_2,u)| \leq k(t)\|x_2 - x_1\|, \ \forall u \in U(t), \ (x_1,x_2) \in \mathbb{B}(x^*,\varepsilon).$$

**(NH2):** $l$ is Lipschitz near $(x^*(a), x^*(b))$ with Lipschitz constant $K_l$.

**(NH3):** $h$ is upper semicontinuous and for each $t \in [a,b]$ the function $h(t,\cdot)$ is Lipschitz on $x^*(t) + \mathbb{B}(0,\varepsilon)$ with Lipschitz constant $K_h$.

**(NH4:)** $GrU$ is a Borel set.

where $GrU$ is the graph of the multifunction $U : [a,b] \to \mathbb{R}^m$ defined by

$$GrU := \{(t,u) \in [a,b] \times \mathbb{R}^m : u \in U(t)\}.$$

**Theorem 2.3** ***(The Nonsmooth Maximum Principle for (OCP) With State Constraints)****:* Let $(x^*, u^*)$ be a strong local minimum for problem $(OCP)$ with state constraints. and assume that hypotheses (NH1)–(NH4) are satisfied. Then there exist an arc $p \in W^{1,1}([a,b];\mathbb{R}^n)$, a scalar $\lambda_0 \geq 0$, $\mu \in C^{\oplus}([a,b])$, and a measurable function $\gamma(t) : [a,b] \to \mathbb{R}^n$ satisfying $\gamma(t) \in \partial_x^{>} h(t,x^*(t)) \quad \mu - a.e.$ such that the following conditions are satisfied.

*(i) The Nontriviality Condition* [NT]:

$$(p, \mu, \lambda_0) \neq (0,0,0),$$

*(ii) The Euler Adjoint Equation* [AE]:

$$-\dot{p}(t) \in \partial_x^C \langle q(t), f(t,x^*(t),u^*(t))\rangle,$$

*(iii) The Weierstrass Condition* [W]:

$\forall \ \ u \in U(t),$

$$\langle q(t), f(t,x^*(t),u)\rangle \leq \langle q(t), f(t,x^*(t),u^*(t))\rangle \ \ \text{a.e.},$$

*(iv) The Transversality Condition* [T]:

$$(p(a), -q(b)) \in \lambda_0 \partial l(x^*(a), x^*(b)) + (\eta_1, \eta_2),$$

for some $(\eta_1, \eta_2) \in N_E^C(x^*(a), x^*(b))$,

*(v)* : $\quad \text{supp}\{\mu\} \subset I(x^*)$,

where $I(x^*) := \{t : h(t, x^*(t)) = 0\}$ , $q$ is as in (6.7), and the partial subdifferential $\partial_x^>$ is defined by

$$\partial_x^> h(t, x) := \text{co}\,\{\gamma : \exists (t_i, x_i) \xrightarrow{h} (t, x) : h(t_i, x_i) > 0\ \forall i,\ \nabla_x h(t_i, x_i) \to \gamma\}. \tag{6.8}$$

Several extended versions, and even more strengthened form of nonsmooth maximum principles for state constrained optimal control problems have been developed over the years. We refer readers [12, 15, 16, 21, 55] for the detailed presentations and to [17, 18] for the recent developments in the nonsmooth maximum principle.

## 2.7 Degeneracy in Optimal Control Problems

The *Degeneracy* phenomenon arising in optimal control problems with state constraints is one of the most troublesome issues while deriving the necessary conditions of optimality (NCO) for optimal control problems (OCPs). It was first pointed out in the Russian literature and has been the subject of subsequent extensive study over the last few decades ( see, for example [56]). We say that the Maximum principle degenerates when the necessary conditions supply no useful information about the minimizers. This happens for example when the state trajectory is fixed and lies on the boundary of the state constraint set. In these cases a nonzero, but nevertheless trivial, set of multipliers exists [54]. It was shown in [2] that this phenomenon occurs because of the incompleteness of the standard variants of Pontryagin's maximum principle for problems with state constraints.

To illustrate this situation, let us now suppose that the pair $(x^*, u^*)$ solve the problem $(OCP)$. Consider, for example,

$$E_a = \{x_a\} \text{ and } h(a, x_a) = 0$$

for some point $x_a \in \mathbb{R}^n$ such that $\nabla_x h(a, x_a) \neq 0$ (i.e., the case when the left endpoint is fixed at $x_a$ and $x_a$ lies in the boundary of the state constraint set). Then the necessary conditions in equations (i)–(v) of Theorem 2.3 are satisfied for *any* feasible process $(x^*, u^*)$ (local minimizer or not) with the nonzero multiplier set

$$\lambda = 0, \quad \mu = \delta_{\{a\}}, \quad \text{and} \quad p = -\nabla_x h(a, x_a), \tag{7.9}$$

where $\delta_{\{a\}}$ denotes the unit measure (or Dirac measure) concentrated at $\{a\}$. In this situation we say that the Maximum principle degenerates. This situation (i.e. $\lambda = 0$ ) is also treated as the *Abnormality* of the maximum principles.

There are several ways to overcome the degeneracy situation and to obtain the *Nondegenerate Maximum Principles* for constrained optimal control problems. The way of avoiding this degeneracy phenomenon is to force the cost functional to be involved in the maximum principles by adding a condition like $\lambda \neq 0$. One (among many others) and the simplest condition is imposed by choosing $\lambda = 1$ which is known as the famous and probably the most cited Kuhn-Tucker conditions [41]. Thus the maximum principles under the condition $\lambda = 1$ is called the *Normal form* and the strengthened forms of the NCO capable of avoiding some type of degeneracy are called *Nondegenerate forms*.

However, the arising degeneracy or abnormality (i.e. when $\lambda = 0$) in the problems is not usual, it is a special case. When this phenomenon arises, some condition like normality, $\lambda > 0$ should be imposed. The conditions that identify problems in which we can guarantee the existence of a nondegenerate or normal set of multipliers are usually refered to as *Constraint Qualifications (CQ)*. A detail study on how the degeneracy phenomenon occurs in optimal control problems and what are the different ways to tackle this situation for obtaining *nondegenerate maximum principles* can be found in[44]. We also refer readers to [1, 2, 8, 29, 45, 54, 56] for more survey as well as the recent developments on the degeneracy phenomenon.

# Chapter 3

# Nonsmooth Maximum Principles for State Constrained Problems

In this chapter we present the first theoretical result we hope to accomplish during the PhD thesis. We emphasize that these are now being finalized. Our first result is a new *Nonsmooth Maximum Principle* for state constrained problems. It breaks into two: the Nonsmooth Maximum Principle is proved in the convex case and then convexity is removed.

## 3.1 Optimal Control Problems with State Constraints

First we develop a nonsmooth maximum principle for optimal control problems with pure state constraints in the presence of a convexity assumption. To achieve our purpose we intertwine established approaches used for state constraints with up to date developments. Indeed, we follow closely the approach developed in [55] (see also [23] and [25]) where necessary conditions for pure state constrained problems are derived. Our proofs differ from those in [23] since we apply a new nonsmooth maximum principle for state constrained problems, derived in [18] (instead of results in [22] as in [23] ). There is however a price to pay; here we assume that the solution of $(P)$ is a $W^{1,1}$ strong minimum in contrast with [18] where a weaker notion of minimum, that of *local minimum of radius* $R$ is used (in this respect see also [14]). The convexity assumption we impose here may be seen as a major hindrance to some applications, but it will be removed later on.

For the convenience of the reader we define our problem again. Indeed, the problem of interest, denoted throughout by $(P)$ is that of minimizing the cost function

$$l(x(a), x(b)) + \int_a^b L(t, x(t), u(t))\ dt$$

subject to the differential equation

$$\dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e.} \quad t \in [a, b],$$

the state constraint

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

the boundary conditions

$$(x(a), x(b)) \in E,$$

and the control constraints

$$u(t) \in U(t) \quad \text{a.e.} \quad t \in [a, b].$$

Here the interval $[a, b]$ is fixed, the state $x(t) \in \mathbb{R}^n$ and the control $u(t) \in \mathbb{R}^m$. Thus the function $f$ describing the dynamics is $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. Moreover $h$ and $L$ are scalar functions $h : [a, b] \times \mathbb{R}^n \to \mathbb{R}$, $L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $U$ is a set-valued function and $E \subset \mathbb{R}^n \times \mathbb{R}^n$.

We will refer to $(P)$ as a *standard optimal control problem* and whenever the state constraint is absent we denote such problem as $(S)$.

Throughout this chapter we assume that the following basic assumptions are in force:

- $(t, (x, u)) \to (L(t, (x, u)), f(t, (x, u)))$ are $\mathcal{L} \times \mathcal{B}$-measurable,
- the multifunction $t \to U(t)$ has $\mathcal{L} \times \mathcal{B}$-measurable graph,
- the set $E$ is closed,
- $f$ is locally Lipschitz with respect to $x$ and
- $l$ is locally Lipschitz.

Again for $(P)$ (or $(S)$) a pair $(x, u)$ comprising an absolutely continuous function $x$, the state, and a measurable function $u$, the control, is called an *admissible process* if it satisfies all the constraints.

## 3.2 Assumptions

Throughout this report, the pair $(x^*, u^*)$ will always denote the solution of the optimal control problem under consideration. We shall also consider two important assumptions on the data of our problem $(P)$ that we now state.

Let us take any function $\phi$ defined in $[a, b] \times \mathbb{R}^n \times \mathbb{R}^m$ and taking values in $\mathbb{R}^n$ or $\mathbb{R}$.

**A1** There exist constants $k_x^\phi$ and $k_u^\phi$ such that for almost every $t \in [a, b]$ and every $(x_i, u_i)$ $(i = 1, 2)$ such that

$$x_i \in \{x : |x - x^*(t)| \leq \epsilon\}, \qquad u_i \in U(t)$$

we have

$$|\phi(t, x_1, u_1) - \phi(t, x_2, u_2)| \leq k_x^\phi |x_1 - x_2| + k_u^\phi |u_1 - u_2|.$$

**A2** The set valued function $t \to U(t)$ is closed valued and there exists a constant $c > 0$ such that for almost every $t \in [a, b]$ we have

$$|u(t)| \leq c \quad \forall u \in U(t).$$

When **A1** is imposed on $f$ and/or $L$, then the Lipschitz constants are denoted by $k_x^f$, $k_u^f$, $k_x^L$ and $k_u^L$. We will discuss these assumptions later on.

**A3** For all $x$ such that $|x(t) - x^*(t)| \leq \epsilon$ the function $t \to h(t, x)$ is continuous. Furthermore, there exists a constant $k_h > 0$ such that the function $x \to h(t, x)$ is Lipschitz of rank $k_h$ for all $t \in [a, b]$.

The need to impose continuity of $t \to h$ instead of merely upper semi-continuity is discussed in [23].

Recall that our basic assumptions are in force. Also suppose that both $f$ and $L$ satisfy **A1** and that **A2** holds. Those familiar with optimal control theory will surely see that assumption **A2** is quite strong since we are assuming the controls to be bounded. This requirement simplifies the proofs of the forthcoming results where limits of sequence of controls needed to be taken. On the other hand, it is quite common in applications.

For future use, observe that our assumptions also assert that following conditions are satisfied:

$$|\phi(t, x^*(t), u) - \phi(t, x^*(t), u^*(t))| \leq k_u^\phi |u - u^*(t)| \text{ for all } u \in U(t) \text{ a.e. } t \tag{2.1}$$

and there exists an integrable function $k$ such that

$$|\phi(t, x^*(t), u)| \leq k(t) \text{ for all } u \in U(t) \text{ a.e. } t. \tag{2.2}$$

In the above $\phi$ is to be replaced by $f$ and $L$. Moreover, it is a simple matter to see that the sets

$$f(t, x, U(t)) \text{ and } L(t, x, U(t)) \text{ are compact for all } x \in x^*(t) + \epsilon\mathbb{B}. \tag{2.3}$$

## 3.3 Auxiliary Results

In this section we now turn to problem $(S)$. We state and discuss an adaptation of Theorem 3.1 in [18] that will be essential to our analysis in the forthcoming sections. We call it "an adaptation" because we state it here under stronger assumptions than those in [18] to be used in our forthcoming analysis.

**Theorem 3.1** Let $(x^*, u^*)$ be a strong local minimum for problem $(S)$. If the basic assumptions are satisfied, $f$ and $L$ satisfy **A1** and $U$ is closed valued, then there exist $p \in W^{1,1}([a, b]; \mathbb{R}^n)$ and a scalar $\lambda_0 \geq 0$ satisfying the *nontriviality condition* [NT]:

$$||p||_\infty + \lambda_0 > 0,$$

the *Euler adjoint inclusion* [EI]:

$$(-\dot{p}(t), 0) \in \partial^C_{x,u}\Big(\langle p(t), f(t, x^*(t), u^*(t))\rangle - \lambda_0 L(t, x^*(t), u^*(t))\Big) - \{0\} \times N^C_{U(t)}(u^*(t)) \text{ a.e.},$$

the global *Weierstrass condition* [W]:

$\forall \;\; u \in U(t)$,

$$\langle p(t), f(t, x^*(t), u)\rangle - \lambda_0 L(t, x^*(t), u) \leq \langle p(t), f(t, x^*(t), u^*(t))\rangle - \lambda_0 L(t, x^*(t), u^*(t)) \text{ a.e.},$$

and the *transversality condition* [T]:

$$(p(a), -p(b)) \in N^L_E(x^*(a), x^*(b)) + \lambda_0 \partial^L l(x^*(a), x^*(b)).$$

Observe first that Theorem 3.1 in [18] assumes that $(x^*, u^*)$ is merely a *local minimum of radius $R$* and assumption **A2** is replaced by a weaker condition where the Lipschitz constants are integrable functions.

The conclusions of the main result in [22] differs from Theorem 3.1 because of the absence of [W]. Condition [EI] is a main feature of the necessary conditions established in [22]. For [EI] to hold it is important to assume that both $f$ and $L$ are Lipschitz with respect to $(x, u)$. The Lipschitz continuity with respect to the control instead of measurability as in the classical case may come as a disadvantage. However this is the feature responsible for the fact that Proposition 4.1 in [22] holds. This Proposition asserts that the necessary conditions are also sufficient for linear-convex problems (as defined in [22]) in the normal form (when $\lambda_0 = 1$). It is a simple matter to see that the proof of Proposition 4.1 in [22] adapts easily to show that Theorem 3.1 is also a sufficient condition.

In [22] the results are proved under the assumption that $(x^*, u^*)$ is a *weak local minimum*[1] instead of a strong local minimum. However, and as pointed out in [17], we only need to restrict the controls to $U(t) \cap \mathbb{B}_\epsilon(u^*(t))$ to see that conditions of Theorem 3.1 holds for such weak notion of minimum.

To finish this discussion we point out that the conditions given by the classical nonsmmoth maximum principle (see [15]) are [NT], [W], [T] and [EI] is replaced by

$$-\dot{p}(t) \in \partial_x^C \Big( \langle p(t), f(t, x^*(t), u^*(t)) \rangle - \lambda_0 L(t, x^*(t), u^*(t)) \Big). \tag{3.4}$$

We refer the reader to [17] for a discussion on (3.4) and [EI].

## 3.4 The Convex Case

We now consider an additional assumption on the "velocity set":

**C** The velocity set $\{(v, l) : v = f(t, x, u),\ l = L(t, x, u),\ u \in U(t)\}$ is convex for all $(t, x) \in [a, b] \times \mathbb{R}^n$.

We also introduce the following subdifferential

$$\bar{\partial}_x h(t, x) := \text{co}\, \{\lim \xi_i \ : \ \xi_i \in \nabla_x h(t_i, x_i),\ (t_i, x_i) \to (t, x)\}. \tag{4.5}$$

**Theorem 3.2** Let $(x^*, u^*)$ be a strong local minimum for problem $(P)$. Suppose that $f$ and $L$ satisfy **A1**, and that assumptions **A2** and **C** hold and $h$ satisfies **A3**, then there exist $p \in W^{1,1}([a, b]; \mathbb{R}^n)$, $\gamma \in L^1([a, b]; \mathbb{R})$, a measure $\mu \in C^\oplus([a, b]; \mathbb{R})$, and a scalar $\lambda_0 \geq 0$ satisfying

[1]The process $(x^*, u^*)$ is said to be a weak local minimum when it minimizes the cost over all admissible processes $(x, u)$ such that $|x(t) - x^*(t)| \leq \epsilon$ for all $t \in [a, b]$ and $|u(t) - u^*(t)| \leq \epsilon$ for almost all $t \in [a, b]$.

(i) $\mu\{[a,b]\} + ||p||_\infty + \lambda_0 > 0$,

(ii) $(-\dot{p}(t), 0) \in \partial^C_{x,u}\Big(\langle q(t), f(t, x^*(t), u^*(t))\rangle - \lambda_0 L(t, x^*(t), u^*(t))\Big)$
$-\{0\} \times N^C_{U(t)}(u^*(t))$ a.e.,

(iii) $\forall \;\; u \in U(t)$,
$\langle q(t), f(t, x^*(t), u)\rangle - \lambda_0 L(t, x^*(t), u) \leq \langle q(t), f(t, x^*(t), u^*(t))\rangle - \lambda_0 L(t, x^*(t), u^*(t))$ a.e.,

(iv) $(p(a), -q(b)) \in N^L_E(x^*(a), x^*(b)) + \lambda_0 \partial l(x^*(a), x^*(b))$,

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

(vi) $\text{supp}\{\mu\} \subset \{t \in [a,b] : h(t, x^*(t)) = 0\}$,

where

$$q(t) = \begin{cases} p(t) + \int_{[a,t)} \gamma(s)\mu(ds) & t \in [a,b) \\ p(t) + \int_{[a,b]} \gamma(s)\mu(ds) & t = b. \end{cases} \tag{4.6}$$

**Remark:** Before proceeding it is important to note that the proof of Theorem 3.1 shows that the normal cone $N^C_{U(t)}(u^*(t))$ appearing in [EI] can be rewritten simply as $K_U \partial^C d_{U(t)}(u^*(t))$ where $K_U$ is a fixed constant depending only on the constants of **A1** and **A2** (see remark on p. 4522 in [18]). After such replacement we get indeed a sharper version of [EI]. This sharper version of [EI] leads to a sharper version of (ii) in Theorem 3.2. The reader should keep in mind that such sharper versions are used in our analysis and they are important specially when taking limits. However, since [22], this kind of conditions have been stated with the normal cone and for consistence we keep such option in here.

The sketch of the proof of the Theorem 3.2 is presented in Section 3.7.

## 3.5 Non Convex Case

In this section we present a nonsmooth maximum principle for the optimal control problem $(P)$ when the convexity assumption imposed in the previous section is removed.

Now we replace the subdifferential $\bar{\partial}_x h$ by a more refined subdifferential $\partial^>_x h$ defined by

$$\partial^>_x h(t,x) := \text{co}\,\{\xi : \exists (t_i, x_i) \xrightarrow{h} (t,x) : h(t_i, x_i) > 0 \;\forall i,\; \nabla_x h(t_i, x_i) \to \xi\}. \tag{5.7}$$

**Theorem 3.3** Let $(x^*, u^*)$ be a strong local minimum for problem $(P)$. Assume that $f$ and $L$ satisfy **A1**, $h$ satisfies **A3** and that **A2** as well as the basic assumptions are satisfied. Then there exist an absolutely continuous function $p$, integrable functions $\xi$ and $\gamma$, a non-negative measure $\mu \in C^{\oplus}([a, b]; \mathbb{R})$, and a scalar $\lambda_0 \geq 0$ such that conditions (i)–(vi) of Theorem 3.2 hold with $\partial_x^{>} h$ as in (5.7) replacing $\bar{\partial}_x h$ and where $q$ is as defined in (4.6).

The above theorem adapts easily when we assume $(x^*, u^*)$ to be a weak local minimum instead of a strong local minimum. It is sufficient to replace $U(t)$ by $U(t) \cap \mathbb{B}_\epsilon(u^*(t))$. Our results in Theorem 3.3 can be extended to cover with a $W^{1,1}$ local minimum for problem $(P)$. This can be easily accomplished following the approach in [56].

**Theorem 3.4** Let $(x^*, u^*)$ be merely a $W^{1,1}$ local minimum for problem $(P)$. Then the conclusions of Theorem 3.3 hold.

## 3.6 Discussion of the Results

The main aim of Maximum Principle (MP) is to get the 'smallest' set of candidates of local minimizers which ensures the optimal solution of the problem. The *Classical Maximum Principle* (or Pontryagin Maximum Principle) is a necessary condition of optimality for optimal control problems. But for *normal linear-convex* problems, Pontryagin maximum principle (PMP) is a necessary and sufficient condition for optimality.

In the case of *Classical Nonsmooth Maximum Principle* it is not guaranteed that the optimality conditions are necessary and sufficient because of the nonsmoothness of the problems. An example in this regard can be consulted in [22]. In order to fix this situation, de Pinho and Vinter came up with the necessary conditions of optimality in the "Euler form". The main ingredient of these necessary conditions is the "Joint" adjoint inclusion. These necessary conditions are a sufficient condition for normal linear convex problem (see Proposition 4.1 in [22]). However, these necessary conditions failed to be a maximum principle since the *Weierstrass Condition* was not validated. In 2002, de Pinho *et al.* [23] extended the work of de Pinho and Vinter to state constraints and they also show that such generalization remains a sufficient condition for the normal linear-convex problems. Recently Clarke and de Pinho derived a new nonsmooth maximum principle (see [18]) in the vein of [22]. As before this result remains a sufficient condition for normal linear-convex problems.

As in [23], in this thesis we generalize the result of Clarke and de Pinho [18] to state constrained problems. As before, our Nonsmooth Maximum Principle (NMP) is a sufficient condition for normal linear-convex problems. Of course, the results in our state constrained problems are worth in so far as they apply to nonsmooth problems. In fact they coincide with known results for the smooth case.

## 3.7 Sketch of the Proofs

All the results are proved assuming that $L \equiv 0$. The case of $L \neq 0$ is treated by a standard and well known technique.

### 3.7.1 Convex Case

We now present the sketch of the Proof of Theorem 3.2.

- First we validate that the Theorem can be established for the simpler problem

$$
(\mathrm{Q}) \quad \begin{cases} \text{Minimize } l(x(b)) \\ \text{subject to} \\ \qquad \dot{x}(t) = f(t, x(t), u(t)) & \text{a.e. } t \in [a,b] \\ \qquad u(t) \in U(t) & \text{a.e. } t \in [a,b] \\ \qquad h(t, x(t)) \leq 0 & \text{for all } t \in [a,b] \\ \qquad (x(a), x(b)) \in \{x_a\} \times E_b. \end{cases}
$$

  Problem $(Q)$ is a special case of $(P)$ in which $E = \{x_a\} \times E_b$ and $l(x_a, x_b) = l(x_b)$.

  Now the proof consists of the following steps

  Q1 Define a sequence of problems penalizing the state-constraint violation. The sequence of problems is

$$
(Q_i) \quad \begin{cases} \text{Minimize } l(x(b)) + i \displaystyle\int_a^b h^+(t, x(t))\, dt \\ \text{subject to} \\ \quad \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [a,b] \\ \quad (x(a), x(b)) \in \{x_a\} \times E_b, \end{cases}
$$

  where $h^+(t,x) := \max\{0, h(t,x)\}$.

  Q2 Assume that

  **[IH]** $\lim_{i \to \infty} \inf\{Q_i\} = \inf\{Q\}$.

Q3 Set $W$ to be the set of measurable functions $u : [a, b] \to \mathbb{R}^m$, $u(t) \in U(t)$ a.e. such that a solution of the differential equation $\dot{x}(t) = f(t, x(t), u(t))$, for almost every $t \in [a, b]$, with $x(t) \in x^*(t) + \varepsilon\mathbb{B}$ for all $t \in [a, b]$ and $x(a) = x_a$ and $x(b) \in E_b$. We provide $W$ with the $L^1$ metric defined by $\Delta(u, v) :=\| u - v \|_{L_1}$ and set

$$J_i(u) := l(x(b)) + i \int_a^b h^+(t, x(t))\, dt.$$

Then $(W, \Delta)$ is a complete metric space in which the functional $J_i \colon W \to \mathbb{R}$ is continuous.

Q4 Apply Ekeland's theorem to the problems of the form

$$(O_i) \begin{cases} \text{Minimize} & J_i(u) \\ \text{subject to} & u \in W \end{cases}$$

which are closely related to $(Q_i)$. Observe the fact that $u^*$ (corresponding to $x^*$) is an admissible solution for each of these problems: we have $J_i(u^*) = l(x^*(b)) = \inf Q$.

Let $\varepsilon_i = J_i(u^*) - \inf Q_i$. The conclusions of application of Ekeland's theorem drives us to the sequence of optimal control problems showing that the process $(x_i, u_i)$ solves the following optimal control problem:

$$(E_i) \quad \begin{cases} \text{Minimize } l(x(b)) + i \displaystyle\int_a^b h^+(t, x(t))\, dt + \sqrt{\varepsilon_i} \int_a^b |u(t) - u_i(t)|\ dt \\ \text{subject to} \\ \quad \dot{x}(t) = f(t, x(t), u(t)) \qquad \text{a.e. } t \in [a, b] \\ \quad u(t) \in U(t) \qquad \text{a.e. } t \in [a, b] \\ \quad x(a) = x_a \\ \quad x(b) \in E_b. \end{cases}$$

The fact is that $\varepsilon_i \to 0$ allows us to prove that $u_i$ converges strongly to $u^*$ and $x_i$ converges uniformly to $x^*$.

Q5 We apply Theorem 3.1 to each $(E_i)$ obtaining the existence of an absolutely continuous function $p_i$ and a scalar $\lambda_i \geq 0$ such that

$$(p_i(t), \lambda_i) \neq 0 \text{ for all } t, \tag{7.8}$$

$$\begin{aligned} (-\dot{p}_i(t), n_i(t)) \in\ & \partial^C_{x,u} \{\langle p_i(t), f(t, x_i(t), u_i(t))\rangle - i\lambda_i h^+(t, x_i(t)) \\ & -\sqrt{\varepsilon_i}\lambda_i |u(t) - u_i(t)|\} \text{ a.e.} \end{aligned} \tag{7.9}$$

$$n_i(t) \in N^C_{U(t)}(u_i(t)) \quad \text{for all } t, \tag{7.10}$$

$$\begin{gathered} u \in U(t) \Longrightarrow \\ \langle p_i(t), f(t, x_i(t), u)\rangle - \sqrt{\varepsilon_i}\lambda_i \, |u - u_i(t)| \leq \langle p_i(t), f(t, x_i(t), u_i(t))\rangle \text{ a.e.} \end{gathered} \tag{7.11}$$

$$- p_i(b) \in N^L_{E_b}(x_i(b)) + \lambda_i \partial^L l(x_i(b)) \tag{7.12}$$

Q6 Rewriting these conditions and taking limits as in [23] we get the required conclusions, i.e., we get the existence of an absolutely continuous function $p : [a, b] \to \mathbb{R}^n$, an integrable function $\gamma : [a, b] \to \mathbb{R}^n$, a measure $\mu \in C^{\oplus}([a, b]; \mathbb{R})$ and a scalar $\lambda_0 \geq 0$ such that

$$\mu\{[a, b]\} + ||p||_\infty + \lambda_0 > 0, \tag{7.13}$$

$$(-\dot{p}(t), n(t)) \in \partial^C_{x,u}\langle q(t), f(t, x^*(t), u^*(t))\rangle - \{0\} \times N^C_{U(t)}(u^*(t)) \;\text{ a.e.} \tag{7.14}$$

$$\forall \;\; u \in U(t), \;\; \langle q(t), f(t, x^*(t), u)\rangle \leq \langle q(t), f(t, x^*(t), u^*(t))\rangle \;\text{ a.e.} \tag{7.15}$$

$$-q(b) \in N^L_{E_b}(x^*(b)) + \lambda_0 \partial l(x^*(b)), \tag{7.16}$$

$$\gamma(t) \in \bar{\partial} h(t, x^*(t)) \quad \mu\text{-a.e.}, \tag{7.17}$$

$$\text{supp}\{\mu\} \subset \{t \in [a, b] : h(t, x^*(t)) = 0\}, \tag{7.18}$$

where $q$ is as in (4.6).

Q7 Finally we show that **C** implies IH. We omit all the details here because this can be shown by an adaptation from the argument in [23].

The remaining of the proof has three stages. We first extend Theorem 3.2 to problems where $x(a) \in E_a$, and $E_a$ is a closed set. This is done following the lines in the end of the proof of Theorem 3.1 in [55].

Next we consider the case when the cost is $l = l(x(a), x(b))$. This is done applying the technique used in Step 2 of section 6 in [25]. And finally, following the approach in section 6 in [25], we derive necessary conditions when $(x(a), x(b)) \in E$, completing the proof.

### 3.7.2 Non Convex Case

We now proceed to present the sketch of the proof of Theorem 3.3. Our proof consists of several steps.

*Step 1:* We first consider the following 'minimax' optimal control problem where the state constraint functional $\max_{t\in[a,b]} h(t,x(t))$ appears in the cost.

$$(\tilde{R}) \begin{cases} \text{Minimize } \tilde{l}(x(a), x(b), \max_{t\in[a,b]} h(t,x(t))) \\ \text{over } x \in W^{1,1} \text{ and measurable functions } u \text{ satisfying} \\ \ \dot{x}(t) = f(t,x(t),u(t)) \text{ a.e. } t \in [a,b] \\ \ u(t) \in U(t) \text{ a.e. } t \in [a,b] \\ \ (x(a), x(b)) \in E_a \times \mathbb{R}^n. \end{cases}$$

where $\tilde{l} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is a given function and $E_a \subset \mathbb{R}^n$ is a given closed set. We observe that $(\tilde{R})$ is the optimal control problem with free endpoint constraints.

We impose here the following additional assumption **A4**, the necessity of which for the forthcoming development of our proof will come out soon.

**A4** The integrable function $\tilde{l}$ is Lipschitz continuous on a neighbourhood of

$$(x^*(a), x^*(b), \max_{t\in[a,b]} h(t,x^*(t)))$$

and $\tilde{l}$ is monotone in the $z$ variable, in the sense that $z' \geq z$ implies $\tilde{l}(y,x,z') \geq \tilde{l}(y,x,z)$, for all $(y,x) \in \mathbb{R}^n \times \mathbb{R}^n$.

*Step 2:* In this step, we present an *Auxiliary Result* in the form of the Proposition 3.5. The advantage of considering this Proposition is that this problem is a special case where the maximum principles presented in the Theorem 3.2 of convex case can be applied.

**Proposition 3.5** Let $(x^*, u^*)$ be a strong local minimum for problem $(\tilde{R})$. Assume the basic hypotheses, **A1**, **A2** and **A3** and the data for the problem $(\tilde{R})$ satisfies the hypothesis **A4**. Then there exist an absolutely continuous function $p : [a,b] \to \mathbb{R}^n$, integrable functions $\xi : [a,b] \to \mathbb{R}^m$ and $\gamma : [a,b] \to \mathbb{R}^n$, a non-negative measure $\mu \in C^{\oplus}([a,b];\mathbb{R})$, and a scalar $\lambda_0 \geq 0$

such that

$$\mu\{[a,b]\} + ||p||_\infty + \lambda_0 > 0, \tag{7.19}$$

$$(-\dot p(t), \xi(t)) \in \partial^C_{x,u}\langle q(t), f(t, x^*(t), u^*(t))\rangle \quad \text{a.e.} \tag{7.20}$$

$$\xi(t) \in N^C_{U(t)}(u^*(t)) \quad \text{a.e.} \tag{7.21}$$

$$(p(a), -q(b), \int_{[a,b)} \mu(ds)) \in N^L_{E_a}(x^*(a)) \times \{0,0\} + \lambda_0 \partial \tilde l(x^*(a), x^*(b), \max_{t\in[a,b]} h(t, x^*(t)), \tag{7.22}$$

$$\gamma(t) \in \bar\partial h(t, x^*(t)) \quad \mu\text{-a.e.}, \tag{7.23}$$

$$\forall \;\; u \in U(t), \;\; \langle q(t), f(t, x^*(t), u)\rangle \le \langle q(t), f(t, x^*(t), u^*(t))\rangle \;\; \text{a.e. }, \tag{7.24}$$

$$\text{supp}\{\mu\} \subset \left\{t \in [a,b] : h(t, x^*(t)) = \max\nolimits_{s\in[a,b]} h(s, x^*(s))\right\}, \tag{7.25}$$

*where $q$ is defined as in (4.6).*

The proof of Proposition 3.5 is done in the light of the Proposition 9.5.4 of [56]. Thus we are omitting the details by refereing readers to [7].

*Step 3:* Let us now consider the set

$$\begin{aligned} V := \{&(x,u,e) : (x,u) \text{ satisfies } \dot x(t) = f(t, x(t), u(t)), \\ &u(t) \in U(t) \text{ a.e., } \; e \in \mathbb{R}^n, (x(a), e) \in E \text{ and } \|x - x^*\|_{L^\infty} \le \epsilon\} \end{aligned} \tag{7.26}$$

and let $d_V : V \times V \to \mathbb{R}$ be a function defined by

$$d_V((x,u,e),(x',u',e')) = |x(a) - x'(a)| + |e - e'| + \int_a^b |u(t) - u'(t)| dt \tag{7.27}$$

For all $i$, we choose $\epsilon_i \downarrow 0$ and define the function

$$\tilde l_i(x, y, x', y', z) := \max\{l(x,y) - l(x^*(a), x^*(b)) + \epsilon_i^2, z, |x' - y'|\}.$$

Then it is easy to verify that $d_V$ defines a metric on the set $V$ and $(V, d_V)$ is a complete metric space with the following properties:

(i) If $(x_i, u_i, e_i) \to (x, u, e)$ in the metric space $(V, d_V)$, then $\|x_i - x\|_{L^\infty} \to 0$,

(ii) The function $(x,u,e) \to \tilde l_i(x(a), e, x(b), e, \max_{t\in[a,b]} h(t, x(t)))$ is continuous on $(V, d_V)$.

We now consider the optimization problem

$$\text{Minimize } \{\tilde{l}_i(x(a), e, x(b), e, \max_{t\in[a,b]} h(t, x(t))) : (x, u, e) \in V\}$$

We observe that

$$\tilde{l}_i(x^*(a), x^*(b), x^*(b), x^*(b), \max_{t\in[a,b]} h(t, x^*(t))) = \epsilon_i^2.$$

Also the cost $\tilde{l}_i$ is non-negative valued, it follows that $(x^*, u^*, x^*(b))$ is an $\epsilon_i^2$-minimizer for the above minimization problem. After the application of Ekeland's theorem we deduce that $(x_i, y_i, w_i \equiv 0, u_i)$ is a strong local minimum for the optimal control problem

$$(\tilde{R}_i) \begin{cases} \text{Minimize } \tilde{l}_i(x(a), y(a), x(b), y(b), \max_{t\in[a,b]} h(t, x(t))) \\ \quad + \varepsilon_i[|x(a) - x_i(a)| + |y(a) - y_i(a)| + w(b)] \\ \text{over } x, y, w \in W^{1,1} \text{ and measurable functions } u \text{ satisfying} \\ \quad \dot{x}(t) = f(t, x(t), u(t)), \dot{y}(t) = 0, \dot{w}(t) = |u(t) - u_i(t)| \text{ a.e.}, \\ \quad u(t) \in U(t) \text{ a.e.}, \\ \quad (x(a), y(a), w(a)) \in E \times \{0\}. \end{cases}$$

The cost function of $(\tilde{R}_i)$ satisfies all the assumptions of the Proposition 3.5 and thus this is an example of optimal control problem where the special case of maximum principle of Proposition 3.5 applies.

Thus, we deduce the existence of absolutely continuous functions $p_i \in W^{1,1}$, $d_i \in \mathbb{R}^n$, $r_i \in \mathbb{R}$, integrable functions $\xi_i$ and $\gamma_i$, a non-negative measure $\mu_i \in C^{\oplus}([a, b]; \mathbb{R})$ satisfying

$$\begin{gathered} (-\dot{p}_i(t), -\dot{d}_i(t), -\dot{r}_i(t), \dot{x}_i(t), \dot{y}_i(t), 0, \xi_i(t)) \in \\ \text{co } \partial(\langle q_i(t), f(t, x_i(t), u_i(t))\rangle + r_i|u(t) - u_i(t)|) \quad \text{a.e.} \\ \xi_i(t) \in \text{co } N_{U(t)}(u_i(t)) \quad \text{a.e.} \quad t \in [a, b], \\ \forall \;\; u_i \in U(t), \quad \langle q_i(t), f(t, x_i(t), u)\rangle + r_i|u(t) - u_i(t)|) \leq \langle q_i(t), f(t, x_i(t), u_i(t))\rangle \quad \text{a.e.} \;, \\ (p_i(a), d_i(a), r_i(a), -q_i(b), -d_i(b), -r_i(b), \textstyle\int_{[a,b]} \mu_i(dt)) \in \\ N_{E\times\{0\}}(x_i(a), y_i(a), w_i(a)) \times \{0, 0, 0, 0\} + \partial\{\tilde{l}_i(x, y, x', y', z) \\ +\varepsilon_i[|x(a) - x_i(a)| + |y(a) - y_i(a)| + w_i(b)]\}, \\ \gamma_i(t) \in \bar{\partial}_x h(t, x_i(t)) \quad \mu\text{-a.e.}, \\ \text{supp}\{\mu_i\} \subset \left\{t : h(t, x_i(t)) = \max\nolimits_{s\in[a,b]} h(s, x_i(s))\right\}. \end{gathered} \tag{7.28}$$

where $q_i := p_i + \int \gamma_i(s)\mu_i(ds)$ in the above relations.

Now rewriting these conditions and taking limits we will obtain the required conclusion.

# Chapter 4

# Maximum Principles for Mixed Constrained Problems–Convex Case

In this chapter we present a new nonsmmoth maximum principle for problem $(P')$ with both pure state and mixed state-control constraints under some convexity assumptions. Problems with mixed state control constraints, amply studied in a smooth framework (see, for example, [1], [3], [15], [26], [27],[33], [46], [49], [51]) have received little attention. Here we only focus on the convex case. The result we report here was announced in [6]. During this thesis we hope to treat the nonconvex case, but we have not fully studied this problem.

## 4.1 The Problem

The problem of interest is

$$(P') \quad \begin{cases} \text{Minimize } l(x(a), x(b)) \\ \text{subject to} \\ \quad \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [a, b] \\ \quad h(t, x(t)) \leq 0 \quad \text{for all } \; t \in [a, b] \\ \quad (x(t), u(t)) \in S(t) \quad \text{a.e. } t \in [a, b] \\ \quad (x(a), x(b)) \in E. \end{cases}$$

We consider the *basic hypotheses* on the problem data throughout. They are the following: $f$ and $L$ are $\mathcal{L} \times \mathcal{B}^{n+m}$, $S$ is $\mathcal{L} \times \mathcal{B}$, $E$ is closed and $l$ is locally Lipschitz.

## 4.2 Auxiliary Results

In this section we present a simplified version of one of the main results in [18] that will be of importance in the forthcoming developments.

Take a fixed interval $[a,b]$ and a set $S$ of $[a,b]\times\mathbb{R}^n\times\mathbb{R}^n$. Define

$$S(t) := \{(x,u) : (t,x,u) \in S\} \text{ for all } t \in [a,b]. \tag{2.1}$$

Assume for the time being that $E \subset \mathbb{R}^n \times \mathbb{R}^n$ and $l : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Consider the following problem:

$$(C)\ \begin{cases} \text{Minimize } l(x(a),x(b)) + \displaystyle\int_a^b L(t,x(t),u(t))dt \\ \text{subject to} \\ \qquad \dot{x}(t) = f(t,x(t),u(t)) \text{ a.e. } t \in [a,b] \\ \quad (x(t),u(t)) \in S(t) \text{ a.e. } t \in [a,b] \\ \quad (x(a),x(b)) \in E. \end{cases}$$

where $L : [a,b]\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}$.

For some $\varepsilon > 0$[1] define

$$S_*^\epsilon(t) = \{(x,u) \in S(t) :\ |x - x^*(t)| \le \epsilon\}.$$

In generic terms we assume that a function $\phi(t,x,u)$ satisfies $[L_*^\epsilon]$ if:

$[\mathbf{L}_*^\epsilon]$ There exist constants $k_x^\phi$ and $k_u^\phi$ such that for almost every $t \in [a,b]$ and every $(x_i,u_i) \in S_*^\epsilon(t)$ $(i = 1,2)$ we have

$$|\phi(t,x_1,u_1) - \phi(t,x_2,u_2)| \le k_x^\phi|x_1 - x_2| + k_u^\phi|u_1 - u_2|.$$

If this assumption is imposed on $f$, then the Lipschitz constants are denoted by $k_x^f$ and $k_u^f$. As for $S(t)$ we consider the following **bounded slope condition**:

$[\mathbf{BS}_*^\epsilon]$ There exists a constant $k_S$ such that for almost every $t \in [a,b]$ the following condition holds

$$(x,u) \in S_*^\epsilon(t),\ (\alpha,\beta) \in N_{S(t)}^P(x,u) \Longrightarrow |\alpha| \le k_S|\beta|.$$

The two previous hypotheses are strengthening of the analogous hypotheses in [18]. For the sake of uniformity and the forthcoming analysis we need to position an extra hypothesis on the set $S_*^\epsilon(t)$. We assume that:

[1] The $\varepsilon$ here can be taken to be equal to the parameter defining the strong local minimum.

$[\mathbf{CS}_*^\epsilon]$ The set $S_*^\epsilon(t)$ is closed and there exists an integrable function $c$ such that for almost every $t \in [a,b]$ the following holds

$$S_*^\epsilon(t) \text{ is closed and } (x,u) \in S_*^\epsilon(t) \implies |(x,u)| \leq c(t).$$

We observe that although $[CS_*^\epsilon]$ is a strong assumption it is nevertheless of importance in our future development. Necessary conditions of optimality for $(C)$ are given by the following theorem:

**Theorem 4.1** *(*adaption of Theorem 7.1 in [18]*) Let $(x^*, u^*)$ be a strong local minimum for problem $(C)$. Assume that the basic hypotheses, that $f$ and $L$ satisfy $[\mathbf{L}_*^\epsilon]$ and that $[BS_*^\epsilon]$ and $[CS_*^\epsilon]$ hold.*

*Then there exist an absolutely continuous function $p\colon [a,b] \to \mathbb{R}^n$ and a scalar $\lambda_0 \geq 0$ such that*

$$(p(t), \lambda_0) \neq 0 \quad \forall\, t \in [a,b], \tag{2.2}$$

$$(-\dot{p}(t), 0) \in \partial^C_{x,u} \left[ \langle p(t), \bar{f}(t) \rangle - \lambda_0 \bar{L}(t) - K|p(t)| \bar{d}_{S(t)}(t) \right] \ \ a.e. \tag{2.3}$$

$$\begin{aligned} (x^*(t), u) \in S(t) \Longrightarrow & \langle p(t), f(t, x^*(t), u) \rangle - \lambda_0 L(t, x^*(t), u) \\ & \leq \langle p(t), f(t, x^*(t), u^*(t)) - \lambda_0 L(t, x^*(t), u^*(t)) \rangle \ \ a.e. \end{aligned} \tag{2.4}$$

$$(p(a), -p(b)) \in N_E^L(x^*(a), x^*(b)) + \lambda_0 \partial l(x^*(a), x^*(b)), \tag{2.5}$$

*where $\bar{f}(t)$ and $\bar{L}(t)$ represent the function evaluated at $(t, x^*(t), u^*(t))$, $\bar{d}_{S(t)}(t)$ is the distance function to $S(t)$ evaluated at $(x^*(t), u^*(t))$ and $K$ in (2.3) is a constant depending only on $k_x^f$, $k_x^L$, $k_u^f$, $k_u^L$ and $k_S$.*

## 4.3 The Convex Case

We now turn to problem $(P')$. For this problem we derive a nonsmooth maximum principle under the following convexity assumption on the "velocity set":

**[C']** The velocity set $\{v \in \mathbb{R}^n : \ v = f(t,x,u), \ u \in S(t,x)\}$ is convex for all $t \in [a,b]$.

Furthermore we need to impose two more hypotheses on the data of our problem, one related to the state constraint and another to mixed constraints.

**[H1]** For all $x \in x^*(t) + \varepsilon\mathbb{B}$ the function $t \to h(t, x)$ is continuous 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 $t \in [a, b]$.

**[H2]** For almost every $t \in [a, b]$ the following condition holds: for all $u \in S(t, x^*(t))$ and all sequence $x_n \to x^*(t)$ there exists a sequence $u_n \in S(t, x_n)$ such that $u_n \to u$.

In the above the set $S(t, x)$ is defined as

$$S(t, x) = \{u : \ (x, u) \in S(t)\}$$

where $S(t)$ is as in (2.1). For a discussion on the need to impose continuity of $t \to h$ see [23]. Hypothesis [H2] asserts the lower semi-continuity of the multifunction $x \to S(t, x)$ (for definition and properties see [4]).

Assume the basic assumptions. Also suppose that $f$ satisfies $[L_*^\epsilon]$ and that both $[BS_*^\epsilon]$ and $[CS_*^\epsilon]$ hold. Under these assumptions we note for future use that the following conditions are satisfied:

$$|f(t, x^*(t), u) - f(t, x^*(t), u^*(t))| \leq k_u^f |u - u^*(t)| \text{ for all } u \in S(t, x^*(t)) \text{ a.e. } t \tag{3.6}$$

for all $u \in S(t, x^*(t))$ a.e. $t \in [a, b]$ and there exists an integrable function $k$ such that

$$|f(t, x^*(t), u)| \leq k(t) \text{ for all } u \in S(t, x^*(t)) \text{ a.e. } t. \tag{3.7}$$

We will now state our main result.

**Theorem 4.2** *Let $(x^*, u^*)$ be a strong local minimum for problem $(P')$. Assume that the basic hypotheses, [C'], [H1], [H2], $[BS_*^\epsilon]$ and $[CS_*^\epsilon]$ hold and that $f$ satisfies $[L_*^\epsilon]$. Then there exist an absolutely continuous function $p$: $[a, b] \to \mathbb{R}^n$, an integrable function $\gamma : [a, b] \to \mathbb{R}^n$, a measure $\mu \in C^\oplus([a, b]; \mathbb{R})$ and a scalar $\lambda_0 \geq 0$ such that*

(i) $$\mu\{[a, b]\} + ||p||_\infty + \lambda_0 > 0,$$

(ii) $$(-\dot{p}(t), 0) \in \partial_{x,u}^C \langle q(t), f(t, x^*(t), u^*(t))\rangle - N_{S(t)}^C(x^*(t), u^*(t)) \quad a.e.,$$

(iii) $$(x^*(t), u) \in S(t) \Longrightarrow \langle q(t), f(t, x^*(t), u)\rangle \leq \langle q(t), f(t, x^*(t), u^*(t))\rangle \quad a.e.,$$

(iv) $$(p(a), -q(b)) \in N_E^L(x^*(a), x^*(b)) + \lambda_0 \partial l(x^*(a), x^*(b)),$$

(v) $$\gamma(t) \in \bar{\partial} h(t, x^*(t)) \quad \mu\text{-}a.e.,$$

(vi) $$\text{supp}\{\mu\} \subset \{t \in [a, b] : h\,(t, x^*(t)) = 0\}\,,$$

*where $q(t)$ is as in (4.6) and the $\bar{\partial}$ is as in (4.5).*

**Remark 1** *The proof shows that we prove a sharper form of (ii):*

$$(-\dot{p}(t), 0) \in \partial^C_{x,u}\langle q(t), f(t, x^*(t), u^*(t))\rangle - K|q(t)|\partial^C_{x,u} d_{S(t)}(x^*(t), u^*(t)).$$

**Remark 2** *It is also easy to deduce from the proofs that when assumption [H2] is not imposed, a "weaker" version of the necessary conditions for $(P')$ (in the vein of [22]) can be obtained: all the conclusions but (iii) (the Weierstrass condition) hold.*

## 4.4 Sketch of the Proof of Theorem 4.2

The proof of Theorem 4.2 is done using the similar procedures (but in the context of mixed constraints) as in the Theorem 3.2 of Chapter 3. Since our main result is already appeared in [6], so we refer the reader to [6] for details about the sketch of the proof of Theorem 4.2.

# Chapter 5

# Application of Optimal Control to Real Life Problems

Now-a-days optimal control is a multidisciplinary field of research connecting almost all branches of science, engineering and technology, social and management sciences, life science and medicine. Although the early applications of optimal control were mainly in aerospace engineering (selection of optimal flight paths), the modern-day applications include process control, resource economics, robotics, engine management and even the recent applications lie in infectious diseases modeling and control in biology and medicine. In this chapter, our aim is to discuss another crucial part of my thesis "*Applications*" of optimal control to real life problems.

We discuss here the optimal control strategy for the chemotherapy of the infectious HIV model, causing one of the most devastating and life-killing pandemic diseases AIDS. We extend and modify the existing model proposed in [37], by adding the state constraints in the data and use the optimal control techniques to obtain the new solution.

## 5.1 Introduction

More than 30 years after the first detection of *Acquired Immune Deficiency Syndrome* (AIDS) and it's etiological agent *Human Immunodeficiency Virus* (HIV) in the early 1980s, the global health of the whole populations in the world is still under a great threat due to a mysterious and difficult unknown mechanism of HIV infections in the human body. Of the 34.3 million people worldwide living with HIV infections today, more than 24 million are in the developing world. A proper treatment or complete cure from AIDS is yet far away from the reality and even an *anti-HIV vaccine* is still a *dream* to the biologists and physicists [9, 34].

The chemotherapies are the only way of treatments for HIV positive patients till today which are aimed at killing or halting the virus pathogen and thus can serve to help the body fighting against the infections [40]. Several (more that 30 varieties of single and/or combined) antiretroviral drugs for the chemotherapy treatments have been approved by the US Food and Drug Administration (FDA) since 1987s aiming at reducing the viral population and improving the immune response. All these drugs cannot cure the diseases completely, rather they can improve the lives of HIV positive patients for a certain period depending on the optimal chemotherapeutic drug dosage strategies. This brings new hope to the treatment of HIV infection in absence of HIV vaccine, and we are exploring strategies for such treatments using optimal control techniques.

## 5.2 Mathematical Model

The mathematical model of HIV is a set of ordinary differential equations (ODEs) which describes the interactions between the $CD4^{+}T$ cells in the human immune systems and the viruses. Recently several mathematical models which describe a better understanding of the *cell-virus* interactions have been developed just in the past few years and even these are constantly being updated to improve the modeling aspects. We refer readers to ([19, 20, 37, 38, 40, 42, 50, 52] and references therein) for more details about the background and analysis of different models as well as the propagation of diseases.

### 5.2.1 The Existing Model

We mention that a mathematical model represents the dynamics of the *cell-virus interactions* and there exist a number of models representing this dynamics available in the literature. However, the crucial role is played by $CD4^{+}T$ lymphocytes, commonly referred to as *helper- T-cells*. These cells are mainly the target of the virus and the command system in the defense against it. A command from these cells activates $CD8^{+}T$ cells, shortly called *killer-T-cells*, which then fight the virus by killing its main source of reproduction. An infected $CD4^{+}T$ cell can produce in the neighborhood of 500 new viruses before its death and thus becomes the much more important target to destroy then the virus itself [42]. However, when a free HIV virus enters the body and attacks the *uninfected* $CD4^{+}T$ cells, the cells become infected and go through a *neutral stage* before becoming *actively infected* for a certain period. The cells in this *latent/interim stage*, which cannot infect other cells is called the *latently infected*. Thus the $CD4^{+}T$ cells are classified into three classes: *active/uninfected*

$CD4^+T$ cells, whose concentration is represented by a variable $T_A(t)$, and other two types of infected $CD4^+T$ cells are *latently infected* and *actively infected* cells with their concentrations represented by $T_L(t)$ and $T_I(t)$ respectively. The concentration of the free infectious virus is represented by $V(t)$. Now the dynamics of these four populations can be modeled by the following system of ordinary differential equations.

$$\frac{dT_A}{dt} = \frac{s}{1+V(t)} - \mu_{T_A}T_A(t) + rT_A(t)\Big(1 - \frac{T_A(t)+T_L(t)+T_I(t)}{T_{\max}}\Big) - \mu_i V(t)T_A(t), \tag{2.1}$$

$$\frac{dT_L}{dt} = \mu_i V(t)T_A(t) - \mu_{T_L}T_L(t) - \mu_c T_L(t), \tag{2.2}$$

$$\frac{dT_I}{dt} = \mu_c T_L(t) - \mu_{T_I}T_I(t), \tag{2.3}$$

$$\frac{dV}{dt} = N\mu_{T_I}T_I(t) - \mu_i V(t)T_A(t) - \mu_V V(t), \tag{2.4}$$

with the initial conditions

$$T_A(0) = T_{A0}, \quad T_L(0) = 0, \quad T_I(0) = 0, \quad \text{and} \quad V(0) = V_0, \tag{2.5}$$

for the case of infection by free virus, or

$$T_A(0) = T_{A0}, \quad T_L(0) = T_{L0}, \quad T_I(0) = T_{I0}, \quad \text{and} \quad V(0) = V_0, \tag{2.6}$$

for the case of infections by both free virus and infected cells.

In the equations (2.1)-(2.4) the terms with negative signs and multipliers $\mu_{T_A}$, $\mu_{T_L}$, $\mu_{T_I}$, and $\mu_V$ represent the natural deaths of active $CD4^+T$ cells, latently infected $CD4^+T$ cells, actively infected $CD4^+T$ cells, and of the free virus respectively. The details explanations and analysis of the model can be found in [37]. We provide the definitions of the parameters and their clinically approved values in Table 5.1.

### 5.2.2 The Optimal Chemotherapeutic Strategy

The influence of the chemotherapy in the HIV model is represented by a *control function* (or *chemotherapy function*) $u(t)$. This control represents the percentage of effect the chemotherapy has on the viral production. Since the drugs reduce the viral infectivity, We multiply the *chemotherapy function* $u(t)$ with the term $\mu_{T_I}T_I(t)$ in the last equation (2.4) modeling the dynamics of the

Table 5.1: Parameters and Constants used in HIV Model and Their Clinically Approved Values [37]

| Parameters and Constants | Definitions of the Parameters | Clinical values |
|---|---|---|
| $\mu_{T_A}$ | natural death rate of uninfected $CD4^{+}T$ cell | $0.02\ d^{-1}$ |
| $\mu_{T_L}$ | natural death rate of latently infected $CD4^{+}T$ cell | $0.02\ d^{-1}$ |
| $\mu_{T_I}$ | natural death rate of actively infected $CD4^{+}T$ cell | $0.24\ d^{-1}$ |
| $\mu_V$ | natural death rate of free virus population | $2.4\ d^{-1}$ |
| $\mu_i$ | rate $CD4^{+}T$ cells become infected by free virus | $2.4 \times 10^{-5}\ mm^3\ d^{-1}$ |
| $\mu_c$ | rate $T_L$ cells convert to actively infected | $3 \times 10^{-3}\ d^{-1}$ |
| $r$ | growth rate for the $CD4^{+}T$ cell population | $0.03\ d^{-1}$ |
| $N$ | number of free virus produced by $T_I$ cell population | 1200 |
| $T_{\max}$ | maximum $CD4^{+}T$ cell population level | $1.5 \times 10^3\ mm^{-3}$ |
| $s$ | source term for uninfected $CD4^{+}T$ cells where $s$ is the parameter in the term $\dfrac{s}{1+V}$ | $10\ d^{-1}\ mm^{-3}$ |

virus to achieve this effect mathematically. Now the term $N\mu_{T_I}T_I(t)$, representing the number of viruses produced by actively infected $CD4^{+}T$ cells at a rate $N$, is reduced by the factor $u(t)$ due to the killing affect of the drug. Thus the model presented in (2.1)-(2.4) is now reformulated as

$$\frac{dT_A}{dt} = \frac{s}{1+V(t)} - \mu_{T_A}T_A(t) + rT_A(t)\Big(1 - \frac{T_A(t)+T_L(t)+T_I(t)}{T_{\max}}\Big) - \mu_i V(t)T_A(t), \tag{2.7}$$

$$\frac{dT_L}{dt} = \mu_i V(t)T_A(t) - \mu_{T_L}T_L(t) - \mu_c T_L(t), \tag{2.8}$$

$$\frac{dT_I}{dt} = \mu_c T_L(t) - \mu_{T_I}T_I(t), \tag{2.9}$$

$$\frac{dV}{dt} = (1-u(t))N\mu_{T_I}T_I(t) - \mu_i V(t)T_A(t) - \mu_V V(t), \tag{2.10}$$

with the initial conditions as in (2.5) and/or in (2.6).

Here the control class are assumed to be the *measurable functions* defined on the fixed interval $[t_s, t_f]$, with the restriction that $u \in U(t) : 0 \le u(t) \le 1,\ \forall t \in [t_s, t_f]$.

Moreover, the finite interval of treatment is necessary because of the allowable treatment window for the chemotherapy. This is due to the fact that HIV has a strong ability to mutate and develop resistance to the chemotherapy treatment after some finite time frame [48]. On the other hand, the treatment for infinite time period has a hazardous side effects. Therefore treatment time $t$ should be finite and must satisfy the restriction $t \in [t_s, t_f]$. For most of the HIV chemotherapy drugs, the length of treatment is less than 500 days [30]. We also refer readers to [10, 19, 20, 35, 36, 39, 40] for more explanations and details discussions about the above issues.

The key target of this model is that the number of uninfected $CD4^{+}T$ cells should be kept as high as possible while at the same time the negative side-effects and cost of the chemotherapy are to be minimized. Taking all these into consideration, the objective function in the vein of [37] is:

$$\begin{aligned} &\text{Minimize} \quad J(u) = \int_{t_s}^{t_f} \Big( -T_A(t) + \frac{1}{2}Bu^2(t)\Big)dt \\ &\text{subject to} \\ &\text{the dynamics defined in} \quad (2.7) - (2.10), \\ &\text{with the initial conditions as in} \quad (2.6), \\ &\text{and the control constraints} \quad u \in [0,1] \text{ a.e.} \end{aligned} \tag{2.11}$$

where the parameter $B > 0$ represents the desired 'weight' on the benefit and cost. Equation (2.11) says that we are interested to maximize the benefit based on the $T$ cell counts, and minimize the systemic cost based on the percentage effect of the chemotherapy given (i.e. $u$) and thus we are minimizing the negative of the difference. It is worth to note that a quadratic cost is chosen here because of the fact that the cost function is assumed to be a nonlinear function of $u^*$ and there is not a linear relationship between the effects of treatment on $T$ cells or virus. If the control $u(t) = 1$ corresponds to maximal use of chemotherapy, then the maximal cost is represented as $(u(t))^2$. Now the aim is to find an *optimal control* $(u^*)$ such that

$$\min_{u \in [0,1]} J(u) = J(u^*).$$

## 5.3 Our Proposed Model

All the models in the references including the discussed one [37] of HIV infections have been solved applying the Pontryagin maximum principles of the optimal control theory in absence of state constraints. Our intention here is to find a new solution of the model in [37] imposing some state constraints in the data. We see that the $CD4^{+}T$ cells count is very crucial for the treatments of HIV infections. The $CD4^{+}T$ cells count '*less than* $200/mm^3$' indicates the severity of the disease. Our idea behind on imposing the state constraints is that the uninfected $CD4^{+}T$ cells count is restricted to follow a lower bound, say $T_A(t) \geq 200$ so that the $CD4^{+}T$ cells count never goes below $200/mm^3$.

Our proposed model then takes the following form:

$$\frac{dT_A}{dt} = \frac{s}{1+V(t)} - \mu_{T_A} T_A(t) + rT_A(t)\Big(1 - \frac{T_A(t)+T_L(t)+T_I(t)}{T_{\max}}\Big) - \mu_i V(t) T_A(t), \tag{3.12}$$

$$\frac{dT_L}{dt} = \mu_i V(t) T_A(t) - \mu_{T_L} T_L(t) - \mu_c T_L(t), \tag{3.13}$$

$$\frac{dT_I}{dt} = \mu_c T_L(t) - \mu_{T_I} T_I(t), \tag{3.14}$$

$$\frac{dV}{dt} = (1-u(t)) N \mu_{T_I} T_I(t) - \mu_i V(t) T_A(t) - \mu_V V(t), \tag{3.15}$$

$$T_A(t) \geq \tilde{T}, \tag{3.16}$$

with the initial conditions

$$T_A(0) = T_{A0}, \quad T_L(0) = T_{L0}, \quad T_I(0) = T_{I0}, \quad \text{and} \quad V(0) = V_0, \tag{3.17}$$

and the objective functional

$$\text{Minimize} \quad J(u) = -T_A(t_f) + \int_{t_s}^{t_f} Bu(t)dt \tag{3.18}$$

Here $\tilde{T}$ is a lower bound belonging to $\mathbb{R}$.

Our aim is to maximize the $CD4^{+}T$ cells count at the end of the treatment while minimize the effect of the drugs during the treatment. We consider the control $u(t)$ as a linear term in $L^1$ space for obtaining the stronger treatment strategy.

### 5.3.1 New Challenge

It is our challenge now to solve the proposed model (3.12)-(3.16) and find out a new chemotherapeutic strategy for the HIV infections. We will try to overcome this challenge during the PhD thesis. Although the analytical solution of such problem may be difficult due to nonlinearity of the dynamics and the presence of state constraint, we hope to handle this new model analytically and then compare the result by numerical illustration.

# Chapter 6

# Conclusion and Future Directions

In this concluding Chapter, we present a summary of the main *Contributions* of our work. We propose our *future plan* of the present work to be carried out within the rest of time frame. We also pose some open issues which require further extensive research and investigations in this area in future.

## 6.1 Contributions

Our main contributions have been presented in Chapter 3 and Chapter 4. We have derived *three distinct sets* of nonsmooth necessary conditions of optimality for constrained optimal control problems in the form of maximum principles with appropriate assumptions. In this report, we only present our results in a sketched form. The details derivations of our results will be accommodated in the final Thesis.

In Section 3.1 of Chapter 3 a new nonsmooth maximum principles for optimal control problems with state constraints have been derived when the velocity set is *convex*.

Another version of new necessary conditions of optimality for state constrained optimal control problems have been provided in Section 3.5 considering the more general *nonconvex* case of the data.

In Chapter 4 a new nonsmmoth maximum principle for problem with both pure state and mixed state-control constraints under some convexity assumptions have been derived. Having doing that we introduce a *bounded slope conditions* on the data set along with other appropriate hypotheses.

## 6.2 Future Works

Optimal Control is a vast area of distinct research in the Dynamic Optimization, but our work has touched a little part of it. Also in this PhD research, we have a *time constraint* to finish the work within a prescribed time framework, say till 2013. Within the rest of time, we will try to modify and refine our results we already have. We will mainly concentrate to implement the optimal control techniques in some real life problems, like the HIV model by imposing state constraints in the existing data set for obtaining the new solution analytically and comparing the result numerically. It will be of course a challenging work to get an analytical solution of such model (proposed in Section 5.3), because the analytical solution of the proposed model may be difficult due to the highly nonlinearity of the dynamic equations and presence of state constraint. Therefore, we are in parallel searching some problems which are linear and/or nonlinear but comparatively easy to handle analytically first and then illustrate the result numerically. But nevertheless, we hope to do that during the PhD thesis.

Research on Optimal Control with State Constraints is still an open issue. There are many questions related to this issue is still unsolved and/or not properly exploited, specially for the presence of measure (due to state constraints) in the problems. We believe that extensive and continuous involvement in optimal control research may result in to answer many questions and also may bring tremendous achievements for the generation plus. Therefore our further research will be dedicated to such open issues.

# Bibliography


[1] (MR780283) A. V. Arutyunov, "Optimality Conditions. Abnormal and Degenerate Problems," $1^{st}$ edition, Kluwer Academic Publishers, Dordrecht, 2000.

[2] A. V. Arutyunov and S. M. Aseev, *State Constraints in Optimal Control. The Degeneracy Phenomenon*, Systems & Control Letters, **26** (1995), 267–273.

[3] (MR2244084) A. V. Arutyunov and D. Yu. Karamzin, *Necessary conditions for a weak minimum in an optimal control problem with mixed constraints*, Differ. Equ. **41** (2005), 1532–1543.

[4] (MR1048347) J. P. Aubin and H. Frankowska, " Set-Valued Analysis," Birkhuser, Boston, 1990.

[5] D. P. Bertesekas, " Dynamic Programming and Optimal Control," Athena Scientific, 1995.

[6] M. H. A. Biswas and M. d. R. de Pinho. *A nonsmooth maximum principle for optimal control problems with state and mixed constraints–convex case.* AIMS Proceedings of the 8th AIMS International Conference on Dynamical Systems, Differential Equations and Applications, *Discrete and Continuous Dynamical Systems– Supplement 2011,* 174–183, 2011.

[7] M. H. A. Biswas, *On Necessary Conditions for Optimal Control Problems with State Constraints.* University of Porto, Faculty of Engineering, Technical Report, *to appear.*

[8] Oleg I. Bogoyavlenskij, *Necessary Conditions for Existence of Non-Degenerate Hamiltonian Structures* Commun. Math. Phys. **182**, (1996) 253–290.

[9] D. R. Burton, R. C. Desrosiers, R. W. Doms, W. C. Koff, P. D. Kwong, J. P. Moore, G. J. Nabel, J. Sodroski, I. A. Wilson, and R. T. Wyatt, *HIV Vaccine Design and the Neutralizing Antibody Problem*, Nature Immunology, **5** (3), (2004) 233–236.

[10] S. Butler, D. Kirschner and S. Lenhart, *Optimal Control of the Chemotherapy Affecting the Infectivity of HIV*, Mathematical Biology and Medicine, Vol. 6, World Scientific, 1995.

[11] L. Cesari, “ Optimization-Theory and applications, Problems with ordinary differential equations,” Applications of Mathematics, **17** , Springer-Verlag, New York, 1983.

[12] (MR709590) F. Clarke, “ Optimization and Nonsmooth Analysis,” John Wiley, New York, 1983.

[13] (MR1488695) F. H. Clarke, Yu. S. Ledyaev, R. J. Stern and P. R. Wolenski, “Nonsmooth Analysis and Control Theory,” Springer-Verla, New York, 1998.

[14] (MR2117692) F. Clarke, “Necessary conditions in dynamic optimization”, Mem. Amer. Math. Soc., 2005.

[15] (MR2208968) F. Clarke, *The maximum principle in optimal control, then and now*, Control Cybernet., **24** (2005) 709–722.

[16] F. Clarke, *Nonsmooth Analysis in Systems and Control Theory*, Encyclopedia of Complexity and System Science, Springer (2009), 6271–6285.

[17] F. Clarke and MdR de Pinho, *The nonsmooth maximum principle*, Control Cybernet., **38** (2009) 1151–1167.

[18] F. Clarke and MdR de Pinho, *Optimal control problems with mixed constraints*, SIAM J. Control Optim., **48** (2010) 4500–4524.

[19] D. J., Covert and D. Kirschner, *Revisiting Early Models of the Host-Pathogen Interactions in HIV Infection*, Comments on Theoretical Biology, **5** (6), (2000) 383–411.

[20] L. Cooper, *Theory of an Immune System Retrovirus*, Proc. National Academy of Science, **83**, (1986) 9159–9163.

[21] MdR de Pinho, *Mixed constrained control problems*, Journal of Mathematical Analysis and Applications., **278** (2003), 293–307.

[22] (MR1344031) MdR de Pinho and R. B. Vinter, *An Euler-Lagrange inclusion for optimal control problems*, IEEE Trans. Automat. Control, **40** (1995), 1191-1198.

[23] ( MR1874702) MdR de Pinho, M. M. A. Ferreira, and F. A. C. C. Fontes, *An Euler-Lagrange inclusion for optimal control problems with state constraints*, . Dynam. Control Systems **8** (2002), 23–45.

[24] (MR2167877) MdR de Pinho, M. M. A. Ferreira, and F. A. C. C. Fontes, *Unmaximized inclusion necessary conditions for nonconvex constrained optimal control problems*, ESAIM Control Optim. Calc. Var., **11** (2005) 614–632.

[25] (MR2529697) MdR de Pinho, P. Loewen and G. N. Silva, *A weak maximum principle for optimal control problems with nonsmooth mixed constraints*, Set-Valued and Variational Analysis, **17** 2009, 203–2219.

[26] (MR1323870) A. V. Dmitruk, *Maximum principle for the general optimal control problem with phase and regular mixed constraints*, Comput. Math. Model., **4** (1993) 364–377.

[27] A. Ya. Dubovitskii and A.A. Milyutin, "Extremum problems under constraints", Dokl. Akad. Nauk SSSR, **149** (1963), 1065–1089.

[28] A. S. Edelman and S. Zolla-Pazner, *AIDS: a syndrome of immune dysregulation, dysfunction, and deficiency*, FASEB J. **3** (1989) 22-30 .

[29] 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.

[30] K. R. Fister, S. Lenhart and J. S. Mc Nally, *Optimizing Chemotherapy in an HIV Model*, Electron. J. Diff. Eqns., **32**, (1998) 1–12.

[31] E. Giner, *Local Minimizers of Integral Functionals are Global Minimizers*, Proceedings of the American Mathematical Society, **123** (3), (1995), 755–757.

[32] 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**, (2), (1995) 181–281.

[33] (MR0203540) M. R. Hestenes, " Calculus of Variations and Optimal Control Theory," John Wiley, New York, 1966.

[34] IAVI Report, 30 Years of AIDS Vaccine Research, International AIDS Vaccine Initiative, **15** (3), May–June, 2011.

[35] H. R. Joshi, *Optimal control of an HIV immunology model*, Optim. Control Appl. Meth., **23**, (2002) 199-213.

[36] D. Kirschner, R. Mehr and A. Perelson, *Role of the Thymus in Pediatric HIV-1 infection*, Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology, **18** (1998) 95–109.

[37] D. Kirschner, S. Lenhart and S. Serbin, *Optimal Control of the Chemotherapy of HIV*, J. Math. Biol, **35**, (1997) 775–792.

[38] D. Kirschner, A. Perelson and R. Deboer, *The Dynamics of HIV infection of CD4+T Cells*, Mathematical Biosciences, **114**, (1993) 81–125.

[39] D. Kirschner and A. Perelson, Model for the Immune System Response to HIV: AZT Treatment Studies, Mathematical Population Dynamics: Analysis of Heterogeneity. Vol. One: Theory of Epidemics, ( Eds. O. Arino, D. Axelrod, M. Kimmel and M. Langlais) Wuerz Publishing Ltd., Winnepeg, Manitoba, Canada (1994), 295–310.

[40] D. Kirschner and G. F. Webb, *Immunotherapy of HIV-1 Infection*, Journal of Biological Systems, **6**(1), (1998) 71–83.

[41] H. W. Kuhn and A. W. Tucker, *Nonlinear Programming*, Proc. Second Berkeley Symp. on Math. Statist. and Prob.,(J. Neyman, Ed.). Univ. of Calif. Press, (1951), 481–492.

[42] U. Ledzewicz and H. Schattler, *On optimal controls for a general mathematical model for chemotherapy of HIV* Proceedings of the American Control Conference, **5**, (2002) 3454–3459.

[43] P. D. Loewen and R. T. Rockafellar, *New Necessary Conditions for the Generalized Problem of Bolza*, SIAM J. Control and Optimization-Preprint, **1** (1995) .

[44] S. O. Lopes, "Nondegenerate forms of the Maximum Principle for Optimal Control Problems with State Constraints," PhD Thesis, University of Minho, Portugal. 2008.

[45] Sofia O. Lopes and Fernando A.C.C. Fontes, *On Stronger Forms of First-Order Necessary Conditions of Optimality for State-Constrained Control Problems* International Journal of Pure and Applied Mathematics, **49**(4) (2008), 459–466.

[46] (MR1641590) A. A. Milyutin and N. P. Osmolovskii, "Calculus of Variations and Optimal Control," Translations of Mathematical Monographs 180, American Mathematical Society, Providence, Rhode Island, 1998.

[47] (MR2191744) B. Mordukhovich, "Variational analysis and generalized differentiation. Basic Theory", Springer-Verlag, Berlin, 2006.

[48] P. L. Nara, L. Smit, N. Dunlop, W. Hatch, M. Merges, D. Waters, J. Kelliher, R. C. Gallo, P. J. Fischinger and J. Goudsmit, *Emergence of Viruses Resistant to Neutralization by V3-Specific Antibodies in Experimental Human Immunodeficiency Virus Type 1 IIIB Infection of Chimpanzees*, J. Virology 64, 37793791 (1990)

[49] (MR0440440) L. W. Neustadt, "Optimization, A Theory of Necessary Conditions," Princeton University Press, New Jersey, 1976.

[50] M. A. Nowak and R. M. May, *Mathematical Biology of HIV Infections: Antigenic Variation and Diversity Threshold*, Mathematical Biosciences, **106** (1991) 1–21.

[51] (MR2048167) Z. Páles and V. Zeidan, *Optimal control problems with set-valued control and state constraints*, SIAM J. Optim., **14** (2003) 334-358.

[52] A. S. Perelson, D. E. Kirschner and R. D. Boer, *Dynamics of HIV Infection of* $CD4^+T$ *cells*, Mathematical Biosciences, **114** (1993) 81–125.

[53] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mischenko, "The Mathematical Theory of Optimal Multiprocesses," John Wiley, New York, 1962.

[54] F. Rampazzo and R. Vinter, *Degenerate Optimal Control Problems with State Constraints*, SIAM J. Control Optim., **39**(4), (2000) 989-1007.

[55] (MR672197) R.B. Vinter and G. Pappas, *A maximum principle for nonsmooth optimal-control problems with state constraints*, J. Math. Anal. Appl., **89** (1982) 212–232.

[56] (MR1756410) R.B. Vinter, "Optimal Control," Birkhäuser, Boston, 2000.

[57] Patrick D. Woodford, "Optimal Control for Nonsmooth Dynamic Systems," Ph D Thesis, Centre for Process Systems Engineering and Department of Electrical and Electronic Engineering, Imperial College of Science, Technology and Medicine, London SW7 2BY, 1997.