OPTIMAL CONTROL OF SINGULAR DIFFERENTIAL SYSTEMS

BY:

MOHAMMED MOGIB M ALSHAHRANI

December 2011

OPTIMAL CONTROL OF SINGULAR DIFFERENTIAL SYSTEMS

BY

MOHAMMED MOGIB M ALSHAHRANI

A Dissertation Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

In

MATHEMATICS

December 2011

KING FAHD UNIVERSITY OF PETROLEUM & MINIRALS
DHAHRAN, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This dissertation, written by MOHAMMED MOGIB M ALSHAHRANI under the direction of his thesis advisors and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY IN MATHEMATICS.

Dissertation Committee

14/12/2011

Prof. Boris Mordukhovich

Dissertation Committee Chairman

14/12/2011

Prof. Mohamed El-Gebeily

Co-Chairman

14/12/2011

Prof. Suliman Al-Homidan

Member

14/12/2011

Prof. Bilal Chanane

Member

Kassem Mustapha 14/12/2011

Dr. Kassem Mustapha

Member

Dr. Hattan Tawfiq

Department Chairman

Dr. Salam Zummo

Dean of Graduate Studies

عمادة الدراسات العليا
DEANSHIP OF GRADUATE STUDIES

24/12/11

Date

To Sabha, Mogib, Omar, Haya, and Maymounah.

ACKNOWLEDGEMENTS

First and above all, I praise Allah, the almighty for providing me this opportunity and granting me the capability to proceed successfully.

I offer my sincerest gratitude to my supervisor, Professor Boris Mordukhovich, for his advise, encouragement, motivation and prompt replies to my questions. I would like also to express my deep appreciation and gratitude to my co-advisor Professor Mohamed El-Gebeily for his encouragement and effort and without him this dissertation would not have been completed or written. One simply could not wish for a better or friendlier supervisor. I extend my sincere thanks to the members of my dissertation committee, Professors Suliman Al-Homidan, Bilal Chanane and Dr. Kassem Mustapha for their comments and advise.

Many thanks to my close friends Mohammed Nasser, Ali Monahi, Ali Saad, Muteb Alqahtani, and Khalid Al-Nowaiser. Thank you for your encouragement and support. Thank you for being my friends.

I cannot finish without thanking my family. My deep thanks and sincere gratitude to my father Mogib, my mother Haya and my siblings Ali, Nourah, Ibrahim, Ayshah, Ahmad, Fayzah, Saif, Fayez, and Faisal for their unconditional love and sincere prayers.

I am forever grateful and thankful to Allah, the almighty, for blessing me with four beautiful, patient, and understanding children. Mogib, Omar, Haya and Maymounah! You are my burning fuel.

And finally, I know that you did not want to be named, My lovely wife, Dear Sabha. You are the best thing that ever happened to me. Without you, simply, this was impossible. May Allah reward you Jannat Al-Firdous.

TABLE OF CONTENTS

TABLE OF FIGURES

ABSTRACT (ENGLISH)

NAME	:	MOHAMMED MOGIB MOHAMMED ALSHAHRANI
TITLE OF STUDY	:	OPTIMAL CONTROL OF SINGULAR DIFFERENTIAL SYSTEMS
MAJOR FIELD	:	MATHEMATICS
DATE OF DEGREE	:	DECEMBER, 2011

In this dissertation we formulate, for the first time in the literature, an optimal control problem for self-adjoint ordinary differential operator equations in Hilbert spaces and derive necessary conditions for optimal controls to this problem in an appropriate extended form of the Pontryagin Maximum Principle.

ملخص الرسالة

الاسم	:	محمد بن معجب بن محمد الشهراني
عنوان الرسالة	:	التحكم الأمثل للنظم التفاضلية الشاذة
التخصص	:	الرياضيات
تاريخ التخرج	:	ديسيمبر 2011

نقدم في هذه الرسالة، ولأول مرة، مسألة تحكم لمعادلات تابعية تفاضيلة عادية ذاتية التجاور في فضاءات هلبرت ونقوم باستنتاج الشروط الضرورية للتحكمات المثلى لهذه المسألة في صورة ممتدة مناسبة من مبدأ بونترياقن الأعظمي.

Chapter 1

INTRODUCTION

This thesis addresses the following controlled system governed by singular differential operator equations in Hilbert spaces:

$$Lx = f(x,u,t), \qquad u(t) \in U \;\text{ a.e. }\; t \in I = (a,b), \quad -\infty \le a < b \le \infty, \tag{1.1}$$

where L is a self-adjoint extension of the minimal operator L_0 (see Chapter 2) generated by a formally self-adjoint quasi-differential expression l and a positive weight function w satisfying the equation

$$\ell x = \lambda w x \qquad \text{on} \quad I \tag{1.2}$$

in the Hilbert space $\mathcal{H} = L^2(I,w)$ of real-valued square integrable functions with respect to the weight fuctiom w, where $u(\cdot)$ is a measurable control action taking values from the given control set U, and where the function f is real-valued.

Optimal control theory is a remarkable area of Applied Mathematics, which has been developed for various classes of controlled systems governed by ordinary differential, functional differential, and partial differential equations and inclusions; see, e.g., [23, 31] with the vast bibliographies therein. However, we are not familiar with any developments on optimal control of differential operator equations of type (1.1). .

The differential operator equation in (1.1) describes many systems in physics and engineering. Many problems of mathematics can be also categorized to be of this form. Sturm-Liouville differential equations, Schrödinger operators and some Dirac systems belong to the long list of problems that can be studied under the form of (1.1). Weidmann in [45] gives a list of solvable examples in which he studies different problems described in this form and calculates the resolvent, spectral representation, spectrum etc. of the operator L in each of these examples. In [13], a collection of more than 50 examples of Sturm-Liouville differential equations; many of these examples are connected with problems in mathematical physics and applied mathematics.

We denote the set of complex numbers and the set of real numbers by the two symbols $\mathbb{C}$ and $\mathbb{R}$ respectively. We sometimes write Ax for some operator or a function A and an element x in the domain of A to mean the image of x under A. In other words, we use Ax to mean $A(x)$ in the standard convention of notation. We rely on the context to read Ax as the image under A rather than the product of A and x.

This thesis is organized as follows. Chapter 2 gives a comprehensive account of the operator equation in (1.1) with concentration on the basic definitions and results that help the reader to have a clear understanding of the problem we are studying. Our display is no where else to be found in this arrangement. We believe that this chapter when extended can be a solid launching pad and a convenient way to whomever interested in pursuing further studies in the theory of singular differential equations.

In Chapter 3, we introduce the problem of optimal control in general considerations and the necessary optimality conditions of optimizers of these problems. We define and discuss the optimal control problem and we describe necessary optimality conditions. First with no constraints on the control set and then with constrained control set deriving necessary optimality conditions in the form of the Pontryagin maximum principle (PMP). We also give a historical review on the development of optimal control theory. This work [2] is a small contribution to the field of optimal control.

In Chapter 4, we formulate, for the first time in the literature, an optimal control problem for self-adjoint ordinary differential operator equations in Hilbert spaces and derive necessary conditions for optimal controls to this problem in an appropriate extended form of the Pontryagin Maximum Principle.

Chapter 5 summarizes the work accomplished in this thesis and presents some interesting problems for further investigation. We think that some of the problems presented in this chapter can be studied in a Master thesis or even in a PhD dissertation.

Chapter 2

SINGULAR ORDINARY DIFFERENTIAL OPERATORS

The goal of this chapter is to shed some light on the equation (1.1) and its well-developed background that is necessary to understand the new results introduced in this thesis.

Very general quasi-differential forms, and in particular symmetric ones, have been considered by Shin [39]. They were rediscovered by Zettl [46] in a slightly different but equivalent form. Special cases of these very general symmetric quasi-differential expressions have been used extensively by many authors, see [4, 22, 36, 28, 11, 12, 44].

The development of the theory of symmetric differential operators in the books by Naimark [32, 33] and by Akhiezer and Glazman [1] is based on the real symmetric form analogous to (2.4). Although these authors refer to Shin's more general symmetric expressions they make no use of them. In [46] it was shown that the techniques in these books can be applied to a much larger class of symmetric operators generated by these very general differential expressions.

In Section 2.1, we present the basic definitions of the general symmetric quasi-differential expressions and give some properties and examples. In Section 2.2 we discuss the deficiency spaces associated to a symmetric operator. The basic theory of the minimal and maximal operators are presented in Section 2.3. The Glazman-Krein-Naimark (GKN) Theorem that describes the domains of self-adjoint extensions of the minimal operator is presented in Section 2.4.

2.1 Quasi-Differential Expressions

In this section we summarize some basic facts about general quasi-differential expressions of even and odd order and real or complex coefficients for the convenience of the reader. For a more comprehensive discussion of quasi-differential equations, the reader is referred to [46] and to [18] in the scalar coefficient case and to [29] for the general case with matrix coefficients.

We define the general quasi-differential expression following the development in [46, 18]. To do so we let $I = (a, b)$ be an interval with $\infty \leq a < b \leq \infty$, n be an integer greater than 1 and let

$$\begin{aligned} Z_n(I) \quad := \quad & \{Q = (q_{rs})_{r,s=1}^{n}\,, \\ & q_{r,s} = 0, \quad \text{a.e.} \quad \text{on} \quad I, \quad \text{for} \quad 2 \leq r+1 < s \leq n, \\ & q_{r,r+1} \neq 0, \quad \text{a.e.} \quad \text{on} \quad I, \quad q_{r,r+1}^{-1} \in L_{loc}(I) \quad \text{for} \quad 1 \leq r \leq n-1, \\ & q_{r,s} \in L_{loc}(I), \quad \text{for} \quad 1 \leq r, s \leq n\}, \end{aligned} \tag{2.1}$$

where $L_{loc}(I)$ denotes the space of all complex-valued functions that are locally (i.e on each compact subinterval) integrable on I. These matrices, i.e. $Z_n(I)$, are called, Shin-Zettl matrices.

A typical member of this class Q displays the format

$$Q = \begin{bmatrix} * & q_{12} & 0 & 0 & \cdots & 0 \\ * & * & q_{23} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ * & * & \cdots & * & q_{n-2,n-1} & 0 \\ * & * & \cdots & \cdots & * & q_{n-1,n} \\ * & * & * & \cdots & * & * \end{bmatrix}$$

where $*$ stands for any locally integrable function that is not on or above the super diagonal of Q.

Definition 2.1 (quasi-derivatives)

For a fixed choice of $Q \in Z_n(I)$, let

$$V_0 := \{x : I \to \mathbb{C}, \quad x \text{ is measurable}\}.$$

The quasi-derivatives $x^{[k]}$ for $k = 0, \cdots, n$, are defined inductively as

$$\begin{aligned} x^{[0]} &:= x, \quad x \in V_0, \\ x^{[k]} &:= q_{k,k+1}^{-1} \left\{ \left(x^{[k-1]}\right)' - \sum_{s=1}^{k} q_{ks} x^{[s-1]} \right\}, \quad x \in V_k \quad \text{for} \quad k = 1, \cdots, n \end{aligned}$$

where $q_{n,n+1} := 1$ and

$$V_k := \left\{ x \in V_{k-1} : x^{[k-1]} \in AC_{loc}(I) \right\}, \quad \text{for } k = 1, \cdots, n.$$

Here the prime marks the ordinary derivative and $AC_{loc}(I)$ is the set of all complex-valued locally absolutely continuous functions on I, i.e., absolutely continuous on each compact subinterval $[\alpha, \beta]$ of I which means that for any $\epsilon > 0$ there is δ such that

$$\sum_j ||x(t_{j+1}) - x(t_j)|| \leq \epsilon \quad \text{whenever} \quad \sum_j |t_{j+1} - t_j| \leq \delta$$

for the disjoint intervals $(t_j, t_{j+1}] \subset [\alpha, \beta]$. □

The quasi-derivatives $x^{[k]}$ for $k = 0, 1, \cdots, n$, are defined as certain linear combinations of the ordinary derivatives $x^{(k)}$, in terms of a prescribed complex $n \times n$ matrix $Q = Q(t)$ for $t \in I$, of Shin-Zettl type, see [19, 14, 19, 33, 46].

Definition 2.2 (quasi-differential expression)

The quasi-differential expression l_Q associated with Q is defined by

$$l_Q x := i^n x^{[n]}, \qquad (i^2 = -1),$$

on the domain $D(Q) := V_n$. □

Clearly, l_Q is a linear map of $D(Q)$ into $L_{loc}(I)$ and different matrices Q may generate the same linear map. The definition generalizes classical differential expressions of order n on I defined as

$$Mx = p_n x^{(n)} + p_{n-1} x^{(n-1)} + \cdots + p_1 x' + p_0 x \tag{2.2}$$

with complex coefficients $p_k \in L_{loc}(I), k = 0, 1, \cdots, n-1$, and further $p_n \in AC_{loc}(I)$ with $p_n \neq 0$ on I. The corresponding domain for M is

$$D(M) :\equiv \{x := I \to \mathbb{C} | x^{(k)} \in AC_{loc}(J) \text{ for } k = 0, 1, \cdots, n-1\},$$

in terms of the ordinary derivatives $x^{(k)}$, so $x^{(n)}$ and also $Mx \in L_{loc}(J)$. To see this, define the following $n \times n$ matrix $Q \in Z_n(I)$ by

$$Q = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & i^n p_n^{-1} \\ -i^n p_0 & -i^n p_1 & -i^n p_2 & \cdots & -i^n p_{n-2} & (p_n' - p_{n-1})p_n^{-1} \end{bmatrix}.$$

That is, $Q = (q_{rs})_{r,s=1}^{n}$ with

$$\begin{aligned} q_{rs} &= 0, && \text{for } (s \neq r+1 \text{ and } 1 \leq r, s \leq n-2) \text{ or } (r = n-1 \text{ and } 1 \leq s \leq n-1) \\ q_{r,r+1} &= 1, && \text{for } 1 \leq r \leq n-2, \\ q_{n-1,n} &= i^n p_n^{-1}, && \\ q_{n,s} &= -i^{-n} p_s, && \text{for } 1 \leq s \leq n-1, \\ q_{n,n} &= (p_n' - p_{n-1})p_n^{-1}. && \end{aligned}$$

The matrix Q belongs to $Z_n(I)$ since p_n^{-1} is locally integrable on I and $p_n^{-1} \neq 0$ on I. In fact, $p_n^{-1} \in AC_{loc}(I)$; because $p_n \in AC_{loc}(I)$ and $p_n \neq 0$ on I. Indeed, let's compute the quasi-derivatives $x^{[k]}$ as follows

$$x^{[k]} = x^{(k)}, \qquad \text{for} \quad k = 0, 1, \cdots, n-2$$

and then

$$x^{(n-1)} = \left(x^{[n-2]}\right)' = i^n p_n^{-1} x^{[n-1]}$$

which shows that $x^{[n-1]} \in AC_{loc}(I)$ and therefore

$$D(M) = D(Q).$$

Moreover,

$$\begin{aligned}
x^{[n]} &= \left(x^{[n-1]}\right)' + i^{-n}\left\{p_0x + p_1x^1 + \cdots + p_{n-2}x^{(n-2)}\right\} - (p_n' - p_{n-1})p_n^{-1}i^{-n}p_nx^{(n-1)} \\
&= \left(i^{-n}p_nx^{(n-1)}\right)' + i^{-n}\left\{p_0x + p_1x^1 + \cdots + p_{n-2}x^{(n-2)}\right\} - i^{-n}(p_n' - p_{n-1})x^{(n-1)} \\
&= i^{-n}\left\{\left(p_nx^{(n-1)}\right)' + p_0x + p_1x^1 + \cdots + p_{n-2}x^{(n-2)} - (p_n' - p_{n-1})x^{(n-1)}\right\} \\
&= i^{-n}\left\{p_n'x^{(n-1)} + p_nx^{(n)} + p_0x + p_1x^1 + \cdots + p_{n-2}x^{(n-2)} - p_n'x^{(n-1)} + p_{n-1}x^{(n-1)}\right\} \\
&= i^{-n}\left\{p_nx^{(n)} + p_0x + p_1x^1 + \cdots + p_{n-2}x^{(n-2)} + p_{n-1}x^{(n-1)}\right\} \\
&= i^{-n}\left\{Mx\right\}.
\end{aligned}$$

Hence,

$$Mx = i^n x^{[n]}, \qquad \text{for } x \in D(M) = D(A).$$

On the other hand, it is not possible to simplify the quasi-expression, l_Q, or to describe its domain of definition without reference to all the quasi-derivatives. In [17], necessary and sufficient conditions for the quasi-differential expression l_Q to be equivalent to a classical expression M are given.

The quasi-differential expression,l_Q, enjoys many advantages over the classical differential expression 2.2. Among these advantages, see [46], are: They are more general. Smoothness conditions on the coefficients are not needed in deriving the Lagrange identity, Definition 2.5.

Definition 2.3

A differential expression $\mathcal{M}$ on I (either classical M or quasi l_Q as above) is formally self-adjoint or Lagrange symmetric if:

$$\int_I \{\mathcal{M}(x_1)\bar{x}_2 - x_1\overline{\mathcal{M}(x_2)}\}dx = 0$$

for all $x_1, x_2 \in \mathcal{D}_0(\mathcal{M})$, where

$$\mathcal{D}_0(\mathcal{M}) = \{x \in D(\mathcal{M}) |\ \text{supp}(x) \subset I\}.$$

In other words, $\mathcal{D}_0(\mathcal{M})$ is a subset of functions in $D(\mathcal{M})$ whose supports are compact subsets of the interior of I. □

Remark 2.4

If $\mathcal{M} = M$ is a classical differential expression, 2.2, with smooth coefficients,that is

$$p_k \in C^k \quad \text{for} \quad k = 0, 1, \cdots, n, \tag{2.3}$$

then M is formally self-adjoint if and only if M coincides with its Lagrange adjoint M^+:

$$M[x] = M^+[x] :\equiv (-1)^n(\bar{p}_n x)^{(n)} + (-1)^{n-1}(\bar{p}_{n-1}x)^{(n-1)} + \dots \bar{p}_0 x.$$

It is known, see [32] or [10, page 1290], that every formally self-adjoint differential expression M whose coefficients satisfy (2.3) can be expressed in the form

$$Mx = \sum_{k=0}^{[n/2]} (-1)^k \left(a_k x^{(k)}\right)^{(k)} + \sum_{k=0}^{[(n-1)/2]} i \left[\left(b_k x^{(k)}\right)^{(k+1)} + \left(b_k x^{(k+1)}\right)^{(k)}\right] \tag{2.4}$$

where a_k, b_k are real-valued function and $[x]$ denotes the largest integer less than or equal to x. In particular every formally self-adjoint differential expression M with real coefficients satisfying (2.3) is of even order $n = 2m$ and has the form

$$\sum_{k=0}^{m} (-1)^k \left(a_k x^{(k)}\right)^{(k)} \tag{2.5}$$

with a_k real-valued function. For $m = 1$, (2.5) reduces to the Sturm-Liouville operator

$$-(a_1 x')' + a_0 x.$$

On the other hand it can readily be shown, by "removing the parenthesis," that every expression of the form (2.5) with $a_k \in C^\infty$ is a formally self-adjoint expression. It is sufficient to verify $M = M^+$ for all $x \in C^\infty(I)$. However, for general M (with non-smooth coefficients), it is possible to test for Lagrange symmetry only by replacing M by an equivalent quasi-differential expression M_Q, see [16, Appendix A], and then test for Lagrange symmetry for M_Q, as we shall see below. □

Definition 2.5 (Lagrange Identity)

Let $x_1, x_2 \in D(Q)$ for some given quasi-differential expression l_Q. Then we have the following identity, called Lagrange Identity,

$$l_Q(x_1)\overline{x_2} - x_1\overline{l_Q(x_2)} = \frac{d}{dt}[x_1, x_2], \tag{2.6}$$

where

$$[x_1, x_2](t) := i^n \sum_{k=0}^{n-1} (-1)^k x_1^{[n-1-k]}(t)\overline{x_2^{[k]}(t)}, \qquad \text{for } t \in I.$$

□

It should be noted that

$$[x_1, x_2](a) = \lim_{t \to a^+} [x_1, x_2](t) \quad \text{and} \quad [x_1, x_2](b) = \lim_{t \to b^-} [x_1, x_2](t).$$

If we integrate both sides of (2.6) over a finite interval $[\alpha, \beta] \subset I$, we have the Lagrange identity in integral form, also called the Lagrange-Green identity,

$$\int_\alpha^\beta l_Q(x_1)\overline{x_2}dt - \int_\alpha^\beta x_1\overline{l_Q(x_2)}dt = [x_1, x_2]_\alpha^\beta = [x_1, x_2](\beta) - [x_1, x_2](\alpha). \tag{2.7}$$

As we will see in Section 2.3, we need to impose the requirement that l_Q be formally self-adjoint. We can do this by demanding that the matrix $Q \in Z_n(I)$ be Lagrange symmetric. That is , in addition to the conditions in (2.1), we shall require the following condition

$$Q = Q^+. \tag{2.8}$$

Here the Lagrange adjoint Q^+ of $Q \in Z_n(I)$ is defined by

$$Q^+ := -\Lambda_n^{-1} Q^* \Lambda_n, \tag{2.9}$$

where $Q^* = \bar{Q}^t$ (the conjugate transpose of Q, as usual), and $\Lambda_n = (\ell_{rs})$ is a certain fixed constant $n \times n$ matrix with $-1, +1, -1, +1 \ldots$ down the counter-diagonal and zeros elsewhere, that is ,

$$\ell_{rs} = \begin{cases} (-1)^r, & \text{for} \quad r+s = n+1, \\ 0, & \text{otherwise .} \end{cases} \tag{2.10}$$

Then easy computations and using the formulas

$$\Lambda_n^{-1} = \Lambda_n^t = (-1)^{n-1}\Lambda_n,$$

show that for $Q = Q^+$, the Lagrange-Green identity (2.7) can be written, for all $x_1, x_2 \in D(Q)$ and each compact interval $[\alpha, \beta]$ interior to I,

$$\int_\alpha^\beta \{l_Q(x_1)\overline{x_2} - x_1\overline{l_Q(x_2)}\}dt = [x_1, x_2](\beta) - [x_1, x_2](\alpha).$$

Thus when $Q = Q^+$ we observe that $[x_1, x_2](t) \equiv 0$ for all t in the complement of (supp$(x_1) \cap$ supp(x_2)) in I. Hence for $Q = Q^+$, we conclude that l_Q is formally self-adjoint, in the sense of Definition 2.3.

The classical expressions (2.4) can be seen as Lagrange symmetric quasi-differential expressions. To clarify these points, we consider the following examples.

Example 2.6

Let $n = 2$, then (2.4) is

$$Mx = a_0x - (a_1x')' + i\left[(b_0x)' + b_0x'\right] = a_0x - (a_1x')' + (ib_0x)' + ib_0x',$$

with $a_k \neq 0, k = 0, 1$ a.e. on I and a_1, b_0 are differentiable. We want to construct a matrix Q that belongs to $Z_2(I)$ and

$$M = l_Q.$$

So, we begin with

$$Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}.$$

and compute

$$\begin{aligned} x^{[0]} &= x, \\ x^{[1]} &= q_{12}^{-1}(x' - q_{11}x) = q_{12}^{-1}x' - q_{11}q_{12}^{-1}x, \\ x^{[2]} &= (q_{12}^{-1}x')' - (q_{11}q_{12}^{-1}x)' - q_{21}x - q_{22}(q_{12}^{-1}x' - q_{11}q_{12}^{-1}x), \\ &= (q_{12}^{-1}x')' - (q_{11}q_{12}^{-1}x)' - q_{21}x - q_{12}^{-1}q_{22}x' + q_{11}q_{12}^{-1}q_{22}x. \end{aligned}$$

Therefore, with the assumption that q_{12}^{-1} and $q_{11}q_{12}^{-1}$ are differentiable,

$$l_Q = i^2x^{[2]} = (q_{21} - q_{11}q_{12}^{-1}q_{22})x - (q_{12}^{-1}x')' + (q_{11}q_{12}^{-1}x)' + q_{12}^{-1}q_{22}x',$$

and solving $M = l_Q$ for Q gives

$$Q = \begin{bmatrix} ib_0a_1^{-1} & a_1^{-1} \\ a_0 - b_0^2a_1^{-1} & ib_0a_1^{-1} \end{bmatrix}.$$

Assuming now that

$$a_1^{-1}, ib_0a_1^{-1}, a_0 - b_0^2a_1^{-1} \in L_{loc}(I), a_0, a_1, b_0 \quad \text{rea real}$$

gives $M = l_Q$ with $Q \in Z_2(I)$ and $Q = Q^+$ as desired. □

Example 2.7

Let $n = 4$, then (2.4) is

$$Mx = a_0x - (a_1x')' + (a_2x'')'' + i\left[(b_0x)' + (b_1x')'' + b_0x' + (b_1x'')'\right].$$

Following the same process in Example 2.6, we end up with the following matrix.

$$Q = \begin{bmatrix} 0, & 1, & 0, & 0 \\ 0, & -ib_1a_2^{-1} & a_2^{-1}, & 0 \\ -ib_0, & a_1 - b_1^2a_2^{-1}, & -ib_1a_2^{-1}, & 1 \\ -a_0, & -ib_0, & 0, & 0 \end{bmatrix}. \tag{2.11}$$

A direct computation yields

$$l_Qx = \{(a_2x'' + ib_1x')' - a_1x' + ib_1x'' + ib_0x\}' + a_0x + ib_0x'.$$

This expression with

$$a_2^{-1},\ b_1a_2^{-1},\ a_1 + b_1^2a_2^{-1},\ a_0,\ b_0 \in L_{loc} \text{ and real}$$

is the quasi-differential analogue of the classical expression (2.4). It reduces to (2.4) with a_2, a_1, b_1 and b_0 are sufficiently differentiable. □

The matrices in Examples 2.6 and 2.7 belong to a relatively small subset of the set of all matrices Q such that $Q \in Z_2(I)$ and $Q = Q^+$. To see this, we illustrate with some examples.

Example 2.8

The general 2×2 matrix Q satisfying $Q \in Z_2(I)$ and the Lagrange symmetry condition $Q = Q^+$ is given by

$$Q = \begin{bmatrix} a & b \\ c & -\bar{a} \end{bmatrix}$$

where $b \neq 0$ a.e and b, c are real functions. Then l_Q is given by

$$l_Qx = -[b^{-1}(x' - ax)]' - \bar{a}b^{-1}(x' - ax) + cx \tag{2.12}$$

To relate (2.12) to (2.4), let $a = ib_0a_1^{-1} + d, b = a_1^{-1}, c = (a_0a_1 - b_0^2)a_1^{-1}$ where a_0, a_1, b_0 and d are real functions. Now (2.12) becomes

$$l_Q x = [-a_1 x' + (ib_0 + a_1 d)x]' + (ib_0 - a_1 d)x' + (a_0 + a_1 d^2)x. \tag{2.13}$$

When $d = 0$ and a_1, b_0 are differentiable, (2.13) can be written as

$$l_Q x = -(a_1 x')' + a_0 x + i\{(b_0 x)' + b_0 x'\}.$$

This is (2.4) for $n = 2$. When b_0 (but not necessarily d) is zero in (2.13) we get the general real symmetric expression

$$l_Q x = [a_1 x' + a_1 d x]' - a_1 d x' + (a_0 + a_1 d^2)x. \tag{2.14}$$

If a_1 and d are differentiable, (2.14) reduces to

$$l_Q = -(a_1 x')' + [(a_1 d)' + a_0 + a_1 d^2]x. \tag{2.15}$$

Finally when $d = 0$ (2.15) reduces to the familiar Sturm-Liouville operator

$$l_Q x = -(a_1 x')' + a_0 x.$$

□

Example 2.9

Let's examine the fourth order case. The general matrix, Q, satisfying $Q \in Z_4(I)$ and the symmetry condition $Q = Q^+$ is

$$Q = \begin{bmatrix} a & b & 0 & 0 \\ c & d & f & 0 \\ g & h & -\bar{d} & \bar{b} \\ k & -\bar{g} & \bar{c} & -\bar{a} \end{bmatrix} \tag{2.16}$$

with f, h and k real-valued and b, f not zero a.e.Then

$$l_Q x = (x^{[3]})' + \bar{a}x^{[3]} - \bar{c}x^{[2]} + \bar{g}x^{[1]} - kx \tag{2.17}$$

where

$$\begin{aligned} x^{[1]} &= b^{-1}(x' - ax), \\ x^{[2]} &= f^{-1}\left[(x^{[1]})' - dx^{[1]} - cx\right], \\ x^{[3]} &= \bar{b}^{-1}\left[(x^{[2]})' + \bar{d}x^{[2]} - hx^{[1]} - gx\right]. \end{aligned}$$

Observe that (2.11), which generates the expression (2.4) when $n = 4$ is a special case of (2.16). Thus (2.17) represents a much larger class of fourth order symmetric expressions than (2.4). Even in the case when all entries are real and $a = d = g = 0$ so that Q has the form

$$Q = \begin{bmatrix} 0 & b & 0 & 0 \\ c & 0 & f & 0 \\ 0 & h & 0 & b \\ k & 0 & c & 0 \end{bmatrix}$$

we get a more general real fourth-order expression than is normally considered. Letting $b = p^{-1}$ and $f = r^{-1}$ we have

$$l_Q x = (p((r((px')' - cx))' - hpx'))' - cr((px')' - cx) - kx.$$

For $p = 1$ and $c = 0$ in the last expression, we get the more familiar form

$$l_Q x = [(rx'')' - hx']' - kx. \tag{2.18}$$

It should be noted, see [21], that Naimark's development of the even-order real case in [33, chapter V] is based on conditions that $r^{-1}, h, k \in L_{loc}(I)$. Under these conditions the quasi-differential expressions have the form (2.18) not

$$(rx'')'' - (hy')' - ky$$

as stated in [33]. □

2.1.1 Properties of Quasi-Differential Equations

Now, fix $Q \in Z_n(I)$, let $\lambda \in \mathbb{C}, w, g \in L_{loc}(I)$ and consider the following quasi-differential equation

$$l_Q x = \lambda w x + g \quad \text{a.e.} \quad \text{on} \quad I. \tag{2.19}$$

with $w(t) > 0$ a.e. on I.

Definition 2.10

A solution of (2.19) is a function $x : I \to \mathbb{C}$ such that $x^{[k]} \in AC_{loc}(I)$ for all $k = 0, \cdots, n-1$ and satisfies (2.19) a.e. on I. □

Definition 2.11

Given a vector function G and a matrix function $A : I \to \mathbb{C}^{n \times n}$, we define a solution of

$$Y' = AY + G \tag{2.20}$$

to be a vector function $Y : I \to \mathbb{C}^n$ such that Y belongs to $AC_{loc}(I)$, component-wise, and satisfies (2.20) a.e.on I. □

It follows from the definition of l_Q that (2.19) is equivalent to (2.20) with

$$Y = \begin{bmatrix} x^{[0]} \\ x^{[1]} \\ \vdots \\ x^{[n-1]} \end{bmatrix}, \quad A = Q + \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ i^{-n}\lambda w & 0 & \cdots & 0 \end{bmatrix}, \quad \text{and} \quad G = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ g \end{bmatrix}.$$

This means that for a given solution of (2.19) if we form Y, A and G as indicated above, then Y is a solution to (2.20) and conversely, for A and G of the above form if Y is a solution of (2.20) then the first component of Y is a solution of (2.19). The proof of following existence and uniqueness theorem, see [33, 18], takes advantage of this equivalence.

Theorem 2.12

Let $Q \in Z_n(I)$ and let $w, g \in L_{loc}(I)$ with $w(t) > 0$ a.e.on I. Then for any $\lambda \in \mathbb{C}$, any $t_0 \in I$ and $c_k \in \mathbb{C}(k = 0, \cdots, n-1)$ there exists a unique solution defined on I of the initial value problem

$$l_Q x = \lambda w x + g, \quad \text{a.e.} \quad \text{on} \quad I.$$
$$x^{[k]}(t_0) = c_k, \qquad k = 0, \cdots, n-1.$$

Furthermore, if g, c_k and all entries of Q are real-valued, then the unique solution is also real. ■

Proof.

See [33, Chapter V] and [10]. ▮

Definition 2.13

Let $x_1, x_2, \cdots, x_n$ be functions for which ${x_\rho}^{[\sigma]}, \sigma = 0, 1, \cdots, n-1, \rho = 1, \cdots, n$ exist. Then we define the Wronskian $W = W(x_1, x_2, \cdots, x_n)$ as follows

$$W = (w_{rs})_{r,s=1}^n \quad \text{where} \quad w_{rs} = {x_s}^{[r-1]}, \quad 1 \leq r, s \leq n.$$ □

Theorems 2.15, 2.14 and 2.16 are stated for the sake of completeness. The proofs of these theorems are given in [46].

Theorem 2.14

The set of all solutions of $l_Q x - \lambda w x = 0$ forms an $n-$dimensional vector space over $\mathbb{C}$. Furthermore, if all entries of Q are real, then the set of real solutions forms an $n-$dimensional vector space over $\mathbb{R}$. ■

Theorem 2.15

Suppose that $x_1, x_2, \cdots, x_n$ are solutions of $l_Q x - \lambda w x = 0$. If $x_1, x_2, \cdots, x_n$ are linearly dependent on I, then $W(t) \equiv 0$ for every $t \in I$. If for some $t_0 \in I$, $W(t_0) = 0$, then $x_1, x_2, \cdots, x_n$ are linearly dependent. ■

Theorem 2.16

Suppose that $g \in L_{loc}(I)$ and $x_1, x_2, \cdots, x_n$ are linearly independent solutions of $l_Q x - \lambda w x = 0$. Let $t_0 \in I$, and let

$$v_k = (-1)^{n+k} \frac{W(x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n)}{W(x_1, x_2, \cdots, x_n)}.$$

Then, if $l_Q x = \lambda w x + g$, there exist $\alpha_1, \cdots, \alpha_n \in \mathbb{C}$ such that

$$x(t) = \sum_{k=1}^{n} \alpha_k x_k(t) + \sum_{k=1}^{n} x_k(t) \int_{t_0}^{t} v_k(\tau) g(\tau) d\tau,$$

for each $t \in I$. Moreover for any choice of the α_k, the above formula gives a solution of $l_Q x = \lambda w x + g$. ■

Definition 2.17 (Regular and Singular Expression)

We say that the expression l_Q is regular at a, if $a > -\infty$, $Q \in Z_n([a, b))$ and $w \in L_{loc}([a, b))$. Similarly one defines regularity at b. The expression l_Q is called regular if it is regular at a and at b. If l_Q is not regular at a (resp. b), it is said to be singular at a (resp. b). The expression l_Q is said to be singular if it is singular at a or at b. □

Remark 2.18

The above definition implies that l_Q is regular if and only if it is regular at each point t in $I = [a, b]$; this is due to the construction of $Z_n([a, b])$ and the assumption on w. On the other hand, its singularity at a occurs if either $a = -\infty$ or $a \in \mathbb{R}$ but

$$\int_a^c \{|q_{r_0, s_0}(t)| + w(t)\} dt = \infty \quad \text{for some} \quad c \in (a, b) \quad \text{and for some} \quad 1 \leq r_0, s_0 \leq n.$$ □

Remark 2.19

The expression l_Q can be regular at a even though its leading coefficient is zero at a. For example in (2.18) on $I = [a, \infty)$ if h, k are in $L_{loc}(I)$ and $r(t) = \sqrt{t-a}$ for all $t \in I$, then l_Q is regular at a and so the only singular point of l_Q is ∞. This case was called weakly singular in the literature.□

2.2 Deficiency Indices

In this section we define the deficiency indices of symmetric differential operators and state the basic classification results for them.

Definition 2.20

A linear operator A from a separable Hilbert space $\mathcal{H}$ into $\mathcal{H}$ is said to be symmetric if it is Hermitian and its domain $D(A)$ is dense in $\mathcal{H}$, i.e.,

$$\langle Af, g\rangle = \langle f, Ag\rangle \quad \text{for all } f, g \text{ in } D(A) \qquad \square$$

It is clear that an operator A, with a domain of definition dense in $\mathcal{H}$, is symmetric if and only if

$$A \subset A^*.$$

Such an operator has associated with it a pair d^+, d^- where each of d^+ and d^- is a nonnegative integer or $+\infty$. The extended integers d^+, d^- are called the deficiency indices of A and are defined as follows.

Definition 2.21

Let A be a symmetric operator and let λ be a non-real complex number and denote by R_λ the range of $(A - \lambda E)$, E being the identity operator. Define the deficiency space N_λ by

$$N_\lambda = R_{\overline{\lambda}}^{\perp} = \mathcal{H} \ominus R_{\overline{\lambda}}.$$

In other words N_λ is the orthogonal complement in $\mathcal{H}$ of the range of the operator $A - \overline{\lambda}E$. We also define d^+ and d^- by

$$d^+ = \dim(N_{+i}), \quad d^- = \dim(N_{-i}). \qquad \square$$

For the convenience of the reader we recall a few elementary facts from the abstract theory of symmetric operators in Hilbert space. These are well-known; for proofs the reader is referred to

[1, 33].

Lemma 2.22

If λ, μ are both in the upper-half of the complex plane or are both in the lower-half of the plane, then

$$\dim(N_\lambda) = \dim(N_\mu).$$ □

Lemma 2.23

For any non-real number λ, the deficiency spaces $N_\lambda, N_{\bar{\lambda}}$ of the symmetric operator A are the eigenspaces of A^*, the adjoint of A, belonging to $\lambda, \bar{\lambda}$ respectively. In other words, for any non-real complex number λ

$$N_\lambda = \{f \in D(A^*) : A^*(f) = \lambda f\}.$$ □

The next lemma is known as von Neumann's formula for the domain of the adjoint.

Lemma 2.24

Let A be a symmetric operator. Then for any non-real number λ,

$$D(A^*) = D(A) \oplus N_\lambda \oplus N_{\bar{\lambda}};$$

with $D(A), N_\lambda, N_{\bar{\lambda}}$ linearly independent and the sum is a direct sum. □

Definition 2.25

An operator B with domain $D(B)$ is said to be an extension of an operator A with domain $D(A)$, and we write $A \subset B$, if

(1) $D(A) \subset D(B)$ and

(2) $A = B_{D(A)}$, i.e., A coincides with B if B is restricted to $D(A)$. □

If B and A are symmetric and $A \subset B$, then $B^* \subset A^*$; but B is symmetric, i.e. $B \subset B^*$; and so we get

$$A \subset B \subset B^* \subset A^*.$$

Definition 2.26

An operator A with domain $D(A)$ which is dense in a Hilbert space $\mathcal{H}$ is said to be self-adjoint if $A = A^*$. ☐

Lemma 2.27

A symmetric operator A has a self-adjoint extension if and only if its deficiency spaces N_λ and $N_{\overline{\lambda}}$ have the same dimension, i.e.,

$$\text{dim}(N_\lambda) = \text{dim}(N_{\overline{\lambda}}).$$ ☐

Definition 2.28

A symmetric operator A is said to be semi-bounded from below if there is a number M such that, for all $x \in D(A)$, the identity

$$\langle Ax, x\rangle \leq M\|x\|^2$$

holds. Similarly A is said to be semi-bounded from above if for all $x \in D(A)$ there is a number m such the inequality

$$\langle Ax, x\rangle \geq m\|x\|^2$$

holds. In the special case when $m = 0$, A is said to be positive. ☐

Definition 2.29

Let V be a normed space and A be any linear operator on V. The domain of regularity, $\mathbb{C}_{reg}(A)$, of A is the set

$$\mathbb{C}_{reg}(A) = \left\{\mu \in \mathbb{C} : R_\mu := (A - \mu E)^{-1} \text{ exits and bounded } \right\}.$$

A point $\mu \in \mathbb{C}_{reg}(A)$ is called a point of regular type. Furthermore, μ is called a regular point of A if $\mu \in \mathbb{C}_{reg}(A)$ and $D(R_\mu) = V$, in this case R_μ is called the resolvent of the operator A. ☐

Lemma 2.30

If A is a positive symmetric operator, then the negative semi-axis belongs to its domain of regularity $\mathbb{C}_{reg}(A)$. □

It is should be mentioned that if A is self-adjoint then $\mathbb{C}_{reg}(A)$ contains all non-real numbers.

2.3 Minimal and Maximal Operators

Symmetric differential expressions generate symmetric differential operators in an appropriate Hilbert space, in particular the so-called minimal operator. In general this minimal operator is not self-adjoint but has self-adjoint extensions.

Let $n > 1$ be an integer, $I = (a,b), -\infty \leq a < b \leq \infty$, $Q \in Z_n(I), Q = Q^+$ and w be a positive weight function that is locally integrable on I. Now we consider the quasi-differential expression

$$w^{-1}l_Q x = i^n w^{-1} x^{[n]},$$

which has the domain $D(Q)$, where

$$D(Q) = \left\{x : I \to \mathbb{C} : x^{[k]} \in AC_{loc}(I), k = 0, \cdots, n-1\right\}.$$

In this section we define the minimal and maximal operators associated with $w^{-1}l_Q$ and develop their basic properties leading to the Glazman-Krien-Naimark (GKN) Theorem. Indeed, $w^{-1}l_Q$ generates a maximal operator L_1 on

$$D_1 = D(L_1) = \left\{x \in D(Q) : x \text{ and } w^{-1}l_Q x \in \mathcal{L}^2(I,w)\right\},$$

where

$$\mathcal{L}^2(I,w) = \left\{x : I \to \mathbb{C} : \int_I |x|^2 w dt < \infty\right\}$$

with the inner product

$$\langle x, y\rangle = \int_I x\bar{y}w dt \quad \text{for } \ x, y \in \mathcal{L}^2(I, w).$$

It is clear that D_1 is a linear manifold in $\mathcal{L}^2(I, w)$. It is the largest manifold in $\mathcal{L}^2(I, w)$ on which the operator can be defined in this way. For, the requirement that the quasi-derivatives $x^{[k]} AC_{loc}(I),\ k = 0, \cdots, n-1$ is necessary in order that the expression $w^{-1}l_Q$ shall make sense, and the requirement that $w^{-1}l_Q \in \mathcal{L}^2(I, w)$ is necessary in order that $w^{-1}l_Q$ shall define an operator on $\mathcal{L}^2(I, w)$.

The following lemma called, Patching Lemma, is of great importance in the development of operators generated by $w^{-1}l_Q$ on I.

Lemma 2.31

Let $J = [\alpha, \beta]$ be a compact set, $Q \in Z_n(J), Q = Q^+$ and $w \in L_{loc}(J)$ be a positive weight on the interval J, so the quasi-differential expression $w^{-1}l_Q$ generates a maximal linear operator L_1 on $D_1 \subset \mathcal{L}^2(J, w)$. Then for arbitrary vectors

$$\xi \in \mathbb{C}^n, \quad \eta \in \mathbb{C}^n,$$

there exists an n-vector function $y(t)$ for $t \in J$, with components $x^{[k]}, k = 0, \cdots, n-1$ such that

$$y(\alpha) = \xi, \quad y(\beta) = \eta.$$

Furthermore, $x = x^{[0]} \in D_1$. □

Proof.

See Naimark [33]. ∎

Now let $D_0' = D_0'(Q)$ denote the set of all functions in D_1 which vanish outside of a compact subinterval, which may be different for different functions, of the interior of I, that is

$$D_0' = \{x \in D_1 : \text{supp}(x) = [\alpha, \beta], \text{ for some } [\alpha, \beta] \subset I\}.$$

Define

$$L_0' x = L_0'(Q)x = w^{-1} l_Q x \quad \text{ for all } x \in D_0'.$$

In other words, $L_0' = (L_1)_{D_0'}$ or $L_0' \subset L_1$.

Lemma 2.32

(i) If x is in D_0' and y is in D_1, then

$$\langle L_0' x, y\rangle = \langle x, L_1 y\rangle.$$

(ii) The operator L_0' is Hermitian, that is,

$$\langle L_0' x, y\rangle = \langle x, L_0' y\rangle \quad \text{for all } \; x, y \in D_0'.$$

(iii) The set D_0' is dense in $\mathcal{L}^2(I, w)$. □

Proof.

See [18, Section 6]. ∎

Lemma 2.32 shows that L_0' is symmetric and therefore admits a closure. We denote the closure of L_0' by L_0, i.e.,

$$L_0 = \overline{L_0'}.$$

This operator is called the minimal operator generated by $w^{-1} l_Q$ on I. Let $D_0 = D_0(Q)$ be the domain of L_0.

Lemma 2.33

(i) $L_0 = L_1^*($ the adgoint of $L_1)$ and $L_0^* = L_1$.

(ii) If x is in D_0 and a is a regular end point of I, then

$$x^{[k]}(a) = 0, \quad k = 0, 1, \cdots, n-1.$$

□

Proof.

See [33, 18]. ∎

Lemma 2.24 gives the following direct sum of D_1

$$D_1 = D_0 \oplus N_{+i} \oplus N_{-i}.$$

where the deficiency spaces N_{+i} and N_{-i} of L_0 are defined as

$$\begin{aligned} N_{+i} &= \{x \in D(L_0^*) : L_0^* x = ix\} \\ &= \{x \in D_1 : L_1 x = ix\} \\ &= \left\{x \in D(Q) \cap \mathcal{L}^2(I, w) : w^{-1} l_Q x = ix\right\} \end{aligned}$$

and

$$\begin{aligned} N_{-i} &= \{x \in D(L_0^*) : L_0^* x = -ix\} \\ &= \{x \in D_1 : L_1 x = -ix\} \\ &= \left\{x \in D(Q) \cap \mathcal{L}^2(I, w) : w^{-1} l_Q x = -ix\right\} \end{aligned}$$

From this we can conclude that d^+, the upper deficiency index, is the maximum number of linearly independent solutions of

$$w^{-1} l_Q x = \lambda x \text{ on } I \tag{2.21}$$

in the space $\mathcal{L}^2(I, w)$ for any λ in the upper half-plane and d^-, the lower deficiency index, is the maximum number of linearly independent solutions of (2.21) in the space $\mathcal{L}^2(I, w)$ for any λ in the lower half-plane. By Lemma 2.22, d^+ is independent of the particular number chosen in the complex upper half plane. Similarly d^- does not depend on the particular number λ chosen from the complex lower half-plane. Thus d^+ and d^- depend only on the coefficients of Q, w, and I. We indicate this dependence by writing

$$d^+ = d^+(Q, w, I) \quad \text{and} \quad d^- = d^-(Q, w, I).$$

Since (2.21) has exactly n linearly independent solutions we see that the integers d^+, d^- must satisfy the basic inequality

$$0 \leq d^+, \quad d^- \leq n.$$

Observe that if l_Q is regular on a compact interval I then $d^+ = d^- = n$ since in this case all solutions of (2.21) are in $\mathcal{L}^2(I, w)$. If l_Q is real, i.e. the entries of Q are all real-valued, we have the following lemma.

Lemma 2.34

Let l_Q be real. Then the deficiency indices of the minimal operator L_0 generated by $w^{-1}l_Q$ satisfy

$$0 \leq d^+ = d^- \leq n. \qquad \square$$

Let $c \in I$ be any point in I. Then we have the following result sometimes referred to as Kodaira's formula.

Lemma 2.35

$$d^+(Q, w, (a, b)) = d^+(Q, w, (a, c)) + d^+(Q, w, (c, b)) - n. \qquad \square$$

Proof.

See [33, Section 17.5]. ∎

It is this theorem that reduces the problem of computing d^+ or d^- on intervals with two singular end points to the case with only one singular end point. We now state the basic classification result for deficiency indices of general symmetric differential expressions on intervals with one regular and one singular end point.

Theorem 2.36

Suppose l_Q is regular at each point of an interval $I = [a, b)$ but the end point b is singular. Then the deficiency indices d^+, d^- of l_Q satisfy the inequalities.

(a) If $n = 2m$ $(m \geq 1)$ is even then

$$\frac{1}{2}n = m \leq d^+(Q, w, I), d^-(Q, w, I) \leq 2m = n \tag{2.22}$$

(b) If $n = 2m + 1$ $(m \geq 1)$ is odd then

(i) when m is even

$$\begin{aligned} \tfrac{1}{2}(n-1) &= m &\leq d^+(Q,w,I) &\leq 2m+1 = n, \\ \tfrac{1}{2}(n+1) &= m+1 &\leq d^-(Q,w,I) &\leq 2m+1 = n. \end{aligned} \tag{2.23}$$

(ii) when m is odd

$$\begin{aligned} \tfrac{1}{2}(n+1) &= m+1 &\leq d^+(Q,w,I) &\leq 2m+1 = n \\ \tfrac{1}{2}(n-1) &= m &\leq d^-(Q,w,I) &\leq 2m+1 = n. \end{aligned} \tag{2.24}$$

All these inequalities are best possible. ■

Proof.

See [18]. ▮

Lemma 2.37

If all solutions of

$$w^{-1}l_Q x = \lambda x \text{ on } I \tag{2.25}$$

are in $\mathcal{L}^2(I, w)$ for some λ in $\mathbb{C}$ then all solutions of (2.25) are in $\mathcal{L}^2(I, w)$ for any λ in $\mathbb{C}$.(Note that

λ is allowed to be real in this theorem.) □

Proof.

See [18, Theorem 9.1]. ▮

In view of Lemma 2.35, we can restrict ourselves to the case when I has one regular and one singular end point. Let $I = [a, b)$ where a is a regular and b singular. In summary we may make the following remarks

(1) If all the coefficients of l_Q are real, then $l_Q x = \lambda w x$ if and only if $l_Q \bar{x} = \bar{\lambda} w \bar{x}$. From this and the fact that $x \in \mathcal{L}^2(I, w)$ if and only if $\bar{x} \in \mathcal{L}^2(I, w)$ it follows that $d^+ = d^-$ in this case.

(2) Let $n = 2m$ be given; then any integer between m and $2m$ occurs as the deficiency index of some symmetric expression, see [36].

(3) The lower bounds for d^+, d^- given by (2.23) are achieved by simple odd order constant coefficient expressions l_Q.

(4) In [27], it was shown that d^+, d^- can be different also in the even order complex coefficient case.

(5) In [22], it was shown that all possibilities not ruled out by Theorem 2.36 and

$$|d^+ - d^-| \leq 1$$

actually occur.

2.4 Self Adjoint Extensions

Given a Lagrange symmetric (formally self-adjoint) differential expression l_Q, i.e. $Q \in Z_n(I),\ Q = Q^+$, and a positive weight function w, we consider self-adjoint realizations of the equation

$$l_Q x = \lambda w x \quad \text{on} \quad I = (a, b),\ -\infty \leq a < b \leq \infty \tag{2.26}$$

in the Hilbert space $\mathcal{L}^2(I,w)$. A self-adjoint realization of (2.26) in the Hilbert space $\mathcal{L}^2(I,w)$ (or self-adjoint extension of L_0) is an operator L satisfying

$$L_0 \subset L = L^* \subset L_1.$$

Lemma 2.27, Section 2.2 page 22, asserts that there exits self-adjoint L to L_0 if and only if $d^+ = d^-$. Unfortunately, See Section 2.3, this not the case in general. Therefore, we assume that

$$d = d^+ = d^-.$$

The deficiency index $d = 0$ if and only if $L_0 = L_1$, in which case L_0 is the only self-adjoint operator generated by $w^{-1}l_Q$ on $\mathcal{L}^2(I,w)$.

Since D_0 is a linear subspace in the Hilbert space D_1, we can construct the quotient or identification D_1/D_0, consisting of D_0-cosets like $\{x + D_0\}$ for each $x \in D_1$. This leads to the following definition.

Definition 2.38

Consider the maximal and minimal operators L_1 on D_1 and L_0 on D_0, respectively, as generated by $w^{-1}l_Q$ on $\mathcal{L}^2(I,w)$. Then define the quotient space

$$\mathfrak{S} = D_1/D_0, \qquad \square$$

which is a complex vector space of dimension $(d^+ + d^-) \le 2n$. Further denote the natural projection of D_1 onto $\mathfrak{S}$

$$\mathfrak{P} : D_1 \to \mathfrak{S}, \quad x \to \mathfrak{P}x = \{x + D_0\}, \tag{2.27}$$

and we introduce the notation, for each $x \in D_1$,

$$\hat{x} = \mathfrak{P}x, \quad \hat{x} \in \mathfrak{S}, \quad (\text{where } \hat{x} = \{x + D_0\}). \tag{2.28}$$

Self-adjoint extensions of L_0 are characterized by describing their domains. The following theorem is a version of the highly celebrated Glazman-Krein-Naimark (GKN-EZ) as extended by Everitt and Zettl [19, 15].

Theorem 2.39

Consider the quasi-differential expression

$$w^{-1}l_Q x = i^n w^{-1} x^{[n]}$$

on the interval I with $Q = Q^+ \in Z_n(I)$ and assume that

$$0 \leq d = d^+ = d^- \leq n, \quad (\text{for } n > 1).$$

Let L_1 on D_1 and L_0 on D_0 be the maximal and minimal operators , respectively, as generated by $w^{-1}l_Q$ on $\mathcal{L}^2(I, w)$. Then there exists a one-to-one correspondence between the set $\{L\}$ of all self-adjoint operators L on $D(L)$, as generated by $w^{-1}l_Q$ on $\mathcal{L}^2(I, w)$, and the set of $\{\mathfrak{L}\}$ of all $d-$space $\mathfrak{L}$ in the complex $2d-$space $\mathfrak{S} = D_1/D_0$. Namely, take the correspondence $L \leftrightarrow \mathfrak{L}$ given by the bijection

$$\mathcal{E} : \{\mathfrak{L}\} \to \{L\},$$

defined as

$$\mathcal{E}(\mathfrak{L}) = \mathfrak{P}^{-1}\mathfrak{L},$$

where $\mathfrak{P} : D_1 \to \mathfrak{S}$, as in (2.27). Hence, we conclude that

$$x \in D(L) \quad \text{if and only if} \quad \hat{x} \in \mathfrak{L},$$

or that $D(L)$ is precisely the pre-image of $\mathfrak{L}$ under the natural projection

$$\mathfrak{P} : D(L)(\subset D_1) \to \mathfrak{L} \subset \mathfrak{S},$$

that is,

$$D(L)/D_0 = \mathfrak{L}.$$

■

Proof.

See [16, Section II, Theorem 1]. ■

This theorem says that for each set of functions $x_1, x_2, \cdots, x_d \in D_1$ such that $\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_d$ is a basis for $\mathfrak{L}$ (that is $[x_r, x_s] = 0$ for $1 \leq r, s \leq d$) the domain $D(L)$ of the corresponding self-adjoint operator L is

$$D(L) = \{x \in D_1 : [x, x_s] = 0 \text{ for } s = 1, \cdots, d\},$$

or equally,

$$D(L) = c_1 x_1 + \cdots + c_d x_d + D_0,$$

where $c_1, \cdots, c_d$ are arbitrary complex constants. Therefore,

$$[x, x_s] = 0 \text{ for } s = 1, \cdots, d$$

are d homogeneous linear boundary conditions determining the function x in $D(L)$.

The GKN Theorem (Theorem 2.39) characterizes all self-adjoint realizations of linear symmetric (formally self-adjoint) ordinary differential equations in terms of maximal domain functions. These functions depend on the coefficients and this dependence is implicit and complicated. In the regular case an explicit characterization in terms of two-point boundary conditions can be given. In the singular case when the deficiency index d is maximal the GKN characterization can be made more explicit by replacing the maximal domain functions by a solution basis for any real or complex value of the spectral parameter λ. In the much more difficult intermediate cases, not all solutions contribute to the singular self-adjoint conditions.

The characterization of self-adjoint extensions is still an active area of research see for example [43, 40, 44, 12].

We conclude this section by giving the following theorem that characterizes the resolvents of self-adjoint extensions of the operator L_0.

Theorem 2.40

For a point of regular type μ, the resolvent $R_\mu = (L - \mu E)^{-1}$ (E the identity operator on $\mathcal{L}^2(I, w)$) of an arbitrary self-adjoint extension of the operator L_0 is an integral operator whose kernel satisfies the conditions

$$\int_I |K(t,s,\mu)|^2 \, ds < \infty \quad \text{for all } t \in I,$$

$$\int_I |K(t,s,\mu)|^2 \, dt < \infty \quad \text{for all } s \in I.$$

For an operator L_0 with deficiency indices (n, n), the kernel $K(t, s, \mu)$ is a Hillbert-Schmidt kernel, i.e., it satisfies

$$\int_I \int_I |K(t,s,\mu)|^2 \, dt ds < \infty.$$

■

Proof.

See [33, §19.3]. ▮

Chapter 3

OPTIMAL CONTROL

Late in 1950s, Pontryagin and his coworkers with their development of the maximum principle laid down the foundation stone of Optimal Control as a distinct area of research. Optimal Control theory is an outcome of the calculus of variations, with a history that goes back to over three hundred years. Optimal Control addresses in a unified way many optimization problems arising in many scientific fields ranging from mathematics and engineering to biomedical and management sciences. Aerospace engineering is considered a rich supply of problems beyond the reach of traditional analytical and computational methods. During the 1960s and 1970s the American and Russian space programs gave a lot of momentum to the field of Optimal Control.

This chapter is organized as follows. We define and discuss the optimal control problem in Section 3.1. Sections 3.2 and 3.3 describe necessary optimality conditions. In Section 3.2, we discuss such conditions with no constraints on the control set. In Section 3.3, we derive necessary optimality conditions in the form the Pontryagin maximum principle (PMP). A historical review on the development of optimal control theory is given in Section 3.4.

3.1 The Optimal Control Problem

An optimal control problem (OCP) is typically an optimization problem where the objective is to find a vector or, more generally, a function u[1], called the control, that causes a system to satisfy some physical constraints and at the same time optimizes a performance criterion. In optimal control, one seeks a solution to the following problem

$$\left.\begin{aligned} &\text{minimize } \mathcal{J}[u,x] := \phi_0(x(b)) \\ &\text{subject to} \\ &\qquad u(t) \in U, \quad \text{a.e.} \\ &\qquad \dot{x}(t) = f(x(t),u(t),t),\ x(a) = x_0 \quad t \in I = [a,b] \end{aligned}\right\} \qquad \text{(OCP)}$$

In (OCP), the variable $x(t)$, called the state (or phase) variable, at instant t is an element of a Banach space X, called the state-space. The function ϕ_0 is real-valued on X and is assumed to be differentiable; though this assumption can be relaxed (cf. [31]). The set U constitutes a metric space and the control u is required to be an element of U at any instant of time t in the closed interval I almost everywhere. The function $f : (x,u,t) \to X$ is a function of $X \times U \times I$ into X.

A deeper look at (OCP) tells that a typical optimal control problem is governed by a dynamical system that itself is to be managed by the controls u while constrained point-wisely in U. The control input steers a system from a prescribed initial state, $x(a) = x_0$, to some final state in an optimal manner; that is maximizing or minimizing a certain performance criterion.

[1]"u" was chosen because it is the first letter of the Russian word "upravlenie" meaning "control".

3.1.1 Dynamical Systems

The ordinary differential equation

$$\dot{x}(t) = f(x(t), u(t), t),\ x(a) = x_0, \quad t \in I \tag{3.1}$$

is an important part of the optimal control problem. It describes the underlying physical aspects of the system. Here t is the independent variable, usually called time. Systems where the function f does not depend explicitly on time are called autonomous. Systems can also be classified into linear and nonlinear depending on whether f is linear or nonlinear.

A solution $x(\cdot)$ of (3.1) is called a response of the system corresponding to the control $u(\cdot)$ for the initial condition $x(a) = x_0$. Precisely, a solution to (3.1) is define as follows.

Definition 3.1

A solution $x(\cdot)$ to the differential equation (3.1) is a function $x : I \to X$ that is Fréchet differentiable for a.e. $t \in I$ and satisfies (3.1) and the following formula, called Newton-Leibniz.

$$x(t) = x(a) + \int_a^t f(x(s), u(s), s)ds, \quad \text{for all } t \in I. \tag{3.2}$$

□

It is well known that for $X = \mathbb{R}^n, x(t)$ is a.e. differentiable on T and satisfies the Newton-Leibniz formula if and only if it is absolutely continuous on I. However, for infinite-dimensional spaces X even the Lipschitz continuity may not imply the a.e. differentiability. On the other hand, there is a complete characterization of Banach spaces X, where the absolute continuity of every $x : I \to X$ is equivalent to its a.e. differentiability and the fulfillment of the Newton-Leibniz formula. This is the class of spaces with the so-called Radon-Nikodym property (RNP).

Definition 3.2 (Radon-Nikodym property)

A Banach space X has the Radon-Nikodym property if for every finite measure space (Ξ, Σ, μ) and for each μ-continuous vector measure $m : \Sigma \to X$ of bounded variation there is $g \in L^1(\mu; \Xi)$ such that

$$m(E) = \int_E g \, d\,\mu \quad \textit{for} \quad E \in \Sigma.$$

□

It is important to observe that the latter list contains every reflexive space and every weakly compactly generated dual space, hence all separable duals. On the other hand, the classical spaces $l^\infty, L^1[0,1]$, and $L^\infty[0,1]$ don't have the RNP.

Throughout this chapter, we shall assume that the function f is continuous in x, u, t and continuously differentiable with respect to x.

3.1.2 Admissible Controls

The system under consideration, (3.1), is assumed to be controllable. In other words, the system is equipped with controllers that direct its behavior over the course of its progression. These controllers are the control variables u. In optimal control problems the control variables are confined to belong to a specific control region U, which might be any set of a metric space. In many applications, the region U is chosen to be closed and bounded. The physical meaning of this choice is usually obvious. For example, the amount of temperature, current, voltage, fuel injected in an engine, etc. can be taken as control variables and clearly these quantities cannot take on arbitrary large values.

The choice of a control region and the control variables lead to the following definition.

Definition 3.3

An admissible control $u(\cdot)$ is a measurable function defined on some interval $[a, b]$ and satisfies the point-wise constraint

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

Sometimes, we will refer to the collection of all admissible control-trajectory pairs, denoting it by $\mathcal{A}$, to mean the set generated by the controls $u(\cdot)$ in the sense of Definition 3.3 and the corresponding trajectories in the sense of Definition 3.1.

3.1.3 Performance Measure

A performance measure (also called effectiveness criterion or cost functional) is a mathematical expression designed in a way that gives a quantitative assessment of the system performance and indicates, when optimized, a desirable behavior from the system. The performance measure is chosen to translate the physical requirement of the system into mathematical terms.

In general, the performance measure $\mathcal{J}$ is a functional from $\mathcal{A}$ to $\mathbb{R}$ and may be defined as

$$\mathcal{J}[u, x] := \phi(x(b)), \tag{3.3}$$

where $\phi : X \to \mathbb{R}$ is a real-valued function. This form is called the Mayer form and the endpoints of I, although can be considered as free variables, will be fixed. It will be assumed that both ϕ and ϕ_x are continuous.

The performance measure may also be written in Lagrange form as,

$$\mathcal{J}[u, x] := \int_a^b l(x(t), u(t), t)dt. \tag{3.4}$$

Here $l : X \times U \times I \to \mathbb{R}$ is a real-valued function and is assumed to be continuous together with its derivative with respect to x.

A more general form of the cost functional is the so-called Bolza form. This form combines a terminal term and an integral term as follows

$$\mathcal{J}[u,x] := \phi(x(b)) + \int_a^b l(x(t),u(t),t)dt. \tag{3.5}$$

Mathematically, these forms are equivalent. Introducing a new state variable $\tilde{x} = x + y$ such that

$$\dot{y}(t) = l(x(t),u(t),t) \quad \text{for all } t \in I, \text{ and } y(a) = 0$$

transforms the cost functional from Lagrange form (3.4) to Mayer form (3.3) with

$$\phi(\tilde{x}(b)) = y(b).$$

On the other hand, Mayer form (3.3) can be readily seen as Lagrange form (3.4) with

$$l(x(t),u(t),t) = \frac{1}{b-a}\phi(x(b)) \quad \text{for all } t \in I.$$

Lastly, Bolza form (3.5) can be written in either Mayer or Lagrange form using the above techniques. Conversely both forms (3.3) and (3.4) can be seen as special cases of Bolza form (3.5) with $l \equiv 0$ in the first and $\phi \equiv 0$ in the second.

3.1.4 Constrained OCP

Different kinds of extra constraints may be imposed on OCP that restrict both the state and the control variables. In an optimal control problem, point constraints, path constraints or isoperimetric constraints can be enforced as equality or inequality constraints.

- POINT CONSTRAINTS or terminal constraints are sometimes used to force the optimal trajectory to belong to a specific set at the terminal time. These may occur as inequality constraints like

$$\psi(x(b)) \leq 0,$$

or as equality constraints like

$$\psi'(x(b)) = 0.$$

- ISOPERIMETRIC CONSTRAINTS. An isoperimetric constraint is one that involves the integral of a given functional over part or all of I.

  $$\int_a^b h(x(t), u(t), t)dt \leq C.$$

  A problem with isoperimetric constraints can be equivalently transformed to one with terminal constraints in the same manner we transformed the Lagrange form of the cost functional to the Mayer form.

- PATH CONSTRAINTS. Equality or inequality type constraints can be used to restrict the state and control variables over the entire interval I or any nonempty subinterval. For example a path constraint may be introduced as

  $$\Psi(x(t), u(t), t) \leq 0, \quad t \in I.$$

Definition 3.4 (Feasible control-trajectory pair)

An admissible control u is said to be feasible if

1. the corresponding trajectory x is defined over the entire interval I, and
2. both of u and x satisfy all the (point and path) constraints over I.

The pair (u, x) is then called a feasible pair. □

Before we discuss the optimality conditions, we give the following definition for a global optimal solution to the (OCP).

Definition 3.5 (Global Optimal Pair)

A feasible pair $(\bar{u}(\cdot), \bar{x}(\cdot))$ to (OCP) and any other physical constraints is said to be optimal if

$$\mathcal{J}[\bar{u}, \bar{x}] \leq \mathcal{J}[u, x] \qquad \text{for all} \quad (u, x) \in \mathcal{A}. \qquad \square$$

Although many difficulties are to be expected when studying the existence problem of an optimal solution or even a feasible one, we will assume the existence of such an optimal pair $(\bar{u}(\cdot), \bar{x}(\cdot))$. We shall discuss in this section different approaches to describe optimality conditions which any optimal pair must satisfy.

To draw a more complete picture of the development of optimal control and to walk in the footsteps of the pioneers of the field, we shall give, in Section 3.2, a brief description of the optimality conditions developed by Euler and Lagrange more than three hundreds years ago. In Section 3.3, we give a precise statement of one version of the Maximum Principle, one for continuous-time systems with smooth dynamics in infinite-dimensional spaces.

3.2 Euler-Lagrange Equations

For simplicity and to focus on the methodology instead of the technicalities arising when working in infinite dimensions, we will consider the following version of the optimal control problem (OCP) with $X = \mathbb{R}^n, U = \mathbb{R}^m$.

$$\text{minimize} \quad \mathcal{J}(u) := \int_a^b l(x(t), u(t), t) dt \tag{3.6}$$

$$\text{subject to} \quad \dot{x}(t) = f(x(t), u(t), t); \qquad x(a) = x_0 \tag{3.7}$$

with a fixed initial time a and terminal time b. The difference between (OCP) and (3.6, 3.4) is that there is no restrictions on the control variables (i.e. the control set U is the whole $\mathbb{R}^m$).

Theorem 3.6 (Euler-Lagrange Conditions)

Consider the problem (3.6)-(3.7) for $u \in C[a,b]$, with fixed endpoints $a < b$, where l and f are continuous in (x,u,t) and have continuous first partial derivatives with respect to x and u for all $(x,u,t) \in \mathbb{R}^n \times \mathbb{R}^m \times [a,b]$. Suppose that u^* is a minimizer for the problem, and let $x^* \in C^1[a,b]$ denote the corresponding response. Then, there is a vector function $p^* \in C^1[a,b]$ such that the triple (u^*, x^*, p^*) satisfies the system

$$\dot{x}(t) = f(x(t), u(t), t); \qquad x(a) = x_0 \tag{3.8}$$

$$\dot{p}(t) = -l_x(x(t), u(t), t) - f_x(x(t), u(t), t)^\top p(t); \qquad p(b) = 0 \tag{3.9}$$

$$0 = l_u(x(t), u(t), t) + f_u(x(t), u(t), t)^\top p(t). \tag{3.10}$$

for $a \leq t \leq b$. These equations are known collectively as the Euler-Lagrange equations, and 3.9 is often referred to as the adjoint equation (or the costate equation).

Before we consider any examples, let's discuss the following remarks

- The above conditions consist of m algebraic equations (3.10), together with $2 \times n$ ODEs (3.8,3.9) and their respective boundary conditions. These boundary conditions are split, i.e., some are given at $t = a$ and others at $t = b$. Such problems, known as two-point boundary value problems, are more difficult to solve than initial-value problems.

- If $f(x(t), u(t), t) = u(t)$ with $n = m$, then (3.10) gives

$$\bar{p}(t) = -l_u(\bar{x}(t), \bar{u}(t), t)$$

and from (3.9) we have the Euler equation

$$\frac{d}{dt} l_u(\bar{x}(t), \dot{\bar{x}}, t) = l_x(\bar{x}(t), \dot{\bar{x}}, t),$$

together with the boundary conditions

$$[l_u(\bar{x}(t), \dot{\bar{x}}, t)]_{t=b}.$$

This shows that Euler-Lagrange equations include the optimality necessary conditions derived for problems of the calculus of variations.

- It is convenient to introduce the Hamiltonian function $\mathbb{H} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}$ associated with the optimal control problem (3.6,3.7) as

$$\mathbb{H}(x, p, u, t) = l(x, u, t) + p^{\mathsf{T}} f(x, u, t). \tag{3.11}$$

Therefore, Euler-Lagrange equations (3.8-3.10) can be rewritten as

$$\dot{x}(t) = \mathbb{H}_p; \qquad x(a) = x_0 \tag{3.12}$$

$$\dot{p}(t) = -\mathbb{H}_x; \qquad p(b) = 0 \tag{3.13}$$

$$0 = \mathbb{H}_u, \tag{3.14}$$

for $t \in [a, b]$. Note that a necessary condition for the triple $(\bar{u}, \bar{x}, \bar{p})$ to be a local minimum of $\mathcal{J}$ is that $\bar{u}(t)$ be a stationary point of the Hamiltonian function with $\bar{x}(t)$ and $\bar{p}(t)$ at each $t \in [a, b]$. In some cases, one can express $u(t)$ in terms of $x(t)$ and $p(t)$ from (3.14), and then substitute into (3.12,3.13) to get a two-point boundary value problem in the variables x and p.

Example 3.7

Consider the optimal control problem

$$\text{minimize} \quad \mathcal{J}(u) := \int_0^1 \left[\frac{1}{2}u^2(t) - x(t)\right] dt \tag{3.15}$$

$$\text{subject to} \quad \dot{x}(t) = 2\left[1 - u(t)\right]; \qquad x(0) = 1. \tag{3.16}$$

The Hamiltonian function for this problem

$$\mathbb{H}(x, p, u, t) = \frac{1}{2}u^2 - x(t) + 2p(t)(1 - u).$$

Any candidate solution $(\bar{u}, \bar{x}, \bar{p})$ to this problem must satisfy the Euler-Lagrange equations. That is

$$\begin{aligned} \dot{x}(t) =& \mathbb{H}_p = 2\left[1 - \bar{u}(t)\right]; & x(0) =& 1 \\ \dot{p}(t) =& -\mathbb{H}_x = 1; & \bar{p}(1) =& 0 \\ 0 =& \mathbb{H}_u = \bar{u}(t) - 2\bar{p}(t). \end{aligned}$$

The adjoint equation gives

$$\bar{p}(t) = t - 1,$$

and from the last condition $\mathbb{H}_u = 0$ we have

$$\bar{u}(t) = 2(t - 1).$$

Finally, substituting $\bar{u}$ into (3.16) gives

$$\dot{\bar{x}}(t) = 6 - 4t; \quad \bar{x}(0) = 1.$$

By integrating, we get

$$\bar{x}(t) = -2t^2 + 6t + 1.$$

It is worth noting that $\mathbb{H}$ is constant along $(\bar{u}, \bar{x}, \bar{p})$. Indeed, we have

$$\mathbb{H}(\bar{x}(t), \bar{p}(t), \bar{u}(t), t) = -5.$$ □

We conclude this section by giving a brief account of optimality sufficient conditions, called Mangasarian Sufficient Conditions, for the problem (3.6,3.7).

Theorem 3.8 (Mangasarian Sufficient Conditions)

Consider the problem (3.6)-(3.7) for $u \in C[a,b]$, with fixed endpoints $a < b$, where l and f are continuous in (x,u,t) and have continuous first partial derivatives with respect to x and u , and are convex in x and u for all $(x,u,t) \in \mathbb{R}^n \times \mathbb{R}^m \times [a,b]$. Suppose that $(u^*, \bar{x}, \bar{p})$ satisfies the Euler-Lagrange equations (3.8-3.10). Suppose also that

$$\bar{p}(t) \geq 0, \qquad \text{for } a \leq t \leq b. \tag{3.17}$$

Then $\bar{u}$ is a global minimizer for the problem (3.6,3.7).

Remark 3.9

In the case where f is linear in (x,u), the result holds without any sign restriction on $\bar{p}$, i.e. without (3.17).

Example 3.10

In Example 3.7, the integrand is convex in (u,x) on $\mathbb{R}^2$, and the right-hand side of (3.16) is linear in u and independent of x. Moreover the candidate solution

$$\bar{u}(t) = 2(t-1), \quad \bar{x}(t) = -2t^2 + 6t + 1, \quad \bar{p}(t) = t - 1$$ □

satisfies the Euler-Lagrange equations (3.8-3.10) for each $t \in [0,1]$. So $\bar{u}(t)$ is a global minimizer for the problem irrespective of the sign condition (3.17) due to the linearity of (3.16) (see Remark

3.9).

For more on sufficient conditions in optimal control theory, we refer the reader to [38] and [26] and the references therein.

3.3 Pontryagin Maximum Principle

Our goal in this subsection is to derive necessary optimality conditions in the form of the Pontryagin maximum principle for the problem (OCP) where the governing dynamic system is an ordinary differential equation in infinite-dimensional spaces that explicitly involve constrained control inputs $u(\cdot)$ as follows:

$$\dot{x} = f(x, u, t),\ u(t) \in U \ \text{ a.e. } t \in [a, b]. \tag{3.18}$$

The system (3.18) is of smooth dynamics, which means that f is continuously differentiable with respect to the state variable x around an optimal solution to be considered. Despite this assumption, the control system (3.18) and optimization problems over its feasible controls and trajectories essentially involve non-smoothness due to the control geometric constraints $u(t) \in U$ a.e. $t \in [a, b]$ defined by control sets U of a general nature. For instance, it is the case with the simplest/classical optimal control problems with $U = \{0, 1\}$.

Now given an optimal solution $(\bar{u}(\cdot), \bar{x}(\cdot))$ to (OCP), we assume the following to be true throughout this subsection.

(A1) the state space X is Banach;

(A2) the control set U is a Souslin subset (i.e., a continuous image of a Borel subset) in a complete and separable metric space;

(A3) there is an open set $O \subset X$ containing $\bar{x}(t)$ such that f is Frèchet differentiable in x with both $f(x,u,t)$ and $\nabla_x f(x,u,t)$ continuous in (x,u), measurable in t, and norm-bounded by a summable function for all $x \in O, u \in U,$ and a.e. $t \in [a,b]$;

(A4) the function ϕ_0 is Frèchet differentiable at $\bar{x}(b)$.

Note that the control set U may depend on t in a general measurable way, which allows one to use standard measurable selection results; see, e.g., the book [3] with the references therein.

The Hamilton-Pontryagin function for (3.18) is defined as

$$H(x,p,u,t) := \langle p, f(x,u,t)\rangle \text{ with } p \in X^*.$$

We now give the following version of the maximal principle due to [31, p. 238].

Theorem 3.11 (maximum principle for smooth control systems)

Let $(\bar{u}(\cdot), \bar{x}(\cdot))$ be an optimal solution to problem (OCP) under the assumptions (A1)-(A4). Then the following maximum conditions holds:

$$H(\bar{x}(t), p(t), \bar{u}(t), t) = \max_{u \in U} H(\bar{x}(t), p(t), u, t) \text{ a.e. } t \in [a,b], \tag{3.19}$$

where an absolutely continuous mapping $p : [a,b] \to X^*$ is a trajectory for the adjoint system

$$\dot{p} = -\nabla_x H(\bar{x}, p, \bar{u}, t) \text{ a.e. } t \in [a,b] \tag{3.20}$$

with the transversality condition

$$p(b) = -\nabla \phi_0(\bar{x}(b)). \tag{3.21}$$

A solution (adjoint arc) to system (3.20) is understood in the integral sense similarly to (3.2), i.e.,

$$p(t) = p(b) + \int_t^b \nabla_x H(\bar{x}(s), p(s), \bar{u}(t), s) ds, \quad t \in [a,b],$$

■

with $\nabla_x H(\bar{x}, p, \bar{u} \cdot t) = \langle p, \nabla_x f(\bar{x}, \bar{u}, t) \rangle$.

Proof.

Let $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ be an optimal solution to problem (OCP), and let $p(\cdot)$ be the corresponding solution to the adjoint system (3.20) with the boundary/transversality condition (3.21). We are going to show that the maximum condition (3.19) holds for a.e.$t \in [a, b]$. Assume on the contrary that there is a set $T \subset [a, b]$ of positive measure such that

$$H(\bar{x}(t), p(t), \bar{u}(t), t) < \sup_{u \in U} H(\bar{x}(t), p(t), u, t) \text{ for } t \in T.$$

Then using standard results on measurable selections under the assumptions made, we find a measurable mapping $v : T \to U$ satisfying

$$\triangle_v H(t) := H(\bar{x}(t), p(t), v(t), t) - H(\bar{x}(t), p(t), \bar{u}(t), t) > 0, \quad t \in T.$$

Let $T_0 \subset [a, b]$ be a set of Lebesgue regular points (or points of approximate continuity) for the function $H(t)$ on the interval $[a, b]$, which is of full measure on $[a, b]$ due to the classical Denjoy theorem. Given $\tau \in T_0$ and $\varepsilon > 0$, consider a needle variation of the optimal control built by

$$u(t) := \begin{cases} v(t), & t \in T_\varepsilon := [\tau, \tau + \varepsilon) \cap T_0, \\ \bar{u}(t), & t \in [a, b] \backslash T_\varepsilon. \end{cases}$$

Now let $x(\cdot)$ be the corresponding solution to $u(\cdot)$ in the sense of (3.2) and denote

$$\Delta \bar{u}(t) := u(t) - \bar{u}(t), \quad \Delta \bar{x}(t) := x(t) - \bar{x}(t), \quad \Delta \mathcal{J}[\bar{u}] := \phi_0(x(b)) - \phi_0(\bar{x}(b)).$$

The perturbed control $u(\cdot)$ differs from the $\bar{u}(\cdot)$ only on the small time set T_ε, where $u(t) \in U$ a.e.; the name "needle variation" comes from this.

Since ϕ_0 is assumed to be Frechet differentiable at $\bar{x}(b)$, we have the representation

$$\Delta\mathcal{J}[\bar{u}] = \phi_0(x(b)) - \phi_0(\bar{x}(b)) = \langle\nabla\phi_0(\bar{x}(b)), \Delta\bar{x}(b)\rangle + o(||\Delta\bar{x}(b)||).$$

Using integration by parts which holds for Bochner integrals, one gets

$$\begin{aligned}
\textstyle\int_a^b\langle p(t), \triangle\dot{\bar{x}}(t)\rangle dt &= \langle p(t), \triangle\bar{x}(t)\rangle|_b^a - \textstyle\int_a^b\langle\dot{p}(t), \triangle\bar{x}(t)\rangle dt, \\
&= \langle p(b), \triangle\bar{x}(b)\rangle - \langle p(a), \triangle\bar{x}(a)\rangle - \textstyle\int_a^b\langle\dot{p}(t), \triangle\bar{x}(t)\rangle dt.
\end{aligned}$$

Since $\Delta\bar{x}(a) = 0$, we have the following identity

$$\langle p(b), \triangle\bar{x}(b)\rangle = \int_a^b\langle\dot{p}(t), \triangle\bar{x}(t)\rangle dt + \int_a^b\langle p(t), \triangle\dot{\bar{x}}(t)\rangle dt.$$

Because of (3.21), we arrive at

$$\triangle\mathcal{J}[\bar{u}] = -\int_a^b\langle\dot{p}(t), \triangle\bar{x}(t)\rangle dt - \int_a^b\langle p(t), \triangle\dot{\bar{x}}(t)\rangle dt + o(||\triangle\bar{x}(b)||).$$

Let us transform the second integral above. Using the equation

$$\triangle\dot{\bar{x}} = f(\bar{x}(t) + \triangle\bar{x}(t), \bar{u}(t) + \triangle\bar{u}(t), t) - f(\bar{x}(t), \bar{u}(t), t),$$

the definition of the Hamilton-Pontryagin function $H(x, p, u, t)$, and (A3), we have

$$\begin{aligned}
\textstyle\int_a^b\langle p(t), \triangle\dot{\bar{x}}(t)\rangle dt &= \textstyle\int_a^b\left[H(\bar{x}(t) + \triangle\bar{x}(t), p(t), \bar{u}(t) + \triangle\bar{u}(t), t) - H(\bar{x}(t), p(t), \bar{u}(t), t)\right]dt \\
&= \textstyle\int_a^b\left[H(\bar{x}(t), p(t), \bar{u}(t) + \triangle\bar{u}(t), t) - H(\bar{x}(t), p(t), \bar{u}(t), t)\right]dt \\
&\quad + \textstyle\int_a^b\left\langle\frac{\partial H(\bar{x}(t), p(t), \bar{u}(t), t)}{\partial x}, \triangle\bar{x}(t)\right\rangle dt + \textstyle\int_a^b o(||\triangle\bar{x}(t)||)dt.
\end{aligned}$$

Now Letting

$$\triangle_u H(\bar{x}(t), p(t), \bar{u}(t), t) := H(\bar{x}(t), p(t), u(t), t) - H(\bar{x}(t), p(t), \bar{u}(t), t),$$

we come to the following increment formula

$$\begin{aligned}\triangle\mathcal{J}[\bar{u}] &= -\int_a^b \triangle_u H(\bar{x}(t), p(t), \bar{u}(t), t)dt - \int_a^b \left\langle \frac{\partial\triangle_u H(\bar{x}(t), p(t), \bar{u}(t), t)}{\partial x}, \triangle\bar{x}(t)\right\rangle dt \\ &\quad - \int_a^b o(||\Delta\bar{x}(t)||)dt + o(||\Delta\bar{x}(b)||).\end{aligned}$$

Let's assume, for the time being, that there exists a constant $K > 0$ independent of (τ, ε) such that

$$\|\triangle\bar{x}(t)\| \leq K\varepsilon \quad \text{for all} \quad t \in I. \tag{3.22}$$

Then we have

$$o(||\triangle\bar{x}(b)||) = o(\varepsilon), \quad \int_a^b o(||\triangle\bar{x}(t)||)dt = o(\varepsilon), \text{ and}$$

$$\begin{aligned}-\int_a^b \left\langle \frac{\partial\triangle_u H(\bar{x}(t), p(t), \bar{u}(t), t)}{\partial x}, \triangle\bar{x}(t)\right\rangle dt &\leq \int_\tau^{\tau+\varepsilon} \left|\left\langle \frac{\partial\triangle H_v(\bar{x}(t), p(t), \bar{u}(t), t}{\partial x}, \triangle\bar{x}(t)\right\rangle\right| dt \\ &\leq K\varepsilon \int_\tau^{\tau+\varepsilon} \left\|\frac{\partial\triangle H_v(\bar{x}(t), p(t), \bar{u}(t), t}{\partial x}\right\| dt = o(\varepsilon),\end{aligned}$$

The choice of $\tau \in T_0$ as a Lebesgue regular point of the function $\triangle_v H(t)$ and the construction of the Bochner integral yield

$$\int_\tau^{\tau+\varepsilon} \triangle_v H(t)dt = \varepsilon\left[H(\bar{x}(\tau), p(\tau), v(\tau), \tau) - H(\bar{x}(\tau), p(\tau), \bar{u}(\tau), \tau)\right] + o(\varepsilon).$$

Thus we get the representation

$$\triangle\mathcal{J}[\bar{u}] = -\varepsilon\left[H(\bar{x}(\tau), p(\tau), v(\tau), \tau) - H(\bar{x}(\tau), p(\tau), \bar{u}(\tau), \tau)\right] + o(\varepsilon),$$

which implies that $\triangle\mathcal{J}[\bar{u}] < 0$ along the above needle variation of the optimal control $\bar{u}(\cdot)$ for all $\varepsilon >$

0 sufficiently small. This clearly contradicts the optimality of $\bar{u}(\cdot)$.

To complete the proof we have to show that (3.22) is valid. To do so, we notice first that for the trajectory increment $\triangle\bar{x}(t)$ we have

$$\triangle\bar{x}(t) = 0 \text{ for all } t \in [a, \tau].$$

Denote by ℓ the uniform Lipschitz constant for $f(\cdot, v(t), t)$ whose existence is guaranteed by (A3). For simplicity we suppose that ℓ is independent of t although the assumptions made allow it to be summable on $[a, b]$ with no change of the result. Since $\triangle\bar{x}(\tau) = 0$, and by (3.2) we have

$$\triangle\bar{x}(t) = \int_\tau^t [f(\bar{x}(s) + \triangle\bar{x}(s), v, s) - f(\bar{x}(s), \bar{u}(s), s)]\, ds, \quad \tau \le t \le \tau + \varepsilon.$$

Denoting

$$\triangle_v f(\bar{x}(s), \bar{u}(s), s) := f(\bar{x}(s), v, s) - f(\bar{x}, \bar{u}(s), s),$$

we have

$$\begin{aligned} ||\triangle\bar{x}(t)|| &= \textstyle\int_\tau^t ||f(\bar{x}(s) + \triangle\bar{x}(s), v, s) - f(\bar{x}(s), \bar{u}(s), s)||ds \\ &\le \textstyle\int_\tau^t ||\triangle_v f(\bar{x}(s), \bar{u}(s), s)||ds + \ell \int_\tau^t ||\triangle\bar{x}(s)||ds. \end{aligned}$$

Using the notation

$$\alpha(t) := \int_\tau^t ||\triangle_v f(\bar{x}(s), \bar{u}(s), s)||ds \qquad \text{and} \qquad \beta(t) := ||\triangle\bar{x}(t)||,$$

the above estimate can be written as

$$\beta(t) \le \alpha(t) + \ell \int_\tau^t \beta(s)ds, \quad \tau \le t \le \tau + \varepsilon,$$

which yields by the classical Gronwall lemma that

$$||\triangle \bar{x}(t)|| \leq \left(\int_{\tau}^{t} ||\triangle_v f(\bar{x}(s), \bar{u}(s), s)||ds\right) e^{\text{l}(t-\tau)} \leq K\varepsilon$$

for $t \in [\tau, \tau+\varepsilon],$ where$K = K(v)$ is independent of ε and $\tau.$ It remains to estimate $\triangle \bar{x}(t)$ on the last interval $[\tau+\varepsilon, b]$, where it satisfies the equation

$$\triangle \dot{\bar{x}}(t) = f(\bar{x}(t) + \triangle \bar{x}(t), \bar{u}(t), t - f(\bar{x}(t), \bar{u}(t), t \quad \text{with} \quad ||\triangle \bar{x}(\tau+\varepsilon)|| \leq K\varepsilon$$

the solution of which is understood in the integral sense (3.2). Since

$$\begin{aligned} ||\triangle \bar{x}(t)|| &\leq ||\triangle \bar{x}(\tau+\varepsilon)|| \int_{\tau+\varepsilon}^{t} ||f(\bar{x}(s) + \triangle \bar{x}(s), \bar{u}(s), s) - f(\bar{x}(s), \bar{u}(s), s||ds \\ &\leq K\varepsilon + \text{l} + \int_{\tau+\varepsilon}^{t} ||\triangle \bar{x}(s)||ds, \quad \tau+\varepsilon \leq t \leq b, \end{aligned}$$

we again apply the Gronwall lemma and arrive, by increasing K if necessary at the desired estimate of $||\triangle \bar{x}(t)||$ on the whole interval $[a, b]$.

∎

Example 3.12

Consider the following problem

$$\text{minimize} \quad \int_0^1 x_1(t)dt \tag{3.23}$$

$$\text{subject to} \quad \dot{x_1}(t) = u(t), \qquad x_1(0) = 1, \tag{3.24}$$

$$u(t) \in [-1, 1]. \tag{3.25}$$

First, we write the problem in Mayer's form by introducing an additional state variable x_2 which satisfies the equation

$$\dot{x}_2(t) = x_1(t), \qquad x_2(0) = 0.$$

The problem (3.23-3.25) now can be cast as

$$\text{minimize} \quad J[u] = x_2(1) \tag{3.26}$$

$$\text{subject to} \quad \dot{x}_1(t) = u(t), x_1(0) = 1, \quad \dot{x}_2(t) = x_1(t), x_2(0) = 0, \tag{3.27}$$

$$u(t) \in [-1, 1]. \tag{3.28}$$

Note that $X = \mathbb{R}^2, U = [-1,1], I = [0,1], x \equiv (x_1 \quad x_2)^T \in \mathbb{R}^2, f \equiv (u \quad x_2)^T \in \mathbb{R}^2$ and $\phi_0(x(t)) = x_2(t)$. The Hamilton-Pontryagin function for this problem is

$$H(x,p,u,t) = p^T \cdot f, \qquad p = (p_1 \quad p_2)^T \in \mathbb{R}^2;$$

That is

$$H(x,p,u,t) = p_1 u + p_2 x_1.$$

This is a linear function in u, and therefore the control $\bar{u}$ that maximizes H is

$$\bar{u}(t) = \begin{cases} 1, & \text{if} \quad \bar{p}_1(t) > 1, \\ -1, & \text{if} \quad \bar{p}_1(t) < 1, \\ \text{undefined}, & \text{if} \quad \bar{p}_1(t) = 0. \end{cases}$$

According to (3.20) and (3.21), p satisfies

$$\begin{aligned} \dot{\bar{p}}_1(t) &= -\bar{p}_2(t), \qquad \bar{p}_1(1) = 0, \\ \dot{\bar{p}}_2(t) &= 0, \qquad \bar{p}_2(1) = -1, \end{aligned}$$

and therefore

$$\left.\begin{aligned} \bar{p}_1(t) &= t-1, \\ \bar{p}_2(t) &= -1, \end{aligned}\right\} \quad \text{for all} \quad t \in [0,1].$$

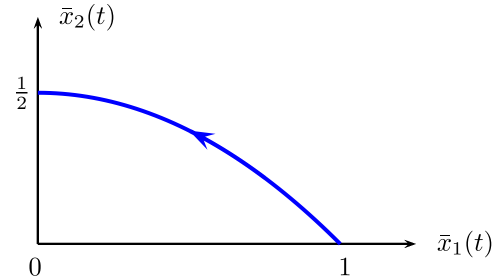

Figure. 3.1: Potential Optimal State Trajectory for Example 3.12.

But $p_1(t) \leq 0$ for all $t \in [0, 1]$, which dictates that $\bar{u}(t) = -1$. This gives through (3.27)

$$\left.\begin{array}{rcl} \bar{x}_1(t) & = & 1 - t, \\ \bar{x}_2(t) & = & -\frac{1}{2}t^2 + t, \end{array}\right\} \text{ for all } \quad t \in [0, 1].$$

The control $\bar{u}$, the response trajectory $(\bar{x}_1, \bar{x}_2)$ and the adjoint arc $(\bar{p}_1, \bar{p}_2)$ constitute a candidate for an optimal solution to Example 3.12 with optimal value to the cost function $\mathcal{J}[\bar{u}] = 1/2$. Figure 3.1 shows $\bar{x}_1$ and $\bar{x}_2$ in the x_1x_2-plane. In Figures 3.2 and 3.3, the graphs of potential optimal state and adjoint trajectories $\bar{x}_1, \bar{p}_1$ and $\bar{x}_2, \bar{p}_2$. □

3.4 A Historical Note

Optimal control had its origins in the calculus of variations in the 17$^{\text{th}}$century (Fermat, Newton, Leibnitz, and the Bernoulis). Johann Bernoulli in 1696 challenged the mathematicians of his era to solve the brachistochrone problem. Five mathematicians responded to the challenge: Leibnitz, l'Hospital, Tschirnhaus, Newton and Johann's brother Jakob Bernoulli. In 1697, Bernoulli published all the solutions. The calculus of variations was developed further in the 18$^{\text{th}}$century by Euler and Lagrange and in the 19$^{\text{th}}$century by Legendre, Jacobi, Hamilton, and Weierstrass. In the early 20$^{\text{th}}$century, Bolza and Bliss put the final touches of rigor on the subject. In 1957, Bellman gave a new view of Hamilton-Jacobi theory which he called dynamic programming. Mc-

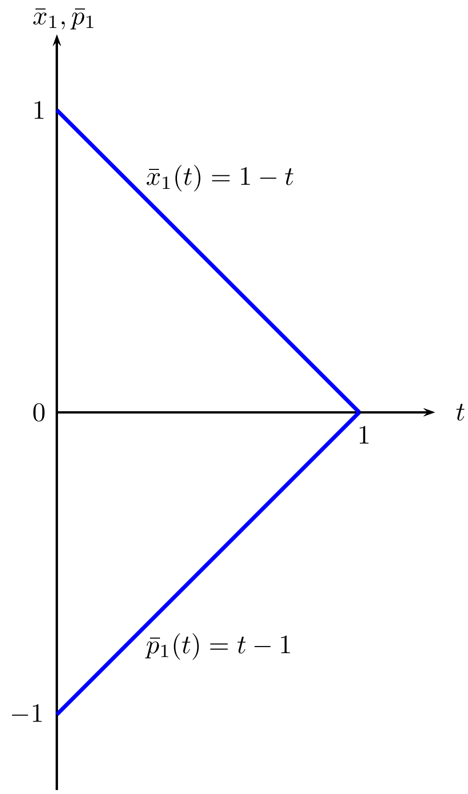

Figure. 3.2: Potential Optimal State and Adjoint Trajectories ($\bar{x}_1$ and $\bar{p}_1$) for Example 3.12.

shane (1939) and Pontryagin(1962) extended the calculus of variations to handle control variable inequality constraints, the latter announcing his elegant maximum principle [35]. The truly enabling element for use of optimal control theory was the digital computer, which became available commercially in the 1950's. In the late 1950's and early 1960's Lawden, Leitmann, Miele, and Breakwell demonstrated possible uses of the calculus of variations in optimizing aerospace flight paths using shooting algorithms, while kelley and Bryson developed gradient algorithms that eliminated the inherent instability of shooting methods. Also in the early 1960's Simon, Chang, Kalman, Bucy, Battin, Athans, and many others showed how to apply the calculus of variations to design optimal output feedback logic for linear dynamic systems in the presence of noise using digital control. Clarke [6, 7], Vinter [25, 42] and Mordukhovich [30, 31] studied more general forms of the optimal control problem with a relaxation of the differentiability conditions

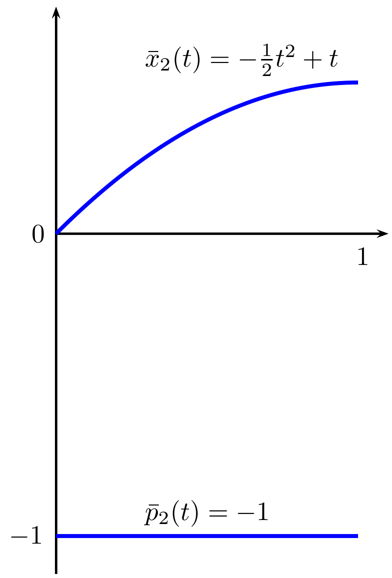

Figure. 3.3: Potential Optimal State and Adjoint Trajectories ($\bar{x}_2$ and $\bar{p}_2$) for Example 3.12.

necessary in the classical results. For more on the history of optimal control, we refer the reader to [5, 41, 37].

The Pontryagin maximum principle is the central result of optimal control theory. In the half-century since its appearance, the underlying theorem has been generalized, strengthened, extended, reproved and interpreted in a variety of ways. Clarke in [8] discusses the evolution of the Pontryagin maximum principle, focusing primarily on the hypotheses required for its validity and giving necessary conditions for optimal control problems formulated in terms of differential inclusions. More recently Clarke [9] reviews one of the principal approaches to obtaining the maximum principle in a powerful and unified context, focusing upon recent results that represent the culmination of over thirty years of progress using the methodology of nonsmooth analysis. A short history of the discovery of the maximum principle in optimal control theory by Pontryagin and his associates is presented by Gamkrelidze in [20]. The reader, with further interest in Pontryagin maximum principle, can visit the well-designed course in [24].

CHAPTER 4

OPTIMAL CONTROL OF SINGULAR DIFFERENTIAL OPERATORS IN HILBERT SPACES

In this chapter we formulate, for the first time in the literature, an optimal control problem for self-adjoint ordinary differential operator equations in Hilbert spaces and derive necessary conditions for optimal controls to this problem in an appropriate extended form of the Pontryagin Maximum Principle.

Section 4.1 is an introductory one where the problem under study is presented. In Section 4.2 we give a brief introduction to the theory of self-adjoint differential operator equations, highlighting the main landmarks that show remarkable features these systems have, which are largely used in what follows. This is based is the seminal work by Akhiezer and Glazman [1], Naimark [33], Weidmann [45], and Zettl [46], [47] among others.

In Section 4.3 we obtain new existence results for self-adjoint differential operator equations, which play a crucial role in the proof of the Maximum Principle of Theorem 4.1 given in Section 4.4.

4.1 Introduction

This chapter addresses the following controlled system governed by singular differential operator equations in Hilbert spaces:

$$Lx = f(x,u,t), \qquad u(t) \in U \ \text{ a.e. } \ t \in I = (a,b), \quad -\infty \leq a < b \leq \infty, \tag{4.1}$$

where L is a self-adjoint extension of the minimal operator L_0 (see Section 4.2) generated by a formally self-adjoint differential expression l_Q and a positive weight function w satisfying the equation

$$l_Q x = \lambda w x \qquad \text{on} \quad I \tag{4.2}$$

in the Hilbert space $\mathcal{H} = L^2(I,w)$ of real-valued square integrable functions, where $u(\cdot)$ is a measurable control action taking values from the given control set U, and where the function f is complex-valued. The inner product $\langle \cdot, \cdot \rangle$ and the norm $\| \cdot \|$ on $\mathcal{H}$ are defined, respectively, by

$$\langle x_1, x_2 \rangle := \int_I x_1(t)x_2(t)w(t)dt,$$

$$\|x\|^2 := \int_I |x(t)|^2 w(t)dt.$$

In what follows we assume that the expression $l := l_Q$ in (4.2) is of even order $2n$, with $Q \in Z_{2n}(I), Q = Q^+$ and Q is real, see Definition 2.2. Recall (cf. [33]) that l is given in the form

$$l(x) = \sum_{i=0}^{n} (-1)^i (r_i x^{(i)})^{(i)}$$

with real-valued coefficients $r_i \in C^i[I]$ for all $i = 0, \dots, n$. Recall that the expression l is regular if the I is finite and

$$r_n^{-1}, r_{n-1}, \dots, r_0 \in L(I, w),$$

i.e., these functions are integrable on the whole interval I. Otherwise l is called singular. Furthermore, the endpoint a is regular if $a > -\infty$ and if $r_n^{-1}, r_{n-1}, \dots, r_0 \in L((a, \beta), w)$ for all $\beta < b$; otherwise a is singular. The regularity and singularity of the other endpoint b is defined similarly. Observe that the expression l is regular if and only if both endpoints a and b have this property.

We now fix a point c such that $a < c < b$ and consider the following optimal control problem of the Mayer type for controlled equation (4.1):

$$\text{minimize} \quad J[u, x] = \phi(x(c)) \text{ over } (u, x) \in \mathcal{A}. \tag{4.3}$$

Here the cost function ϕ is real-valued and the set $\mathcal{A}$ is the collections of admissible pairs $(u(\cdot), x(\cdot))$ with measurable controls $u(\cdot)$ satisfying the pointwise constraint $u(t) \in U$ a.e. $t \in I$ and the corresponding solutions $x(\cdot)$ to (4.1) described by

$$x(t) = \int_I K(t, \tau) f(x(\tau), u(\tau), \tau) d\tau, \quad t \in I \tag{4.4}$$

see Section 4.2 for more details. If b is regular, we may take $c = b$. Although any state variable x must satisfy boundary conditions; being an element of D; see Section 4.2, particularly Theorem 4.2). Since no additional constraints are imposed on $x(\cdot)$ at $t = b$, problem (4.3) is labeled a free-endpoint problem of optimal control. Any admissible pair $(u,x)\in\mathcal{A}$ are called feasible solution to the control problem (4.3). A feasible solution $(\bar{u},\bar{x})$ is (globally) optimal for this problem if

$$J[\bar{u},\bar{x}]\le J[u,x]\quad \text{whenever } (u,x)\in\mathcal{A}.$$

Optimal control theory is a remarkable area of Applied Mathematics, which has been developed for various classes of controlled systems governed by ordinary differential, functional differential, and partial differential equations and inclusions; see, e.g., [23, 31] with the vast bibliographies therein. However, we are not familiar with any developments on optimal control of differential operator equations of type (4.1).

To proceed further, take an arbitrary admissible control $u(\cdot)$ and define the operator F_u on $\mathcal{H}$ by

$$F_u(x):=f(x(\cdot),u(\cdot),\cdot)\quad \text{on}\quad I. \tag{4.5}$$

The main goal of this chapter is deriving necessary optimality conditions for a fixed optimal solution $(\bar{u}(\cdot),\bar{x}(\cdot))$ to problem (4.3). Involving this optimal pair and operator (4.5), we impose the following standing assumptions:

(H1) F_u maps $\mathcal{H}$ into $\mathcal{H}$ and there exists an open set $O\subset\mathcal{H}$ containing $\bar{x}$ such that the functions $(x,u)\mapsto F_u(x)$ and $(x,u)\mapsto F'_u(x)$ are continuous on $\mathcal{A}$ and the operators $F'_u(\bar{x})$ are uniformly bounded for all admissible controls u.

(H2) For each admissible control u the operator F_u is weakly continuous.

(H3) For each admissible control u the operator F_u is monotone, i.e.,

$$\langle F_u(x_1)-F_u(x_2),x_1-x_2\rangle\le\eta\|x_1-x_2\|^2,\qquad \text{for all } x_1,x_2\in H,$$

where $\eta\in\mathbb{R}$ independent of u.

(H4) There exists a real number $\gamma>\eta$, assumed to be positive without loss of generality, such that

$$\langle Lx, x\rangle \geq \gamma\|x\|^2 \quad \text{for any } x \in D,$$

where D is the domain of L to be defined in Section 4.2.

(H5) For every needle variation u (see Section 4.4) of $\bar{u}$ on measurable sets $I_\epsilon \subset I$ of measure ϵ we have

$$\|F_u(\bar{x}) - F_{\bar{u}}(\bar{x})\| = o(\epsilon).$$

(H6) The function ϕ is Fréchet differentiable at the point $\bar{x}(c)$.

(H7) The control set U in (4.1) is a Souslin subset (i.e., a continuous image of a Borel subset) of some Banach space.

To formulate the main result, we introduce the appropriate counterpart of the Hamilton-Pontryagin function for system (4.1) defined by

$$H(x, p, u, t) := (p + P(\phi(x(c))K(c, t)))\, f(x, u, t), \quad p \in D_1 \tag{4.6}$$

where P is a projection operator onto the range of L_0 to be discussed in Section 4.2; see particularly Lemma 4.4 therein.

Theorem 4.1 (Maximum Principle)

Let $(\bar{u}(\cdot), \bar{x}(\cdot))$ be an optimal solution to problem (4.3) under the assumptions imposed in (H1)–(H7). Then there exists an adjoint arc $p \in D_1$ such that

$$H(\bar{x}(t), p(t), \bar{u}(t), t) = \max_{u \in U} H(\bar{x}(t), p(t), u, t) \text{ a.e. } t \in I, \tag{4.7}$$

$$L_1(p) = -\nabla_x H(\bar{x}(t), p(t), \bar{u}(t), t) \text{ a.e.} \tag{4.8}$$

and the following transversality condition is satisfied:

$$[p, x_i]_a^b = -\phi_0'(\bar{x}(c))x_i(c), \qquad i = 1, \ldots, d, \tag{4.9}$$

where D_1 is the domain of the operator L_1, and where the functions $x_i, i = 1, \ldots, d$ determine the domain D in the sense of Theorem 4.2. ■

4.2 Self-Adjoint Differential Operator Equations

The expression l in (4.2) generates various operators on $\mathcal{H}$. Among these operators we single out the minimal operator L_0, the maximal operator L_1, and self-adjoint operators L lying between. The maximal operator L_1 is defined by

$$\begin{aligned} D_1 = D(L_1) : &= \{x \in \mathcal{H} : x^{[0]}, x^{[1]}, \ldots, x^{[2n-1]} \in AC_{\text{loc}}(I) \text{ and } x^{[2n]} \in \mathcal{H}\}, \\ L_1(x) : &= l(x), \quad x \in D_1, \end{aligned}$$

where $x^{[i]}$ is the i$^{\text{th}}$quasi-derivative related to l and given by

$$\begin{aligned} x^{[i]} : &= \frac{d^i x}{dt^i}, \qquad i = 0, \ldots, n-1, \\ x^{[n]} : &= r_n \frac{d^n x}{dt^n}, \\ x^{[n+i]} : &= r_{n-i} \frac{d^{n-i} x}{dt^{n-i}} - \frac{d}{dt}\left(x^{[n+i-1]}\right), \qquad i = 1, \ldots, n. \end{aligned}$$

Denote by $AC_{\text{loc}}(I)$ the set of real-valued functions, which are absolutely continuous on every compact subinterval of I. Let $L_0 := L_1^*$ with $D_0 := D(L_0)$, where L_1^* is the adjoint of L_1 uniquely defined due to the fact that D_1 is dense in $\mathcal{H}$. It is shown in [33] that $D_0 \subset D_1$, that D_0 is dense in $\mathcal{H}$, and that $L_0^* = L_1$, which implies in turn that L_0 is a symmetric closed operator.

Pick an arbitrary complex number ν with Im$(\nu) \neq 0$ and denote the range of $(L_0 - \nu E)$ by $\mathcal{R}_\nu$, where E is the identity operator on $\mathcal{H}$. The orthogonal complement of cl R_ν in $\mathcal{H}$ is called the deficiency space of L_0 corresponding to ν and is denoted by $\mathcal{N}_\nu$. It is shown in [33] that $\mathcal{N}_\nu$ is the eigenspace of L_1 corresponding to the eigenvalue $\bar{\nu}$ and that D_1 is decomposed as

$$D_1 = D_0 \dot{+} \mathcal{N}_\nu \dot{+} \mathcal{N}_{\bar{\nu}}. \tag{4.10}$$

It is also shown in [33] that the equality

$$\text{Dim}\,(\mathcal{N}_\nu) = \text{Dim}\,(\mathcal{N}_{\bar{\nu}})$$

holds, where the dimension of $\mathcal{N}_\nu$, $\text{Dim}\,(\mathcal{N}_\nu)$, is called the deficiency index of L_0 on I and is denoted by d. We have in fact that $0 \leq d \leq 2n$.

A self-adjoint realization of the the equation (4.2) in $\mathcal{H}$ is any linear bounded operator L satisfying the relationships

$$L_0 \subset L = L^* \subset L_1.$$

These self-adjoint realizations are distinguished from one another by their domains. Naimark [33] established the following decomposition

$$D = D_0 \dot{+} \text{ span}\,\{\phi_1, \phi_2, \ldots, \phi_d\} \tag{4.11}$$

of the domain of L via an arbitrary orthonormal basis

$$\phi_1, \phi_2, \ldots, \phi_d$$

in the deficiency space $\mathcal{N}_\nu$ of L_0. Observe that D_1 is always a $2d-$dimensional extension of D_0 and that D is a $d-$dimensional extension of D_0. It follows furthermore that D_1 is a $d-$dimensional extension of D.

The fundamental Glazman-Krein-Naimark (GKN) Theorem [16] characterizes these domains as follows.

Theorem 4.2

(GKN characterization of domains). Let $d \in \mathbb{N}$ be the deficiency index of L_0. A linear submanifold D of D_1 is the domain of a self-adjoint extension L of L_0 with deficiency index d if and only if there exist functions $x_1, x_2, \ldots, x_d$ in D_1 satisfying the following conditions:

(i) $x_1, x_2, \ldots, x_d$ are linearly independent modulo D_0;

(ii) $[x_i, x_j]_a^b = 0, \quad i, j = 1, 2, \ldots, d$;

(iii) $D = \{x \in D_1 : [x, x_i]_a^b = 0, \quad i = 1, 2, \ldots, d\}$. ■

The bracket $[\cdot,\cdot]_a^b$ in Theorem 4.2 is called the Lagrange bracket and is defined for any $x, z \in D_1$ and $t \in I$ by

$$[x, z]\,(t) := \sum_{i=1}^{n} \left\{ x^{[i-1]}(t) z^{[2n-i]}(t) - x^{[2n-i]}(t) z^{[i-1]}(t) \right\}. \tag{4.12}$$

It is worth mentioning that the limits in (4.12) as $t \to a^+$ and as $t \to b^-$ exist and are denoted, respectively, by

$$\lim_{t \to a^+} [x, z]\,(t) = [x, z]\,(a), \qquad \lim_{t \to b^-} [x, z]\,(t) = [x, z]\,(b).$$

We can also write the expression

$$[x, z]_{t_0}^{t_1} = [x, z]\,(t_1) - [x, z]\,(t_0)$$

and observe the validity of the Lagrange identity

$$\int_a^b l(x) z dt - \int_a^b x l(z) dt = [x, z]_a^b \;\text{ for any }\; x, z \in D_1. \tag{4.13}$$

Recall that the operator

$$R_\nu := (L - \nu E)^{-1}$$

is known as the resolvent operator of L with respect to the complex number ν. It follows from assumption (H4) that the mapping L is one-to-one and zero is a regular point of L. This implies that the resolvent $R_0 = L^{-1}$ exists as a bounded operator defined on the whole space $\mathcal{H}$.

Furthermore, it is an integral operator with the kernel K, see Lemma 2.40, satisfying

$$\int_I |K(\tau,t)|^2\, w(\tau)d\tau < \infty \quad \text{and} \quad \int_I |K(\tau,t)|^2\, w(t)dt < \infty.$$

Thus for any function $y \in D$ we can be write

$$y = R_0 f = \int_I K(\tau,t)g(\tau)w(\tau)d\tau \quad \text{a.e.} \quad t \in I, \tag{4.14}$$

where g is some element of $\mathcal{H}$.

Lemma 4.3

Let L_0 be a minimal operator generated by l, as before. Then under Assumption (H4) the range of L_0, $\mathcal{R}_0$, is a closed subspace of $L^2(I,w)$. □

Proof.

Let $\{y_k\} \subset R_0$ be a convergent sequence to y. Then there exists a sequence $\{x_k\}$ in D_0 such that $L_0 x_k = y_k$. By Assumption (H4), L_0, being the restriction of L on D_0, is bounded below; therefore,

$$\|x_j - x_i\|^2 \leq (1/\gamma)\langle L_0(x_j - x_i), x_j - x_i\rangle = \langle y_j - y_i, x_j - x_i\rangle \to 0,$$

which shows that $\{x_k\}$ is Cauchy and therefore convergent. So $x_k \to x$ with $x \in \mathcal{H}$. But L_0 is a closed operator; implying that $x \in D_0$ and furthermore, $L_0 x = y$. This shows that y belongs to R_0 and concludes the proof of this lemma. ■

Next we define the projection operator P onto the range $\mathcal{R}_0$ of L_0. First observe from the domain decomposition (4.11) and from Lemma 4.3 that

$$\mathcal{H} = \tilde{\mathcal{R}} = \mathcal{R}_0 \oplus \mathcal{R}_0^{\perp},$$

where $\tilde{\mathcal{R}}$ is the range of L, and where $\mathcal{R}_0^{\perp}$ is the corresponding $d-$dimensional subspace of H. Let

$\{z_i\}_{i=1}^d$ be an orthonormal basis of $\mathcal{R}_0^\perp$, and let $\{x_i\}_{i=1}^d \subset D$ be such that $Lx_i = z_i$ for $i = 1, \dots, d$. It is clear that $\{x_i\}_{i=1}^d$ is linearly independent modulo D_0. Finally, define P on H as

$$P(y) := (E - Q)y, \qquad y \in \mathcal{H}, \tag{4.15}$$

where Q is the projection onto $\mathcal{R}_0^\perp$ given by

$$Q(y) = \sum_{i=1}^{d} \langle y, z_i \rangle z_i, \qquad y \in \mathcal{H}. \tag{4.16}$$

By the fundamental Theorem 4.2, we may assume that

$$D = D_0 \dot{+} \ \text{span}(\{x_1, x_2, \dots, x_d\}). \tag{4.17}$$

Take further $g \in \mathcal{H}$ with $Lx = g$. Then we have the equalities

$$Lx = Lx_0 + \sum_{i=1}^{n} \alpha_i Lx_i = Lx_0 + \sum_{i=1}^{n} \alpha_i z_i,$$

$$Lx = g = P(g) + Q(g).$$

Both elements Lx_0 and $P(g)$ belong to $\mathcal{R}_0$, while $\sum_{i=1}^{n} \alpha_i z_i$ and $Q(g)$ belong to $\mathcal{R}_0^\perp$. Since the sum in (4.11) is in fact a direct sum, it gives us therefore that

$$Lx_0 = P(g) \quad \text{and} \quad \sum_{i=1}^{n} \alpha_i z_i = Q(g).$$

We summarize our discussions in the following lemma, which justifies the well-posedness of the projection operator P that appears in the construction of the Hamilton-Pontryagin function (4.6) used in our main result.

Lemma 4.4

Let $Lx = g$ with $g \in \mathcal{H}$, and let

$$x = x_0 + \sum_{i=1}^{n} \alpha_i x_i \quad \text{with } x_0 \in D_0.$$

Then we have the representation of x_0 via the projection operator:

$$x_0 = R_0(P(g)).$$

□

4.3 Existence of Solutions to Operator Equations

In this section we derive new results on the existence of solutions of the primal operator equation (4.1) in the domain D and of the adjoint equation (4.8) in the domain D_1. Besides of their own independent interest, the results obtained are important for the proof of our main Theorem 4.1 on the Maximum Principle.

We begin with the following lemma, which can be also seen as a consequence of the existence result from [34, Theorem 15]. Although throughout this chapter all the assumptions (H1)–(H7) are imposed to hold, the reader can see from the proofs that only parts of these assumptions are used in the results below.

Lemma 4.5

Equation (4.1) has at least one solution in D for any feasible control $u(\cdot)$. □

Proof.

By assumption (H2) the proof is complete if we show that there exists a $\rho > 0$ such that the inequality

$$\langle L(y) - F_u(y), y \rangle > 0$$

holds for all $y \in D$ with $\|y\| = \rho$. To proceed, take $y \in D$ and then compute

$$\langle L(y) - F_u(y), y\rangle = \langle L(y), y\rangle - \langle F_u(y) - F_u(0), y\rangle - \langle F_u(0), y\rangle.$$

Using assumption (H4) on L, assumption (H3) on F_u, and the classical Cauchy-Schwartz inequality give us

$$\begin{aligned}\langle L(y) - F_u(y), y\rangle &\geq \gamma\|y\|^2 - \eta\|y\|^2 - \|F_u(0)\|\|y\| \\ &= (\gamma - \eta)\|y\|^2 - \|F_u(0)\|\|y\|.\end{aligned}$$

Now choosing $\rho > \|F_u(0)\|/(\gamma - \eta)$ and taking into account that $\gamma > \eta$, we get

$$\langle L(y) - F_u(y), y\rangle > 0 \quad \text{for all } y \in D,$$

which completes the proof of the lemma. ∎

The result of Lemma 4.5 can be treated as the justification of controllability of the primal differential operator system (4.1) with measurable controls.

The next lemma plays a crucial role in justifying the existence of solutions to boundary value problem for the adjoint system (4.8), which is the main result of this section; see Theorem 4.7 below.

Lemma 4.6

Let $h_1 \in \mathcal{H}$ be such that

$$\langle h_1 z, z\rangle \leq \eta\|z\|^2 \text{ for all } z \in \mathcal{H},$$

where η is taken from assumption (H3). Let $d \in \mathbb{N}$ be the deficiency index of L_0, and let the functions $x_1, \ldots, x_d$ are taken from (4.17). Then for any $h_2 \in \mathcal{H}$ and for arbitrary real numbers $\alpha_i, i = 1, \ldots, d$, the equation

$$\left.\begin{aligned}(L_1 x)(t) &= h_1(t)x(t) + h_2(t), \qquad t \in I \\ [x, x_i]_a^b &= \alpha_i, \qquad i = 1, \ldots, d\end{aligned}\right\} \tag{4.18}$$

admits a solution in the domain D. $\square$

Proof.

Let $\{\xi_1, \dots, \xi_d\}$ be a linearly independent set in D_1 modulo D. Construct the following quadratic matrix

$$A := \begin{bmatrix} [\xi_1, x_1]_a^b & [\xi_2, x_1]_a^b & \dots & [\xi_d, x_1]_a^b \\ [\xi_1, x_2]_a^b & [\xi_2, x_2]_a^b & \dots & [\xi_d, x_2]_a^b \\ \vdots & \vdots & \ddots & \vdots \\ [\xi_1, x_d]_a^b & [\xi_2, x_d]_a^b & \dots & [\xi_d, x_d]_a^b \end{bmatrix}$$

and check that this matrix is invertible. Indeed, otherwise there exists a nonzero vector u such that $Au = 0$. This gives

$$\sum_{j=1}^{d} \left([\xi_j, x_i]_a^b\right)_{i=1}^{d} u_j = \left(\left[\sum_{j=1}^{d} u_j \xi_j, x_i\right]_a^b\right)_{i=1}^{d} = 0,$$

and thus we arrive at the equality

$$\left[\sum_{j=1}^{d} u_j \xi_j, x_i\right]_a^b = 0 \qquad \text{for all } i = 1, \dots, d$$

implying by Theorem 4.2 that $\sum_{j=1}^{d} u_j \xi_j \in D$. The latter contradicts the fact that the functions $\xi_j, j = 1, \dots, d$, are linearly independent modulo D.

Using the invertibility of A^{-1}, define $\beta = (\beta_1, \dots, \beta_d)$ by

$$\beta := A^{-1}\alpha,$$

with $\alpha = (\alpha_1, \dots, \alpha_d)^T$ and choose $\tilde{x} \in D$ to be a solution of

$$L\tilde{x} = h_1\tilde{x} + \sum_{i=1}^{d} \beta_i \left(h_1\xi_i - L_1\xi_i\right) + h_2. \tag{4.19}$$

Then we see that the element

$$x := \tilde{x} + \sum_{i=1}^{d} \beta_i \xi_i$$

is certainly a solution to (4.18). It remains to show that equation (4.19) admits a solution in D. To proceed, we define the function

$$F(z) := h_1 z + h_3 \text{ for any } z \in D,$$

where $h_3 := \sum_{i=1}^{d} \beta_i (h_1 x_i - L x_i) + h_2$. The function F is obviously weakly continuous, and furthermore we have

$$\begin{aligned} \langle Lz - F(z), z \rangle &= \langle Lz - h_1 z - h_3, z \rangle = \langle Lz, z \rangle - \langle h_1 z, z \rangle - \langle h_3, z \rangle \\ &> \gamma \|z\|^2 - \eta \|z\|^2 - \|h_3\| \|z\| = (\gamma - \eta) \|z\|^2 - \|h_3\| \|z\|. \end{aligned}$$

This ensures the existence of a solution to (4.18) in D by [34, Theorem 15] with

$$\rho > \frac{\|h_3\|}{\gamma - \eta},$$

which completes the proof of this theorem. ▮

Now we are ready to establish the existence of solutions to the adjoint system (4.8), (4.9) in the required domain D_1.

Theorem 4.7 (existence of solutions to the adjoint system)

The adjoint equation (4.8) with the boundary conditions (4.9) admits a solution in D_1. ■

Proof.

Let $r \in \mathbb{R}$, and let O be a neighborhood of $\bar{x}$ from (H1). Taken any $x \in O$ and observe from (H3) that

$$\langle F_{\bar{u}}(\bar{x} + rx) - F_{\bar{u}}(\bar{x}), rx \rangle \leq \eta r^2 \|x\|^2.$$

Dividing by r^2 both sides of this inequality and taking the limit as $r \to 0$ give us

$$\Big\langle \lim_{r\to 0} \frac{F_{\bar u}(\bar x + rx) - F_{\bar u}(x)}{r}, x \Big\rangle \le \eta \|x\|^2,$$

which yields, by the Fréchet differentiability of F_u at $\bar x$, that

$$\langle F'_{\bar u}(\bar x)x, x\rangle \le \eta\|x\|^2.$$

The latter estimate allows us to complete the proof of the theorem by putting there

$$h_1 := F'_{\bar u}(\bar x) \quad \text{and} \quad h_2 := P(\phi(x(c))K(c,\cdot))F'_{\bar u}(\bar x)$$

and applying finally Lemma 4.6. ∎

4.4 Proof of the Maximum Principle

This section is devoted to the proof of our main result on the Maximum Principle for optimal solutions to problem (4.3) under the standing assumption formulated in Theorem 4.1. The proof is based on the results on the primal and adjoint operator equation presented in the previous sections and the optimal control techniques developed below. We split the proof into several steps.

Given two feasible controls $\bar u(t), u(t) \in U$ a.e. and taking the corresponding solutions $\bar x(\cdot), x(\cdot)$ of system (4.1) defined by (4.14), we write the increments

$$\begin{aligned} \Delta \bar u(t) :&= u(t) - \bar u(t),\\ \Delta \bar x(t) :&= x(t) - \bar x(t),\\ \Delta J[\bar u] :&= \phi(x(c)) - \phi(\bar x(c)). \end{aligned}$$

The first lemma in this section justifies the increment formula for the cost functional J needed in what follows.

Lemma 4.8

In the notation above we have the increment formula

$$\begin{aligned} \Delta J[\bar{u}] &= -\langle p + P(\check{K}_c(\cdot)), \Delta_u F'_{\bar{u}}(\bar{x})\Delta\bar{x}\rangle - \langle p + P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}(\bar{x})\rangle \\ &\quad + o(\|\Delta\bar{x}\|) + o(|\Delta\bar{x}(c)|), \end{aligned} \tag{4.20}$$

where K is the kernel of the resolvent operator R_0, $\check{K}_c := \phi_o(c)K(c,\cdot)$, P is the projection onto the range of L_0 defined in (4.15), and

$$\Delta_u F_{\bar{u}}(\bar{x}) := F_u(\bar{x}) - F_{\bar{u}}(\bar{x}).$$

□

Proof.

By (H6), the cost function ϕ is Fréchet differentiable at $\bar{x}(c)$; thus we have

$$\Delta J[\bar{u}] = \phi(x(c)) - \phi(\bar{x}(c)) = \phi'_0(\bar{x}(c))\Delta\bar{x}(c) + o(|\Delta\bar{x}(c)|). \tag{4.21}$$

If $x_i \in D_1, i = 1, \ldots, d$, are the functions that determine L by Theorem 4.2), then every $x \in D$ can be written as

$$x = x_0 + \sum_{i=1}^{d} \beta_i v_i$$

with some x_0 in D_0. For any arcs $x \in D$ and any $p \in D_1$ satisfy the primal and adjoint systems (these solutions exist due to Lemma 4.5 and Theorem 4.7, respectively) we have

$$\begin{aligned} [p, x]_a^b &= [p, x_0]_a^b + \sum_{i=1}^{d} \beta_i [p, x_i]_a^b \\ &= \phi'_0(\bar{x}(c))x_0(c) - \phi'_0(\bar{x}(c)) \left[x_0(c) + \sum_{i=1}^{d} \beta_i x_i(c) \right] \\ &= \phi'_0(\bar{x}(c))x_0(c) - \phi'_0(\bar{x}(c))x(c). \end{aligned}$$

This gives there the representation

$$\phi_0'(\bar{x}(c))\Delta\bar{x}(c) = \phi_0'(\bar{x}(c))\Delta\bar{x}_0(c) - [p, \Delta\bar{x}]_a^b. \tag{4.22}$$

Now using the Lagrange identity (4.13) and elementary transformations implies that

$$\begin{aligned}
[p, \Delta\bar{x}]_a^b &= \langle Lp, \Delta\bar{x}\rangle - \langle p, L\Delta\bar{x}\rangle \\
&= \langle Lp, \Delta\bar{x}\rangle - \langle p, F_u(x) - F_{\bar{u}}(\bar{x})\rangle \\
&= \langle Lp, \Delta\bar{x}\rangle - \langle p, F_u(x) - F_{\bar{u}}(x)\rangle - \langle p, F_{\bar{u}}(x) - F_{\bar{u}}(\bar{x})\rangle \\
&= \langle Lp, \Delta\bar{x}\rangle - \langle p, \Delta_u F_{\bar{u}}(x)\rangle - \langle p, F_{\bar{u}}'(\bar{x})\Delta\bar{x}\rangle + o(\|\Delta\bar{x}\|) \\
&= \langle Lp, \Delta\bar{x}\rangle - \langle p, F_{\bar{u}}'(\bar{x})\Delta\bar{x}\rangle - \langle p, \Delta_u F_{\bar{u}}(x) - \Delta_u F_{\bar{u}}(\bar{x})\rangle - \langle p, \Delta_u F_{\bar{u}}(\bar{x})\rangle + o(\|\Delta\bar{x}\|) \\
&= \langle Lp, \Delta\bar{x}\rangle - \langle p, F_{\bar{u}}'(\bar{x})\Delta\bar{x}\rangle - \langle p, \Delta_u F_{\bar{u}}(\bar{x})\rangle - \langle p, \Delta_u F_{\bar{u}}'(\bar{x})\Delta\bar{x}\rangle + o(\|\Delta\bar{x}\|) \\
&= \langle (L_1 - F_{\bar{u}}'(\bar{x}))p, \Delta\bar{x}\rangle - \langle p, \Delta_u F_{\bar{u}}(\bar{x})\rangle - \langle p, \Delta_u F_{\bar{u}}'(\bar{x})\Delta\bar{x}\rangle + o(\|\Delta\bar{x}\|).
\end{aligned}$$

Employing further the solution representation (4.14), we get

$$\begin{aligned}
\phi_0'(\bar{x}(c))\Delta\bar{x}_0(c) &= \phi_0'(\bar{x}(c))(x_0(c) - \bar{x}_0(c)) \\
&= \phi_0'(\bar{x}(c))\left[\int_a^b K_c(s)P(F_u(x) - F_{\bar{u}}(\bar{x}))(s)w(s)ds\right] \\
&= \int_a^b \check{K}_c(s)P(F_u(x) - F_{\bar{u}}(x) + F_{\bar{u}}(x) - F_{\bar{u}}(\bar{x}))(s)w(s)ds \\
&= \int_a^b \check{K}_c(s)P(\Delta_u F_{\bar{u}}(x) + F_{\bar{u}}'(\bar{x})\Delta\bar{x})(s)w(s)ds + o(\|\Delta\bar{x}\|) \\
&= \int_a^b \check{K}_c(s)P(F_{\bar{u}}'(\bar{x})\Delta\bar{x} + \Delta_u F_{\bar{u}}(x) - \Delta_u F_{\bar{u}}(\bar{x}) + \Delta_u F_{\bar{u}}(\bar{x}))(s)w(s)ds + o(\|\Delta\bar{x}\|) \\
&= \int_a^b \check{K}_c(s)P(F_{\bar{u}}'(\bar{x})\Delta\bar{x} + \Delta_u F_{\bar{u}}'(\bar{x})\Delta\bar{x} + \Delta_u F_{\bar{u}}(\bar{x}))(s)w(s)ds + o(\|\Delta\bar{x}\|) \\
&= \langle \check{K}_c(\cdot), P(F'\bar{u}(\bar{x})\Delta\bar{x})\rangle + \langle \check{K}_c(\cdot), P(\Delta_u F_{\bar{u}}'(\bar{x})\Delta\bar{x})\rangle + \langle \check{K}_c(\cdot), P(\Delta_u F_{\bar{u}}(\bar{x}))\rangle + o(\|\Delta\bar{x}\|) \\
&= \langle F_{\bar{u}}'(\bar{x})P(\check{K}_c(\cdot)), \Delta\bar{x})\rangle + \langle P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}'(\bar{x})\Delta\bar{x}\rangle + \langle P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}(\bar{x})\rangle + o(\|\Delta\bar{x}\|).
\end{aligned}$$

Substituting the obtained expressions for $[p, \Delta\bar{x}]_a^b$ and $\phi_0'(\bar{x}(c))\Delta\bar{x}_0(c)$ into (4.22) yields

$$\begin{aligned}\phi_0'(\bar{x}(c))\Delta\bar{x}(c) &= \langle F_{\bar{u}}'(\bar{x})P(\check{K}_c(\cdot)), \Delta\bar{x})\rangle + \langle P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}'(\bar{x})\Delta\bar{x}\rangle + \langle P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}(\bar{x})\rangle \\ &\quad -\langle (L_1 - F_{\bar{u}}'(\bar{x}))p, \Delta\bar{x}\rangle + \langle p, \Delta_u F_{\bar{u}}(\bar{x})\rangle + \langle p, \Delta_u F_{\bar{u}}'(\bar{x})\Delta\bar{x}\rangle + o(\|\Delta\bar{x}\|) \\ &= \langle -Lp + F_{\bar{u}}'(\bar{x})p + F_{\bar{u}}'(\bar{x})P(\check{K}_c(\cdot)), \Delta\bar{x})\rangle + \langle p + P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}'(\bar{x})\Delta\bar{x}\rangle \\ &\quad + \langle p + P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}(\bar{x})\rangle + o(\|\Delta\bar{x}\|)\end{aligned}$$

Taking finally formula (4.21) into account, we arrive at

$$\begin{aligned}\Delta J[\bar{u}] &= \langle -Lp + F_{\bar{u}}'(\bar{x})p + F_{\bar{u}}'(\bar{x})P(\check{K}_c(\cdot)), \Delta\bar{x})\rangle + \langle p + P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}'(\bar{x})\Delta\bar{x}\rangle \\ &\quad + \langle p + P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}(\bar{x})\rangle + o(\|\Delta\bar{x}\|) + o(|\Delta\bar{x}(c)|)\end{aligned}$$

and thus complete the proof of the lemma. ∎

Note that the derivation of the increment formula in Lemma 4.8 is different from the usual way know in control theory (compare, i.e., [31, Lemma 6.43]) in the sense that we take advantage of the well-developed theory of the differential operator equations under consideration. The next two lemmas are designed to estimate the trajectory increments in both functional $\Delta\bar{x}$ and pointwise $\Delta\bar{x}(c)$ form by building a single needle variation $u(\cdot)$ of the reference control $\bar{u}(\cdot)$.

To proceed, fix a set $I_\epsilon \subset I$ of finite measure ϵ, take a measurable mapping v such that $v(t) \in U$ a.e. $t \in I_\epsilon$, and define $u(t), t \in I$, as follows:

$$u(t) = \begin{cases} v(t), & t \in I_\epsilon, \\ \bar{u}(t), & t \notin I_\epsilon. \end{cases} \tag{4.23}$$

Lemma 4.9

Let $\Delta\bar{x} = \Delta\bar{x}(\cdot)$ be the increment of $\bar{x}(\cdot)$ corresponding to the needle variation (4.23) of $\bar{u}(\cdot)$. Then we have the functional trajectory increment estimate

$$\|\Delta\bar{x}\| = o(\epsilon). \tag{4.24}$$

Proof.

The semi-boundedness assumption of the operator L in (H4) and the monotonicity property of F_u in (H3) lead us to the relationships

$$\begin{aligned} \gamma\|\Delta\bar{x}\|^2 &\leq \langle L\Delta\bar{x}, \Delta\bar{x}\rangle \\ &= \langle F_u(x) - F_{\bar{u}}(\bar{x}), \Delta\bar{x}\rangle \\ &= \langle F_u(x) - F_u(\bar{x}) + F_u(\bar{x}) - F_{\bar{u}}(\bar{x}), \Delta\bar{x}\rangle \\ &= \langle F_u(x) - F_u(\bar{x}), \Delta\bar{x}\rangle + \langle \Delta_u F_{\bar{u}}(\bar{x}), \Delta\bar{x}\rangle \\ &\leq \eta\|\Delta\bar{x}\|^2 + \|\Delta_u F_{\bar{u}}(\bar{x})\|\|\Delta\bar{x}\|. \end{aligned}$$

Employing further assumption (H5) ensures that

$$(\gamma - \eta)\|\Delta\bar{x}\| \leq \|\Delta_u F_{\bar{u}}(\bar{x})\| = o(\epsilon),$$

and thus we arrive at (4.24). ∎

Lemma 4.10

The following pointwise trajectory increment estimate holds:

$$|\Delta\bar{x}(c)| = o(\epsilon).$$ □

Proof.

By using the pointwise representation of the trajectory (4.4) corresponding to the needle variation $\bar{u}(\cdot)$, we have

$$\begin{aligned}
|\Delta\bar{x}(c)| &= |x(c)-\bar{x}(c)| \\
&= \left|\int_I K_c(s)(F_u(x)-F_{\bar{u}}(\bar{x}))(s)w(s)ds\right| \\
&= \left|\int_I K_c(s)(\Delta_u F_u(x)-\Delta_x F_{\bar{u}}(\bar{x}))(s)w(s)ds\right| \\
&\le \left|\int_{I_\epsilon} K_c(s)(\Delta_u F_u(x))(s)w(s)ds\right| + \left|\int_I K_c(s)(\Delta_x F_{\bar{u}}(\bar{x}))(s)w(s)ds\right|.
\end{aligned}$$

The second term of the above inequality can be split into

$$\left|\int_I K_c(s)(\Delta_x F_{\bar{u}}(\bar{x}))(s)w(s)ds\right| = \left|\int_I K_c(s)F'_{\bar{u}}(\bar{x})(s)\Delta\bar{x}(s)w(s)ds\right| + \left|\int_I K_c(s)o(\epsilon)w(s)ds\right|.$$

Using further the assumed continuity of $F'_u(\bar{x})$ and Lemma 4.9 ensure the estimates

$$\left|\int_I K_c(s)F'_{\bar{u}}(\bar{x})(s)\Delta\bar{x}(s)w(s)ds\right| \le \|K_c\|\|F'_u(\bar{x})\Delta\bar{x}\| \le \|K_c\|\|F'_u(\bar{x})\|\|\Delta\bar{x}\| = o(\epsilon),$$

$$\left|\int_I K_c(s)o(\epsilon)w(s)ds\right| = o(\epsilon),$$

which show in turn that

$$|\Delta\bar{x}(c)| = o(\epsilon).$$

and thus justify our claim. ∎

Lemmas 4.9 and 4.10 enable us to rewrite the increment formula (4.20) of Lemma 4.8 as

$$\Delta J[\bar{u}] = -\langle p+P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}(\bar{x})\rangle - \langle p+P(\check{K}_c(\cdot)), \Delta_u F'_{\bar{u}}(\bar{x})\Delta\bar{x}\rangle + o(\epsilon). \tag{4.25}$$

Now all the ingredients required for the justification of the Maximum Principle in Theorem 4.1 (namely, Lemmas 4.8, 4.9, and 4.10) are ready, and we can proceed with the completion of the proof.

Completion of the proof of the Maximum Principle. Let $(\bar{u}, \bar{x})$ be an optimal solution to problem (4.3), and let p be the corresponding solution to the adjoint system (4.8) satisfying the boundary/transversality conditions (4.9). Let us show that the maximum condition (4.7) is also satisfied for $(\bar{u}, \bar{x})$. To proceed, we argue by contradiction and suppose that there exists a set $T \subset I$ of positive measure such that

$$H(\bar{x}(t), p(t), \bar{u}(t) < \sup_{u \in U} H(\bar{x}(t), p(t), u(t)) > 0, \qquad t \in T.$$

Following the proof of [31, Theorem 6.37] by using the theory of measurable selections and taking into account assumption (H7), we conclude that there is a measurable mapping $v\colon T \to U$ such that

$$\Delta_v H(t) := H(\bar{x}(t), p(t), v(t), t) - H(\bar{x}, p(t), \bar{u}(t), t) > 0, \quad t \in T. \tag{4.26}$$

Now let $T_0 \subset I$ be a set of Lebesgue regular points of the function H on I. It is well known that the set T_0 is of full measure on I. Taking any $\tau \in T_0$ and $\epsilon > 0$, consider a needle variation of type (4.23) built by

$$u(t) := \begin{cases} v(t), & t \in I_\epsilon := [\tau, \tau + \epsilon) \cap T_0, \\ \bar{u}(t), & t \in I \setminus I_\epsilon. \end{cases}$$

The increment formula for the cost functional (4.25) corresponding to $\bar{u}$ and u gives us

$$\Delta J[\bar{u}] \quad = \quad -\int_\tau^{\tau+\epsilon} \Delta_v H(t) w(t) dt + \int_\tau^{\tau+\epsilon} \Delta_v F'_{\bar{u}}(\bar{x}(t)) \Delta\bar{x}(t) w(t) dt + o(\epsilon)$$

Assumption (H1) and Lemma 4.9 ensure that

$$\int_\tau^{\tau+\epsilon} \Delta_v F'_{\bar{u}}(\bar{x}(t)) \Delta\bar{x}(t) w(t) dt = o(\epsilon)$$

due to the estimate

$$\int_\tau^{\tau+\epsilon} \Delta_v F'_{\bar{u}}(\bar{x}(t)) \Delta\bar{x}(t) w(t) dt \le \|\Delta_v F'_{\bar{u}}(\bar{x})\| \|\Delta\bar{x}\|.$$

Since τ is a Lebesgue regular point of $\Delta_v H$, we have

$$-\int_{\tau}^{\tau+\epsilon} \Delta_v H(t)w(t)dt = -\epsilon\left[\Delta_v H(\tau)\right] + o(\epsilon),$$

which implies therefore that

$$\Delta J[\bar{u}] = -\epsilon\left[\Delta_v H(\tau)\right] + o(\epsilon).$$

This shows by (4.26) that $\Delta J[\bar{u}] < 0$ along the above needle variation $u(\cdot)$ for all $\epsilon > 0$ sufficiently small, which contradicts the optimality of the reference control $\bar{u}(\cdot)$ for problem (4.3) and thus completes the proof of Theorem 4.1. ∎

4.5 Illustrating Example

In this section we give an example to illustrate the discussion and results above.

Example 4.11

Consider the following quasi-differential expression

$$lx = -(1/t)(tx')', \quad \text{on} \quad I = [0,1].$$

Here $n = 2$ and $w = t$. This expression is singular since $1/t$ is not integrable at 0. We now solve the quasi-differential equations

$$-(1/t)(tx')' = 0.$$

The solution space is spanned by the set $\{y_1 := 1, y_2 := \ln(t)\}$. The expression in the Hilbert space $\mathcal{H}$ generates a minimal operator L_0. The set $\{y_1, y_2\}$ is linearly independent modulo D_0. Furthermore, both functions belong to the Hilbert space $\mathcal{H} = L^2([0,1], t)$ and their quasi-derivatives are locally absolutely continuous; namely

$$1^{[0]} = 1, 1^{[1]} = t \cdot (1)' = 0, \ln(t)^{[0]} = \ln(t), \ln(t)^{[1]} = t \cdot (\ln(t))' = 1 \in AC_{loc}([0,1]).$$

Hence both of y_1 and y_2 are in D_1, the domain of the maximal operator L_1. This shows that d, the deficiency index of L_0, is equal to 2. The range of L_0, R_0, is a closed subspace of $\mathcal{H}$ by Lemma 4.3 and

$$\mathcal{H} = R_0 \oplus R_0^{\perp}.$$

The space $R_0^{\perp}$ is $2-$dimensional subspace in $\mathcal{H}$. The set $\{y_1, y_2\}$ is a linearly independent set in $R_0^{\perp}$. In fact, any solution of the eigenvalue problem

$$lx = 0, \tag{4.27}$$

which belongs to D_1 is a member of $R_0^{\perp}$. To see this let z be a solution of (4.27) and $y \in R_0$ then there exits $x \in D_0$ such that $y = L_0 x = lx$. Therefore,

$$\langle y, z \rangle = \langle lx, z \rangle = 0.$$

Thus z is orthogonal to R_0; that is $z \in R_0^{\perp}$.

Now let L, with domain D, be a self-adjoint extension of L_0. We now solve the following two boundary value problems

$$\left.\begin{aligned} -(1/t)(tx')' &= 1, \\ x^{[1]}(0) &= 0, \\ 3x^{[1]}(1) + 2x^{[0]}(1) &= 0, \end{aligned}\right\}, \quad \text{and} \quad \left.\begin{aligned} -(1/t)(tx')' &= \ln(t), \\ x^{[1]}(0) &= 0, \\ 3x^{[1]}(1) + 2x^{[0]}(1) &= 0, \end{aligned}\right\},$$

giving the solutions $z_1 = 1 - t^2/4$ for the first, and $z_2 = t^2/4(1 - \ln(t)) - 5/8$ for the second. These functions belong to D; because a solution of a second-order quasi-differential equation subject to this type of boundary conditions is a member of D, see formula (10.4.59) in [48]. In addition, both of them are not in D_0; since $z_1(1) \neq 0$ and $z_2(1) \neq 0$, see [33, IV, §17.4]. This means that z_1

and z_2 are linearly independent modulo D_0. Hence, we have the decomposition

$$D = D_0 + \text{span}(\{w_1, w_2\}),$$

where

$$w_1 = \alpha_1 z_1 + \beta_1 z_2, \quad w_2 = \alpha_2 z_1 + \beta_2 z_2, \quad \text{with} \quad \alpha_k, \beta_k, \ \ k = 1, 2 \in \mathbb{R}.$$

The two functions w_1 and w_2 are the ones mentioned in Theorem 4.2. Also, we have the decompostion for D_1

$$D_1 = D + \text{span}(\{w_3, w_4\}).$$

where

$$\{w_3 := 2, w_4 := \sqrt{2}(2\ln t + 1)\}$$

is an orthonormal set given by $\{y_1, y_2\}$ through Gram-Schmidt process.

We define now the projection operator, $P : D_0 \to R_0$, that appears in the Hamilton-Pontryagin function (4.6) as follows.

$$Px = (1 - Q)(x), \quad x \in D_0,$$

where $Q : \text{span}(\{w_1, w_2\}) \to R_0^\perp = \text{span}(\{w_3, w_4\})$ is defined as

$$Qx = \langle x, w_3\rangle w_3 + \langle x, w_4\rangle w_4.$$

We now turns our attention to dynamical system that governs our optimal control (4.3), that is,

$$\left.\begin{array}{rcl} -(1/t)(tx')'(t) & = & u(t), \quad t \in I \ \text{ a.e. } \ |u(t)| \leq 1, \\ x^{[1]}(0) & = & 0, \\ 3x^{[1]}(1) + 2x^{[0]}(1) & = & 0. \end{array}\right\}$$

We solve the equation $-(1/t)(tx')'(t) = u(t)$ using the method of variation of parameters to obtain the solution

$$y(t) = a_1 + a_2 \ln t + \int_c^t v_1(\tau)u(\tau)d\tau + \ln t \int_c^t v_2(\tau)u(\tau)d\tau,$$

where $c \in (a, b]$, a_1, a_2 are arbitrary scalars, see Lemma 2.16, and

$$v_1(t) = -\ln t, \quad v_2(t) = 1.$$

The two functions v_1 and v_2 are essential in construction of the kernel, $K(c, s)$, of the resolvent of D appeared also in (4.6) in the following manner, see [33, Theorem 1, §19.2].

$$K(c,\tau) = \left\{ \begin{array}{ll} \sum_{k=1}^{2} y_k(c)h_k(\tau), & c \leq \tau \\ \sum_{k=1}^{2} y_k(c)\left[h_k(\tau) + v_k(\tau)\right], & c \leq \tau \end{array} \right\} + \sum_{k=1}^{2} y_k(c)\bar{h}_k(\tau).$$

where h_1 and h_2 are the solutions of the following system

$$\begin{bmatrix} [y_1, w_1]_a^b & [y_2, w_1]_a^b \\ [y_1, w_2]_a^b & [y_2, w_2]_a^b \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} [y_1, w_1]_a v_1 & [y_2, w_1]_a v_2 \\ [y_1, w_2]_a v_1 & [y_2, w_2]_a v_2 \end{bmatrix}$$

An optimal solution $(\bar{u}(\cdot), \bar{x}(\cdot))$ to the problem

$$\left.\begin{array}{l} \text{minimize } \mathcal{J}[u,x] := \phi_0(x(1)) \\ \text{subject to} \\ \qquad \begin{array}{ll} -(1/t)(tx')'(t) & = u(t), \quad t \in I \text{ a.e. } \quad |u(t)| \leq 1, \\ x^{[1]}(0) & = 0, \\ 3x^{[1]}(1) + 2x^{[0]}(1) & = 0. \end{array} \end{array}\right\}$$

satisfies, according to Theorem (4.8),

$$H(\bar{x}(t), p(t), \bar{u}(t), t) := \max_{u \in U} \left(p(t) + P(\phi(\bar{x}(c))K(c,t))\right) u \quad \text{a.e. } t \in I,$$

where $p : I \to \mathbb{C}$ such that $p^{[0]}, p^{[1]} \in AC_{loc}([0,1])$, $p, -(1/t)(tp')' \in L^2([0,1], t)$ and

$$(1/t)(tp')' = \nabla_x H(\bar{x}(t), p(t), \bar{u}(t), t) \quad \text{a.e.}$$

with the transversality conditions

$$\begin{aligned} [p, w_1]_a^b &= -\phi_0'(\bar{x}(c))w_1(c), \\ [p, w_2]_a^b &= -\phi_0'(\bar{x}(c))w_2(c). \end{aligned}$$

□

CHAPTER 5

CONCLUSIONS AND FURTHER RESEARCH

In this thesis we formulated, for the first time in the literature, an optimal control problem for self-adjoint ordinary differential operator equations in Hilbert spaces and derived necessary conditions for optimal controls to this problem in an appropriate extended form the Pontryagin Maximum Principle. Our treatment to derive the Pontryagin Maximal Principle relied heavily on the well-developed theory of quasi-differential expressions and the operators they generate in an appropriate Hilbert space. The reader can see this in our version of the Hamilton-Pontryagin function (4.6) which involves the projection onto the orthogonal complement of the range of a minimal operator L_0 associated with l and the kernel function of the resolvent of its self-adjoint extension, well one of them anyway. The work we developed in this thesis was accepted for publication [2].

We believe that this work opens a door to more work of potential significance on many levels. The following is a list of some problems that we think are worth investigating.

▶ **(Equality and Inequality Constraints)**

The first, and quite natural, problem is to consider a constrained end-point problem rather than a free-end point one. Namely, the problem

$$\left.\begin{aligned}
&\text{minimize } \mathcal{J}[u,x] := \phi_0(x(c)) \\
&\text{subject to} \\
&\qquad Lx = f(x,u,t) \quad \text{a.e.}, \;\; t \in I = (a,b), \quad -\infty \le a < b \le \infty \\
&\qquad u(t) \in U \quad \text{a.e.}, \\
&\qquad \phi_k(x(c)) \le 0 \quad \text{for} \quad k = 1, \cdots, m, \\
&\qquad \phi_k(x(c)) = 0 \quad \text{for} \quad k = m+1, \cdots, m+r,
\end{aligned}\right\}$$

where $\phi_k, k = 0, \cdots, m+r$ are real-valued functions. Under Assumptions (H1)–(H7) and that only $\phi_{m+k}, k = 1, \cdots, r$ are continuous around $\bar{x}(c)$ and Fréchet differentiable at $\bar{x}(c)$, we conjuncture that Theorem 4.1 stands true with the following transversality conditions

$$[p, x_i]_a^b = -\sum_{k=o}^{m+r} \mu_k \phi_k'(\bar{x}(c)) x_i(c), \qquad i = 1, \cdots, d,$$

where $\mu_k, k = 0, \cdots, m+r$ are multipliers satisfying

$$\begin{aligned}
&(\mu_0, \cdots, \mu_{m+r}) \neq 0 \\
&\mu_k \ge 0 && \text{for} \quad k = 0, \cdots, m, \\
&\mu_k \phi_k(\bar{x}(c)) = 0 && \text{for} \quad k = 1, \cdots, m.
\end{aligned}$$

▶ **(Matrix Quasi-Differential Expressions)**

Let $I = (a,b)$ be an interval with $\infty \le a < b \le \infty$, n, m be positive integers. For a given set S, $M_{n,m}(S)$ denotes the set of $n \times m$ matrices with entries in S. If $n = m$, we write also

$M_n(S)$ and if $m = 1$ we write S^n. Let

$$\begin{aligned} Z_{n,m}(I) \quad := \quad & \{\mathbb{Q} = (Q_{rs})_{r,s=1}^{n} \in M_n(M_m(L_{loc}(I))), \\ & Q_{r,s} = 0, \quad \text{a.e.} \quad \text{on} \quad I, \quad \text{for} \quad 2 \leq r+1 < s \leq n, \\ & Q_{r,r+1} \ \text{invertible} \quad \text{a.e.} \quad \text{on} \quad I, \quad Q_{r,r+1}^{-1} \in M_m(L_{loc}(I)) \quad \text{for} \quad 1 \leq r \leq n-1\}. \end{aligned}$$

Let $Q \in Z_{n,m}(I)$. We define

$$V_0 := \{x : I \to \mathbb{C}^m, \quad x \text{ is measurable}\}.$$

The quasi-derivatives $x^{[k]}$ for $k = 0, \cdots, n$, are defined inductively as

$$\begin{aligned} x^{[0]} \quad &:= \quad x, \quad x \in V_0, \\ x^{[k]} \quad &:= \quad Q_{k,k+1}^{-1} \left\{ \left(x^{[k-1]}\right)' - \sum_{s=1}^{k} Q_{ks} x^{[s-1]} \right\}, \quad x \in V_k \quad \text{for} \quad k = 1, \cdots, n \end{aligned}$$

where $q_{n,n+1} := I_m$, the $m \times m$ identity matrix, and

$$V_k := \left\{ x \in V_{k-1} : x^{[k-1]} \in (AC_{loc}(I))^m \right\}, \quad \text{for} \ \ k = 1, \cdots, n.$$

Finally we set

$$l_{\mathbb{Q}} x := i^n x^{[n]} \quad (x \in V_n).$$

The expression l_A is called the quasi-differential expression with matrix coefficients associated with $\mathbb{Q}$. This is a linear operator from V_n to $(L_{loc}(I))^m$, see [28].

An interesting extension of our work is to study the optimal control for operators generated by this generalized quasi-differential expression. Aside from the technicalities expected, we believe that our findings in this thesis can be extended to cover problems defined in terms of matrix quasi-differential expressions.

▶ **(Numerical Aspects)**

As it is clear from our discussion in Chapter 4. We are interested in solving the equation

$$Lx = f$$

with L an arbitrary self-adjoint extension of L_0. Numerical methods such as the Galerkin method proves to be effective and more natural for solving such equations, see e.g. [11]. We see promising prospects in exploring numerical methods specially designed to facilitate the optimal control problem under consideration in this thesis.

▶ **(Differential Inclusions)**

Let X be a Banach space, and let $I := [a, b]$ be a time interval of the real line. Consider a set-valued mapping $F : X \times T \rightrightarrows X$ and define the differential/evolution inclusion

$$\dot{x}(t) \in F(x(t), t) \quad \text{a.e.} \quad t \in [a, b] \tag{5.1}$$

generated by F, where $\dot{x}(t)$ stands for the time derivative of $x(t)$. By a solution to the above inclusion (5.1) we understand a mapping $x : I \to X$, which is Fréchet differentiable for a.e.$t \in I$ and satisfies (5.1) and the Newton-Leibniz formula

$$x(t) = x(a) = \int_a^t \dot{x}(\tau) d\tau \quad \text{for all} \quad t \in I,$$

where the integral is taken in the Bochner sense.

The study of optimal control for dynamic/evolution systems governed by differential inclusions and their finite difference approximations in appropriate Banach spaces is appealing because these models capture more conventional problems of optimal control described by parameterized differential equations. The success in this regards, see [42, 31], is encouraging to reformulate our problem as an inclusion. This idea, though attractive, needs a lot of work in developing the theory to handle a problem of the form

$$Lx \in F \quad \text{a.e.} \quad t \in I.$$

where L is a self-adjoint operator extending a minimal operator L_0 generated by a quasi-differential expression, or maybe a more general form of the one we considered, in a Hilbert space or even in a Banach space.

► **(Optimal Control of Operator Equations)**

A last, but definitely not least, problem is the study optimal control problems for operator equations in the form

$$Ax(t) = f(x, u, t) \quad t \in I \;\; \text{a.e.},$$

where A is a general linear operator defined on a Banach space X. Many interesting questions are in order. Among these are: what is a solution of this equation look like? how to define the Hamilton-Pontryagin function? What kind of assumptions we need to impose on A to develop necessary optimality conditions?

REFERENCES

[1] N. I. AKHIEZER and I. M. GLAZMAN. THEORY OF LINEAR OPERATORS IN HILBERT SPACE. Dover Publications Inc., New York, 1993.

[2] M. M. ALSHAHRANI, M. A. EL-GEBEILY, and B. S. MORDUKHOVICH. Maximum principle for optimal control systems governed by singular ordinary differential operators in hilbert spaces. Dynam. Systems Appl. Accepted.

[3] J. AUBIN and H. FRANKOWSKA. SET-VALUED ANALYSIS. Birkhäuser, Boston, Massachusetts, 1990. doi:10.1007/978-0-8176-4848-0.

[4] J. H. BARRETT. Oscillation theory of ordinary linear differential equations. Advances in Mathematics, volume 3(4):pp. 415–509, 1969. doi:10.1016/0001-8708(69)90008-5.

[5] A. E. BRYSON. Optimal control-1950 to 1985. Control Systems Magazine, IEEE, volume 16(3), 1996. doi:10.1109/37.506395.

[6] F. H. CLARKE. The maximum principle under minimal hypotheses. SIAM Journal on Control and Optimization, volume 14(6):pp. 1078–1091, 1976.

[7] F. H. CLARKE. OPTIMIZATION AND NONSMOOTH ANALYSIS. Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons Inc., New York, 1983.

[8] F. H. CLARKE. Necessary conditions in dynamic optimization. Memoirs of the American Mathematical Society, volume 173(816), 2005.

[9] F. H. CLARKE. The Pontryagin maximum principle and a unified theory of dynamic optimization. Proceedings of the Steklov Institute of Mathematics, volume 268(1):pp. 58–69, 2010. doi:10.1134/S0081543810010062.

[10] N. DUNFORD and J. T. SCHWARTZ. LINEAR OPERATORS. PART II: SPECTRAL THEORY. SELF ADJOINT OPERATORS IN HILBERT SPACE. With the assistance of William G. Bade and Robert G. Bartle. Interscience Publishers John Wiley & Sons
New York-London, 1963.

[11] M. A. EL-GEBEILY, K. M. FURATI, and D. O'REGAN. The finite element-Galerkin method for singular self-adjoint differential equations. J. Comput. Appl. Math., volume 223(2):pp. 735–752, 2009. doi:10.1016/j.cam.2008.02.011.

[12] M. A. EL-GEBEILY, D. O'REGAN, and R. AGARWAL. Characterization of self-adjoint ordinary differential operators. Mathematical and Computer Modelling, volume 54(1-2):pp. 659–672, 2011. doi:10.1016/j.mcm.2011.03.009.

[13] W. N. EVERITT. A Catalogue of Sturm-Liouville Differential Equations, pp. 271–331. Birkhäuser Verlag, Basel/Switzerland, 2005.

[14] W. N. EVERITT and L. MARKUS. Controllability of [r]-matrix quasi-differential equations. Journal of Differential Equations, volume 89(1):pp. 95–109, 1991. doi:10.1016/0022-0396(91)90113-N.

[15] W. N. EVERITT and L. MARKUS. The Glazman-Krein-Naimark theorem for ordinary differential operators, pp. 118–130. Birkhauser Verlag, Basel, Switzerland, Switzerland, 1997. ISBN 3-7643-5775-4.

[16] W. N. EVERITT and L. MARKUS. BOUNDARY VALUE PROBLEMS AND SYMPLECTIC ALGEBRA FOR ORDINARY DIFFERENTIAL AND QUASI-DIFFERENTIAL OPERATORS, volume 61 of Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1999. ISBN 0-8218-1080-4.

[17] W. N. Everitt and D. Race. Some Remarks on Linear Ordinary Quasi-Differential Expressions. Proceedings of the London Mathematical Society, volume s3-54(2):pp. 300–320, 1987. doi:10.1112/plms/s3-54.2.300.

[18] W. N. Everitt and A. Zettl. Generalized symmetric ordinary differential expressions. I. The general theory. Nieuw Arch. Wisk. (3), volume 27(3):pp. 363–397, 1979.

[19] W. N. Everitt and A. Zettl. Differential operators generated by a countable number of quasi-differential expressions on the real line. Proc. London Math. Soc. (3), volume 64(3):pp. 524–544, 1992. doi:10.1112/plms/s3-64.3.524.

[20] R. V. Gamkrelidze. Discovery of the maximum principle. J. Dynam. Control Systems, volume 5(4):pp. 437–451, 1999. doi:10.1023/A:1021783020548.

[21] L. Greenberg and M. Marletta. Numerical methods for higher order sturmliouville problems. Journal of Computational and Applied Mathematics, volume 125(1-2):pp. 367 – 383, 2000. doi:10.1016/S0377-0427(00)00480-5. ¡ce:title¿Numerical Analysis 2000. Vol. VI: Ordinary Differential Equations and Integral Equations¡/ce:title¿.

[22] V. I. Kogan and F. S. Rofe-Beketov. On the question of the defect numbers of symmetric differential operators with complex coefficients. Mat. Fiz. i Funkcional. Anal., (Vyp. 2):pp. 45–60, 237, 1971.

[23] I. Lasiecka and R. Triggiani. CONTROL THEORY FOR PARTIAL DIFFERENTIALEQUATIONS. Cambridge University Press, Cambridge, UK, 2000. Published in two volumes.

[24] A. Lewis. Course on Pontryagin's Maximum Principle, 2006. Http://www.mast.queensu.ca/˜andrew/teaching/MP-course/.

[25] P. D. Loewen and R. B. Vinter. Pontryagin-type necessary conditions for differential inclusion problems. Systems & Control Letters, volume 9(3), 1987. doi:DOI:\%2010.1016/0167-6911(87)90049-1.

[26] H. MAURER, REIHE, R. A. PREPRINTS, R. B. BERICHTE, R. C. MATHEMATISCHE, M. SIMULATION, R. D. ELEKTRISCHE, N. BAUELEMENTE, H. J. OBERLE, and H. J. OBERLE. Second Order Sufficient Conditions for Optimal Control Problems with Free Final Time: The Riccati Approach, 2000.

[27] J. B. MCLEOD. The number of integrable-square solutions of ordinary differential equations. The Quarterly Journal of Mathematics, volume 17(1):pp. 285–290, 1966. doi: 10.1093/qmath/17.1.285.

[28] M. MÖLLER. On the unboundedness below of the Sturm-Liouville operator. Proceedings of the Royal Society of Edinburgh, Section: A Mathematics, volume 129(05):pp. 1011–1015, 1999. doi:10.1017/S030821050003105X.

[29] M. MOLLER and A. ZETTL. Symmetrical Differential Operators and Their Friedrichs Extension. Journal of Differential Equations, volume 115(1):pp. 50–69, 1995. doi:10.1006/jdeq.1995.1003.

[30] B. S. MORDUKHOVICH. Discrete Approximations and Refined EulerLagrange Conditions for Nonconvex Differential Inclusions. SIAM Journal on Control and Optimization, volume 33(3):pp. 882–915, 1995. doi:10.1137/S0363012993245665.

[31] B. S. MORDUKHOVICH. VARIATIONAL ANALYSIS AND GENERALIZED DIFFERENTIATION II, volume 331 of A Series of Comprehensive Studies in Mathematics. Springer Berlin Heidelberg, 1 edition, 2005. ISBN 978-3-540-25438-6.

[32] M. A. NAIMARK. LINEAR DIFFERENTIAL OPERATORS. PART I: ELEMENTARY THEORY OF LINEAR DIFFERENTIAL OPERATORS. Frederick Ungar Publishing Co., New York, 1967.

[33] M. A. NAIMARK. LINEAR DIFFERENTIAL OPERATORS: PART II. Frederick Ungar Publishing Co., Inc., New York, 1968.

[34] D. O'REGAN and M. EL-GEBEILY. Existence, upper and lower solutions and quasilinearization for singular differential equations. IMA J. Appl. Math., volume 73(2):pp. 323–344, 2008. doi:10.1093/imamat/hxn001.

[35] L. S. PONTRYAGIN, V. G. BOLTYANSKII, R. V. GAMKRELIDZE, and E. F. MISHCHENKO. THE MATHEMATICAL THEORY OF OPTIMAL PROCESSES. Translated by D. E. Brown. A Pergamon Press Book. The Macmillan Co., New York, 1964.

[36] T. T. READ. Sequences of deficiency indices. Proc. Roy. Soc. Edinburgh Sect. A, volume 74:pp. 157–164 (1976), 1974/75.

[37] R. W. H. SARGENT. Optimal control. J. Comput. Appl. Math., volume 124(1-2):pp. 361–371, 2000. doi:10.1016/S0377-0427(00)00418-0.

[38] A. SEIERSTAD and K. SYDSAETER. Sufficient Conditions in Optimal Control Theory. International Economic Review, volume 18(2):pp. 367–91, 1977.

[39] D. SHIN. On solutions of the system of quasi-differential equations. C. R. (Doklady) Acad. Sci. URSS (N.S.), volume 28:pp. 391–395, 1940.

[40] J. SUN and W. Y. WANG. Characterization of domains of self-adjoint ordinary differential operators and spectral analysis. Neimenggu Daxue Xuebao Ziran Kexue, volume 40(4):pp. 469–485, 2009.

[41] H. J. SUSSMANN and J. C. WILLEMS. 300 years of optimal control: from the brachystochrone to the maximum principle. Control Systems Magazine, IEEE, volume 17(3), 1997. doi:10.1109/37.588098.

[42] R. B. VINTER. OPTIMAL CONTROL. Birkhäuser, Boston, 2000.

[43] A. WANG, J. SUN, and A. ZETTL. The classification of self-adjoint boundary conditions: Separated, coupled, and mixed. Journal of Functional Analysis, volume 255(6):pp. 1554–1573, 2008. doi:10.1016/j.jfa.2008.05.003.

[44] A. WANG, J. SUN, and A. ZETTL. The classification of self-adjoint boundary conditions of differential operators with two singular endpoints. Journal of Mathematical Analysis and Applications, volume 378(2):pp. 493–506, 2011. doi:10.1016/j.jmaa.2011.01.070.

[45] J. WEIDMANN. SPECTRAL THEORY OF ORDINARY DIFFERENTIAL OPERATORS, volume 1258 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1987.

[46] A. ZETTL. Formally self-adjoint quasi-differential operators. Rocky Mountain J. Math., volume 5:pp. 453–474, 1975.

[47] A. ZETTL. Sturm-Liouville problems. In Spectral theory and computational methods of Sturm-Liouville problems (Knoxville, TN, 1996), volume 191 of Lecture Notes in Pure and Appl. Math., pp. 1–104. Dekker, New York, 1997.

[48] A. ZETTL. STURM-LIOUVILLE THEORY, volume 121 of Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2005. ISBN 0-8218-3905-5.

VITA

- Mohammed Mogib Mohammed Alshahrani
- Born in Kazza, Saudi Arabia on August 23, 1971.
- Received B.Sc. (Honours) in Mathematics from Abha Teachers College in 1997.
- Received M.Sc. in Mathematics from King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia in 2003.
- Worked as a Teacher of Mathematics in Abha, Saudi Arabia, 1998-1999.
- Worked as a Graduate Assistance in the Department of Mathematics, Dammam Teachers College, Dammam, 1998-2003.
- Worked as a Lecturer in the Department of Mathematics, Dammam Teachers College, Dammam, 2003-2007.
- Worked as a Lecturer in the Department of General Studies-Mathematics Discipline, Jubail Industrial College, Jubail, 2007-2008.
- Joined as Lecturer the Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, 2008.
- **Present Address:** Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Box # 1258, Dhahran 31261, Saudi Arabia.

- **Office Phone:**+966-3-860-7748.
- **Permanent Address:** Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Box # 1258, Dhahran 31261, Saudi Arabia.
- **Email:** mmogib@gmail.com, mshahrani@kfupm.edu.sa.