Dynamic Systems and Applications 21 (2012) 219-234

MAXIMUM PRINCIPLE IN OPTIMAL CONTROL OF SINGULAR ORDINARY DIFFERENTIAL EQUATIONS IN HILBERT SPACES

M. M. ALSHAHRANI, M. A. EL-GEBEILY, AND B. S. MORDUKHOVICH

Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Box # 1258, Dhahran 31261, Saudi Arabia
mshahrani@kfupm.edu.sa
Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Box # 1777, Dhahran 31261, Saudi Arabia
mgebeily@kfupm.edu.sa
Department of Mathematics, Wayne State University, Detroit, MI 48202, USA and Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia
boris@math.wayne.edu

ABSTRACT. In this paper 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 the Pontryagin Maximum Principle.

1. INTRODUCTION

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

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

where $\tilde{L}$ is a self-adjoint extension of the minimal operator L_0 (see Section 2) generated by a formally self-adjoint differential expression l and a positive weight function w satisfying the equation

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

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

Research of the last author was partly supported by the USA National Science Foundation under grant DMS-1007132, by the Australian Research Council under grant DP-12092508, and by the Portuguese Foundation of Science and Technologies under grant MAT/11109.
Received November 15, 2011

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 (cf. [5]) that the expression l in (1.2) is of even order $2n$ 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 $i = 0, \ldots, n$. Recall that the expression l is *regular* if the I is finite and

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

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 $r_n^{-1}, r_{n-1}, \ldots, r_0 \in L((a, \beta), w)$ 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 (1.1):

$$\text{minimize} \quad J[u, x] = \phi(x(c)) \text{ over } (u, x) \in \mathcal{A}. \tag{1.3}$$

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 (1.1) described by

$$x(t) = \int_I K(t, \tau) f(x(\tau), u(\tau), \tau) d\tau, \quad t \in I; \tag{1.4}$$

see Section 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 $\tilde{D}$; see Section 2, particularly Theorem 2.1). Since no additional constraints are imposed on $x(\cdot)$ at $t = b$, problem (1.3) 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 (1.3). 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., [3, 4] with the vast bibliographies therein. However, we are not familiar with any developments on optimal control of differential operator equations of type (1.1).

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

$$F_u(x) := f(x(\cdot), u(\cdot), \cdot). \tag{1.5}$$

The main goal of this paper is deriving necessary optimality conditions for a fixed optimal solution $(\bar{u}(\cdot), \bar{x}(\cdot))$ to problem (P). Involving this optimal pair and operator (1.5), we impose the following *standing assumptions*:

(H1) F_u maps $\mathcal{H}$ in to $\mathcal{H}$ 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.,

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

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

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

where $\tilde{D}$ is the domain of $\tilde{L}$ to be defined in Section 2.

(H5) For every needle variation u (see Section 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 (1.1) is a Souslin subset (i.e., a continuous image of a Borel subset) of some Banach space.

To formulate the main result of this paper, we introduce the appropriate counterpart of the *Hamilton-Pontryagin function* for system (1.1) defined by

$$H(x, p, u, t) := (p + P(\phi(x(c))K(c, t)))\, f(x, u, t), \tag{1.6}$$

where P is a projection operator onto the range of L_0 to be discussed in Section 2; see particularly Lemma 2.2 therein.

Theorem 1.1 (Maximum Principle). *Let* $(\bar{u}(\cdot), \bar{x}(\cdot))$ *be an optimal solution to problem* (1.3) *under the assumptions imposed in* (H1)–(H7). *Then there exists an adjoint arc* $p \in D$ *such that*

$$H(\bar{x}(t), p(t), \bar{u}(t), t) = \max_{u \in U} H(\bar{x}(t), p(t), u, t) \ ae \ \ t \in I, \tag{1.7}$$

(1.8) $$L(p) = -\nabla_x H(\bar{x}(t), p(t), \bar{u}(t), t) \ \ a.e.$$

and the following transversality condition is satisfied:

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

where D is the domain of the operator L defined in Section 2*), and where the functions $x_i, i = 1, \dots, d$ determine the domain $\tilde{D}$ in the sense of Theorem* 2.1.

The rest of the paper is organized as follows. In Section 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 [5], Weidmann [7], and Zettl [8], [9] among others.

In Section 3 we obtain new existence results for self-adjoint differential operator equations, which play a crucial role in the prove of the Maximum Principle of Theorem 1.1 given in Section 4.

2. SELF-ADJOINT DIFFERENTIAL OPERATOR EQUATIONS

The expression l in (1.2) generates various operators on $\mathcal{H}$. Among these operators we single out lie the *minimal* operator L_0, the *maximal* operator L, and self-adjoint operators $\tilde{L}$ lying between. The maximal operator L is defined by

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

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

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

Denote by $AC_{\text{loc}}(I)$ the set of real-valued functions, which are absolutely continuous on every compact subinterval of I. Let $L_0 := L^*$ with $D_0 := D(L_0)$, where L^* is the adjoint of L uniquely defined due to the fact that D is dense in $\mathcal{H}$. It is shown in [5] that $D_0 \subset D$, that D_0 is dense in $\mathcal{H}$, and that $L_0^* = L$, which implies in turn that L_0 is a symmetric closed operator.

Pick an arbitrary complex number ν with $\text{Im}(\nu) \neq 0$ 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 $\text{cl}\, R_\nu$ 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 [5] that $\mathcal{N}_\nu$ is the eigenspace of L corresponding to the eigenvalue

$\bar{\nu}$ and that D is decomposed as

$$D = D_0 \dotplus \mathcal{N}_\nu \dotplus \mathcal{N}_{\bar{\nu}}.$$

It is also shown in [5] that the equality

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

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

A *self-adjoint realization* of the the equation (1.2) in $\mathcal{H}$ is any linear bounded operator $\tilde{L}$ satisfying the relationships

$$L_0 \subset \tilde{L} = \tilde{L}^* \subset L.$$

These self-adjoint realizations are distinguished from one another by their domains. Naimark [5] established the following decomposition

$$\tilde{D} = D_0 \dotplus \ \mathrm{span}\,\{\phi_1, \phi_2, \ldots, \phi_d\} \tag{2.1}$$

of the domain of $\tilde{L}$ 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 is always a $2d-$dimensional extension of D_0 and that $\tilde{D}$ is a $d-$dimensional extension of D_0. It follows furthermore that D is a $d-$dimensional extension of $\tilde{D}$.

The fundamental Glazman-Krein-Naimark (GKN) Theorem [2] characterizes these domains as follows.

Theorem 2.1 (GKN characterization of domains)**.** *Let $d \in \mathbb{N}$ be the deficiency index of L_0. A linear submanifold $\tilde{D}$ of D is the domain of a self-adjoint extension $\tilde{L}$ of L_0 with deficiency index d if and only if there exist functions $x_1, x_2, \ldots, x_d$ in D 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) *$\tilde{D} = \{x \in D : [x, x_i]_a^b = 0, \quad i = 1, 2, \ldots, d\}$.*

The bracket $[\cdot,\cdot]_a^b$ in Theorem 2.1 is called the *Lagrange bracket* and is defined for any $x, z \in D$ and $t \in I$ by

$$[x, z]\,(t) := \sum_{i=1}^{n} \left\{ x^{[i-1]}(t) z^{[2n-i]}(t) - x^{[2n-i]}(t) z^{[i-1]}(t) \right\}. \tag{2.2}$$

It is worth mentioning that the limits in (2.2) 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*

$$\int_a^b l(x)z\,dt - \int_a^b x\,l(z)\,dt = [x,z]_a^b \quad \text{for any} \quad x,z \in D. \tag{2.3}$$

Recall that the operator $R_\nu := \left(\tilde{L} - \nu E\right)^{-1}$ is known as the *resolvent operator* of $\tilde{L}$ with respect to the complex number ν. It follows from assumption (H4) that the mapping $\tilde{L}$ is one-to-one and equals zero is a regular point of $\tilde{L}$. This implies that the resolvent $R_0 = \tilde{L}^{-1}$ exists as a bounded operator defined on the whole space $\mathcal{H}$. Furthermore, it is an integral operator with the kernel K 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 $y \in \tilde{D}$ we can be write

$$y = R_0 f = \int_I K(\tau,t)f(\tau)w(\tau)d\tau \quad \text{a.e.} \quad t \in I, \tag{2.4}$$

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

Next we define the projection operator P onto the range $\mathcal{R}_0$ of L_0. First observe from the domain decomposition (2.1) that

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

where $\tilde{R}$ is the range of $\tilde{L}$, 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 $\{x_i\}_{i=1}^d \subset \tilde{D}$ be such that $\tilde{L}x_i = z_i$ for $i = 1, \ldots, d$. It is clear that $\{x_i\}_{i=1}^d$ is linearly independent modulo D_0. Finally, define P on H as

$$P(y) := (E - Q)y, \qquad y \in \mathcal{H}, \tag{2.5}$$

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

$$Q(y) = \sum_{i=1}^d \langle y, z_i\rangle z_i, \qquad y \in \mathcal{H}. \tag{2.6}$$

By the fundamental Theorem 2.1, we may assume that

$$\tilde{D} = D_0 \dotplus \operatorname{span}(\{x_1, x_2, \ldots, x_d\}). \tag{2.7}$$

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

$$\tilde{L}x = \tilde{L}x_0 + \sum_{i=1}^{n} \alpha_i \tilde{L}x_i = \tilde{L}x_0 + \sum_{i=1}^{n} \xi_i z_i,$$

$$\tilde{L}x = g = P(g) + Q(g).$$

Both elements $\tilde{L}x_0$ 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 (2.1) is in fact a direct sum, it gives us therefore that

$$\tilde{L}x_0 = P(g) \quad \text{and} \quad \sum_{i=1}^{n} \alpha_i z_i = Q(g).$$

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 (1.6) used in our main result.

Lemma 2.2. *Let $\tilde{L}x = g$ 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:

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

3. EXISTENCE OF SOLUTIONS TO OPERATOR EQUATIONS

In this section we derive new results on the existence of solutions of the primal operator equation (1.1) in the domain $\tilde{D}$ and of the adjoint equation (1.8) in the domain D. Besides of their own independent interest, the results obtained are important for the proof of our main Theorem 1.1 on the Maximum Principle.

We begin with the following lemma, which can be also seen as a consequence of the existence result from [6, Theorem 15]. Although throughout the paper 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 3.1. *Equation* (1.1) *has at least one solution in $\tilde{D}$ 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

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

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

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

Using assumption (H4) on $\tilde{L}$, assumption (H3) on F_u, and the classical Cauchy-Schwartz inequality give us

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

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

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

which completes the proof of the lemma. □

The result of Lemma 3.1 can be treated as the justification of *controllability* of the primal differential operator system (1.1) with measurable controls.

The next lemma plays a crucial role in justifying the existence of solutions to boundary value problem for the adjoint system (1.8), which is the main result of this section; see Theorem 3.3 below.

Lemma 3.2. *Let $h_1 \in \mathcal{H}$ be such that*

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

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* (2.7). *Then for any $h_2 \in \mathcal{H}$ and for arbitrary real numbers $\alpha_i, i = 1, \ldots, d$, the equation*

$$\left.\begin{aligned} (Lx)(t) &= h_1(t)x(t) + h_2(t), \qquad t \in (a,b) \\ [x, x_i]_a^b &= \alpha_i, \qquad i = 1, \ldots, d \end{aligned}\right\} \tag{3.1}$$

admits a solution in the domain $\tilde{D}$.

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

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

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

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

implying by Theorem 2.1 that $\sum_{j=1}^{d} u_j\xi_j \in \tilde{D}$. The latter contradicts the fact that the functions $\xi_j, j = 1, \ldots, d$, are linearly independent modulo $\tilde{D}$.

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

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

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

$$\tilde{L}\tilde{x} = h_1\tilde{x} + \sum_{i=1}^{d} \beta_i \left(h_1\xi_i - L\xi_i\right) + h_2. \tag{3.2}$$

Then we see that the element

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

is certainly a solution to (3.1). It remains to show that equation (3.2) admits a solution in $\tilde{D}$. To proceed, we define the function

$$F(z) := h_1z + h_3 \;\text{ for any }\; z \in \tilde{D},$$

where $h_3 := \sum_{i=1}^{d} \beta_i\left(h_1x_i - Lx_i\right) + h_2$. The function F is obviously weakly continuous, and furthermore we have

$$\begin{aligned}\langle \tilde{L}z - F(z), z\rangle &= \langle \tilde{L}z - h_1z - h_3, z\rangle = \langle \tilde{L}z, z\rangle - \langle h_1z, z\rangle - \langle h_3, z\rangle\\ &> \gamma\|z\|^2 - \eta\|z\|^2 - \|h_3\|\|z\| = (\gamma - \eta)\|z\|^2 - \|h_3\|\|z\|.\end{aligned}$$

This ensures the existence of a solution to (3.1) in $\tilde{D}$ by [6, 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 (1.8), (1.9) in the required domain D.

Theorem 3.3 (existence of solutions to the adjoint system)**.** *The adjoint equation* (1.8) *with the boundary conditions* (1.9) *admits a solution in D.*

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

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

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 3.2. □

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 (1.3) under the standing assumption formulated in Theorem 1.1. 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 (1.1) defined by (2.4), 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 4.1. *In the notation above we have the increment formula*

$$\begin{aligned} \Delta J[\bar{u}] &= -\langle p+P(\check{K}_c(\cdot)), \Delta_u F'_{\bar{u}}(\bar{x})\Delta\bar{x}\rangle - \langle p+P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}(\bar{x})\rangle \\ &\quad + o(\|\Delta\bar{x}\|) + o(|\Delta\bar{x}(c)|), \end{aligned} \tag{4.1}$$

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 (2.5), *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

$$\Delta J[\bar{u}] = \phi(x(c))-\phi(\bar{x}(c)) = \phi'_0(\bar{x}(c))\Delta\bar{x}(c) + o(|\Delta\bar{x}(c)|). \tag{4.2}$$

If $x_i \in D, i = 1, \ldots, d$, are the functions that determine $\tilde{L}$ by Theorem 2.1), then every $x \in \tilde{D}$ 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 $x \in \tilde{D}$ and any $p \in D$ satisfy the primal and adjoint systems (these solutions exist due to Lemma 3.1 and Theorem 3.3, 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

$$\phi_0'(\bar{x}(c)) \Delta \bar{x}(c) = \phi_0'(\bar{x}(c)) \Delta \bar{x}_0(c) - [p, \Delta \bar{x}]_a^b. \tag{4.3}$$

Now using the Lagrange identity (2.3) and elementary transformations implies that

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

Employing further the solution representation (2.4), we get

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

$$
\begin{aligned}
&+ \Delta_u F_{\bar u}(\bar x))(s)w(s)ds + o(\|\Delta\bar x\|) \\
&= \int_a^b \check{K}_c(s)P(F'_{\bar u}(\bar x)\Delta\bar x + \Delta_u F'_{\bar u}(\bar x)\Delta\bar x + \Delta_u F_{\bar u}(\bar x))(s)w(s)ds + o(\|\Delta\bar x\|) \\
&= \langle \check{K}_c(\cdot), P(F'\bar u(\bar x)\Delta\bar x)\rangle \\
&\quad + \langle \check{K}_c(\cdot), P(\Delta_u F'_{\bar u}(\bar x)\Delta\bar x)\rangle + \langle \check{K}_c(\cdot), P(\Delta_u F_{\bar u}(\bar x))\rangle + o(\|\Delta\bar x\|) \\
&= \langle F'_{\bar u}(\bar x)P(\check{K}_c(\cdot)), \Delta\bar x)\rangle \\
&\quad + \langle P(\check{K}_c(\cdot)), \Delta_u F'_{\bar u}(\bar x)\Delta\bar x\rangle + \langle P(\check{K}_c(\cdot)), \Delta_u F_{\bar u}(\bar x)\rangle + o(\|\Delta\bar x\|).
\end{aligned}
$$

Substituting the obtained expressions for $[p,\Delta\bar x]_a^b$ and $\phi_0'(\bar x(c))\Delta\bar x_0(c)$ into (4.3) yields

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

Taking finally formula (4.2) into account, we arrive at

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

and thus complete the proof of the lemma. □

Note that the derivation of the increment formula in Lemma 4.1 is different from the usual way know in control theory (compare, i.e., [4, 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:

$$
u(t) = \begin{cases} v(t), & t\in I_\epsilon, \\ \bar u(t), & t\notin I_\epsilon. \end{cases} \tag{4.4}
$$

Lemma 4.2. *Let $\Delta\bar x = \Delta\bar x(\cdot)$ be the increment of $\bar x(\cdot)$ corresponding to the needle variation (4.4) of $\bar u(\cdot)$. Then we have the functional trajectory increment estimate*

$$
\|\Delta\bar x\| = o(\epsilon). \tag{4.5}
$$

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

$$
\gamma\|\Delta\bar x\|^2 \;\le\; \langle \tilde L\Delta\bar x, \Delta\bar x\rangle
$$

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

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 (4.5). □

Lemma 4.3. *The following pointwise trajectory increment estimate holds:*

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

Proof. By using the pointwise representation of the trajectory (1.4) corresponding to the needle variation $\bar{u}(\cdot)$, we have

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

The second term of the above inequality can be split into

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

Using further the assumed continuity of $F'_u(\bar{x})$ and Lemma 4.2 ensure the estimates

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

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

which show in turn that

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

and thus justify our claim. □

Lemmas 4.2 and 4.3 enable us to rewrite the increment formula (4.1) of Lemma 4.1 as

$$\Delta J[\bar{u}] = -\langle p + P(\check{K}_c(\cdot)), \Delta_u F_{\bar{u}}(\bar{x})\rangle - \langle p + P(\check{K}_c(\cdot)), \Delta_u F'_{\bar{u}}(\bar{x})\Delta\bar{x}\rangle + o(\epsilon). \tag{4.6}$$

Now all the ingredients required for the justification of the Maximum Principle in Theorem 1.1 (namely, Lemmas 4.1, 4.2, and 4.3) 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 (1.3), and let p be the corresponding solution to the adjoint system (1.8) satisfying the boundary/transversality conditions (1.9). Let us show that the maximum condition (1.7) 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 [4, 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

$$\Delta_v H(t) := H(\bar{x}(t), p(t), v(t), t) - H(\bar{x}, p(t), \bar{u}(t), t) > 0, \quad t \in T. \tag{4.7}$$

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 (4.4) 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 (4.6) corresponding to $\bar{u}$ and u gives us

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

Assumption (H1) and Lemma 4.2 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

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

which implies therefore that

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

This shows by (4.7) 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 (1.3) and thus completes the proof of Theorem 1.1. □

REFERENCES

[1] N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Spaces, Vol. II*, Frederic Ungar Publishing Co., New York, 1963.

[2] W. N. Everitt and L. Markus, *Boundary Value Problems and Symplectic Algebra for Ordinary Differential and Quasi-differential Operators*, Mathematical Surveys and Monographs Series, Vol. 61, American Mathematical Society, Providence, RI, 1999.

[3] I. Lasiecka and R. Triggiani, *Control Theory for Partial Differential Equations*, published in two volumes, Cambridge University Press, Cambridge, UK, 2000.

[4] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, II: Applications*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 331, Springer, Berlin, 2006.

[5] M. A. Neimark, *Linear Differential Operators, Part II*, Frederic Ungar Publishing Co., New York, 1968.

[6] D. O'Regan and M. El-Gebeily, Existence, upper and lower solutions, and quasilinearization for singular differential equations, *IMA J. Appl. Math.* **73** (2008), 323–344.

[7] J. Weidmann, *Spectral Theory of Ordinary Differential Operators*, Lecture Notes in Mathematics, Vol. 1258, Springer, Berlin, 1987.

[8] A. Zettl, Formally self-adjoint quasi-differential operators, *Rocky Mountains J. Math.* **5** (1975), 453–474.

[9] A. Zettl, Sturm-Liouville problems, In: *Spectral Theory and Computational Methods of Sturm-Liouville Problems*, Lecture Notes in Pure and Applied Mathematics, pp. 1–104, Marcel Dekker, New York, 1997.