Second-order Necessary Optimality Conditions for an Optimal Control Problem

Toan Nguyen* and Thuy Le

Abstract. Second-order necessary optimality conditions for an optimal control problem with a nonconvex cost function and state-control constraints are studied in this paper. By establishing an abstract result on second-order necessary optimality conditions for a mathematical programming problem, we obtain second-order necessary optimality conditions for an optimal control problem.

1. Introduction

A wide variety of problems in optimal control problem can be posed in the following form.

Determine a control vector $u \in L^p([0,1], \mathbb{R}^m)$ and a trajectory

$$x \in W^{1,p}([0,1], \mathbb{R}^n), \quad 1 < p < \infty,$$

which minimize the cost

$$(1.1) \qquad J(x, u) = \int_0^1 L(t, x(t), u(t))\, dt,$$

with the state equation

$$(1.2) \qquad \dot{x}(t) = A(t)x(t) + B(t)u(t) \quad \text{a.e. } t \in [0, 1],$$

the initial condition

$$(1.3) \qquad x(0) = c,$$

and the mixed constraint

$$(1.4) \qquad g(t, x(t)) + bu(t) \leq 0 \quad \text{a.e. } t \in [0, 1].$$

Here, $W^{1,p}([0,1],\mathbb{R}^n)$ is the Sobolev space, which consists of absolutely continuous functions $x\colon [0,1] \to \mathbb{R}^n$ such that $\dot{x} \in L^p([0,1],\mathbb{R}^n)$. Its norm is given by

$$\|x\|_{1,p} = |x(0)| + \|\dot{x}\|_p.$$

The notations in (1.1)–(1.4) have the following meanings:

    – x, u are the state variable and the control variable, respectively,

    – $L\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ and $g\colon [0,1] \times \mathbb{R}^n \to \mathbb{R}^m$ are given functions,

    – $A(t) = (a_{kl}(t))_{n\times n}$ and $B(t) = (b_{kl}(t))_{n\times m}$ are matrix-valued functions,

    – $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^*$ are constants.

    Put

$$X = W^{1,p}([0,1],\mathbb{R}^n), \quad U = L^p([0,1],\mathbb{R}^m), \quad Z = X \times U,$$

and

(1.5) $\qquad\qquad E = \{z = (x,u) \in X \times U : (1.2) \text{ and } (1.3) \text{ are satisfied}\}.$

We define a mapping

$$G(z) = G(x,u) = g(\,\cdot\,,x) + bu.$$

When $g(\,\cdot\,,x(\cdot)) + bu(\cdot) \in C([0,1])$ by a traditional approach, we can define a mapping

$$\Phi(x,u) = \max_{t\in[0,1]} \{g(t,x(t)) + bu(t)\}$$

and reduce the problem (1.1)–(1.4) to the mathematical programming problem

$$\text{Minimize } J(x,u) \text{ subject to } (x,u) \in E, \ \Phi(x,u) \le 0.$$

Unfortunately, $g(\,\cdot\,,x(\cdot)) + bu(\cdot)$ belongs to $L^p([0,1],\mathbb{R}^m)$ in general, we fail to define Φ as above. Therefore, we need to find a suitable model for the problem (1.1)–(1.4). Let us define a closed convex cone

$$Q = \{v \in L^p([0,1],\mathbb{R}^m) : v(t) \le 0 \text{ a.e. } t \in [0,1]\}.$$

Then, the problem (1.1)–(1.4) becomes

$$\text{Minimize } J(x,u) \text{ subject to } (x,u) \in E, \ G(x,u) \in Q.$$

Note that E is a closed convex set and Q is a nonempty closed convex cone. By the assumption (H2), we also have $E \ne \emptyset$ (see Lemma 5.2).

This type of problems is considered and investigated in $[7,8,10,22,31,38\text{–}40,42,44]$ and the references therein.

The study of optimality conditions is an important topic in variational analysis and optimization. In order to give a general idea of such optimality conditions, we consider

for the moment the simplest case, when optimization problem is unconstrained. Then, stationary points are given by the first-order optimality condition. It is well known that the second-order necessary condition for stationary points to be locally optimal is that the Hessian matrix should be positive semidefinite. There have been many papers dealing with the first-order optimality condition and second-order necessary condition for mathematical programming problems; see, for example, [2, 3, 5, 9, 17, 19, 32, 36].

By considering a set of assumptions involving different kinds of the critical direction and the Mangasarian-Fromovitz condition, Kawasaki [19] derived second-order optimality conditions for a mathematical programming problem. However, the results of Kawasaki cannot be applied for nonconical constraints. In [9], Cominetti extended the results of Kawasaki. He gave second-order necessary optimality conditions involving Kuhn-Tucker-Lagrange multipliers for optimization problem with geometrical and functional constraints. The novelty of this result with respect to the classical positive semidefiniteness condition on the Hessian of the Lagrangian function is that it contains an extra term, which represents a kind of "second-order derivative" associated with the target set of the functional constraints of the problem.

Besides the study of optimality conditions in mathematical programming, the study of optimality conditions in optimal control is also of interest. There have been many papers dealing with the optimality condition for optimal control; see, for example, [4, 6, 11, 12, 14–16, 20, 23, 24, 26–28, 32–35, 41, 43, 45]. We refer the reader to [4, 11, 12, 16, 20, 23, 32–35] for recent studies on first-order necessary optimality conditions and second-order necessary optimality conditions for continuous optimal control problems. To deal with second-order optimality conditions for the problem (1.1)–(1.4), one often require that Q has a nonempty interior. When $Q \subset L^p([0,1], \mathbb{R}^m)$ with $1 \leq p < \infty$ then its interior is likely empty (see [21, Example 1.1]). Hence, in order to have that Q is of nonempty interior, one must requires that $Q \subset L^\infty([0,1], \mathbb{R}^m)$ and so we need control variable u belonging to $L^\infty([0,1], \mathbb{R}^m)$. By this approach, Páles and Zeidan in [32–35] obtained second-order necessary optimality conditions for the optimal control problem (1.1)–(1.4).

Note that if we consider the problem under the assumption that control variable $u \in L^\infty([0,1], \mathbb{R}^m)$ then, when the constraint sets are unbounded, the problem (1.1)–(1.4) generally has no solutions in $W^{1,p}([0,1], \mathbb{R}^n) \times L^\infty([0,1], \mathbb{R}^m)$ because the objective function does not satisfy coercivity conditions. Besides, multipliers belong to the dual space $L^\infty([0,1], \mathbb{R}^m)^*$, which are measures rather than functions. Therefore, it had better we require $(x, u) \in W^{1,p}([0,1], \mathbb{R}^n) \times L^p([0,1], \mathbb{R}^m)$ with $1 < p < \infty$. For this we first establish the second-order optimality conditions for a mathematical programming problem. We then derive the second-order optimality conditions for an optimal control problem. This approach allows us to obtain necessary optimality conditions when the interior of Q

is empty.

The paper is organized as follows. In Section 2, we state our main result (Theorem 2.3). Section 3 gives basic definitions and preliminaries which will be used in the sequel. To prove the main result, we first derive necessary optimality conditions for a mathematical programming problem (Theorem 4.3), which is given in Section 4, and then apply the obtained results for the problem (1.1)–(1.4). The proof of the main result is given in Section 5. In Section 6, we give some examples to show that if the second-order necessary condition is not satisfied, then the admissible couple is not a solution even it satisfies first-order necessary conditions.

2. Statement of the main result

Recall that a couple $(\overline{x}, \overline{u})$ is said to be *admissible* if it satisfies constraints (1.2)–(1.4). For a given admissible couple $(\overline{x}, \overline{u})$, the symbols

$$\overline{L}(t), \ \overline{L}_x(t), \ \overline{L}_{xx}(t), \ \text{etc.},$$

stand, respectively, for

$$L(t, \overline{x}(t), \overline{u}(t)), \ L_x(t, \overline{x}(t), \overline{u}(t)), \ L_{xx}(t, \overline{x}(t), \overline{u}(t)), \ \text{etc.}$$

An admissible couple $(\overline{x}, \overline{u})$ is said to be a *locally optimal solution* of the problem (1.1)–(1.4) if and only if there exists $\epsilon > 0$ such that for all admissible couples (x, u), the following implication holds:

$$\|(x, u) - (\overline{x}, \overline{u})\|_Z \leq \epsilon \quad \Longrightarrow \quad f(x, u) \geq f(\overline{x}, \overline{u}).$$

To deal with our problem, we impose the following assumptions:

(H1) The function $L \colon [0, 1] \times \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ has properties that $L(\,\cdot\,, x, u)$ is measurable for all $(x, u) \in \mathbb{R}^n \times \mathbb{R}^m$, $L(t, \cdot, \cdot)$ is a function of class C^2 for almost every $t \in [0, 1]$, $L(t, 0, 0) \in L^1([0, 1], \mathbb{R})$ and for each $M > 0$, there exist a positive number k_{LM} and a non-negative function $w_{LM} \in L^\infty([0, 1], \mathbb{R})$ such that

$$|L_x(t, x, u)| + |L_u(t, x, u)| \leq k_{LM}(|x| + |u|^{p-1}) + w_{LM}(t),$$

$$|L_x(t, x, u) - L_x(t, x', u')| \leq k_{LM}(|x - x'| + |u - u'|),$$

$$|L_u(t, x, u) - L_u(t, x, u')| \leq k_{LM} \sum_{j=0}^{p-1-j>0} |u - u'|^{p-1-j}|u'|^j$$

for all $x, x' \in \mathbb{R}^n$ satisfying $|x|, |x'| \leq M$ and $u, u' \in \mathbb{R}^m$. Also for each $M > 0$, there exists a constant $k_{LM} > 0$ such that

$$|L_{xx}(t, x, u) - L_{xx}(t, x', u')| \leq k_{LM}(|x - x'| + |u - u'|),$$

$$|L_{xu}(t, x, u) - L_{xu}(t, x', u')| \leq k_{LM}(|x - x'| + \epsilon|u - u'|^{p-1})$$

(with $\epsilon = 0$ if $1 < p \leq 2$ and $\epsilon = 1$ if $p > 2$) and

$$|L_{uu}(t,x,u) - L_{uu}(t,x,u')| \begin{cases} = 0 & \text{if } 1 < p \leq 2, \\ \leq k_{LM} \sum_{j=0}^{p-2-j>0} |u - u'|^{p-2-j} |u'|^j & \text{if } p > 2 \end{cases}$$

for a.e. $t \in [0,1]$, for all $x, x' \in \mathbb{R}^n$ with $|x|, |x'| \leq M$ and $u, u' \in \mathbb{R}^m$.

(H2) The matrix-valued functions $A \colon [0,1] \to M_{n,n}(\mathbb{R})$ and $B \colon [0,1] \to M_{n,m}(\mathbb{R})$ are measurable and essentially bounded.

(H3) The function $g \colon [0,1] \times \mathbb{R}^n \to \mathbb{R}^m$ is continuous and of class C^2 with respect to the second variable, and satisfies the following properties: $g(\cdot, 0) \in L^p([0,1], \mathbb{R}^m)$ and for each $M > 0$, there exists a constant $l_{gM} > 0$ such that

$$|g_x(t,x)| + |g_{xx}(t,x)| \leq l_{gM},$$
$$|g_x(t,x) - g_x(t,x')| + |g_{xx}(t,x) - g_{xx}(t,x')| \leq l_{gM}|x - x'|$$

for a.e. $t \in [0,1]$, for all $x, x' \in \mathbb{R}^n$ with $|x|, |x'| \leq M$.

(H4) For each $v \in Q$, there exists $u \in L^p([0,1], \mathbb{R}^m)$ such that

$$\overline{g}_x(\cdot)(x - \overline{x}) + bu + \overline{g}(\cdot) - v \in Q,$$

where x and u are satisfied the state equation (1.2) and the initial condition (1.3).

The assumptions (H1) and (H3) guarantee that $J(x,u)$ and $G(x,u)$ are second-order Fréchet differentiable on $W^{1,p}([0,1], \mathbb{R}^n) \times L^p([0,1], \mathbb{R}^m)$ (see [21, Proposition A] or Lemma 5.3 and Lemma 5.4), where the last condition in (H1) ensures that the second-order Taylor expansion is valid. Meanwhile, (H4) guarantees that the Robinson condition is satisfied (see Lemma 5.5). In contrast with the case where the control variable $u \in L^\infty([0,1], \mathbb{R}^m)$, the objective function $J(x,u)$ is difficult to be second-order Fréchet differentiable on $W^{1,p}([0,1], \mathbb{R}^n) \times L^p([0,1], \mathbb{R}^m)$. However, when $L(t,x,u)$ is a polynomial of variables x and u then the assumption (H1) is easily fulfilled. Let us give some illustrative examples to show that $L(t,x,u)$ satisfies (H1).

Example 2.1. Let $p = 11/5$ and $L \colon [0,1] \times \mathbb{R}^2 \times \mathbb{R}^2 \to \overline{\mathbb{R}}$ is defined by

$$L(t,x,u) = x_1^2 + u_1^{11/5} + u_2^{11/5}, \quad \forall x = (x_1, x_2) \in \mathbb{R}^2, \; u = (u_1, u_2) \in \mathbb{R}^2.$$

Then, $L(t,x,u)$ satisfies the assumption (H1).

Indeed, we have

$$L_x(t,x,u) = (2x_1, 0), \quad L_u(t,x,u) = \frac{11}{5}(u_1^{6/5}, u_2^{6/5}), \quad L_{xu}(t,x,u) = L_{ux}(t,x,u) = 0$$

and

$$L_{xx}(t, x, u) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad L_{uu}(t, x, u) = \frac{66}{25} \begin{bmatrix} u_1^{1/5} & 0 \\ 0 & u_2^{1/5} \end{bmatrix}.$$

So,

$$|L_x(t, x, u)| = 2|x_1| \le 2(|x| + |u|^{6/5}),$$

$$|L_x(t, x, u) - L_x(t, x', u')| = 2|x_1 - x_1'| \le 2(|x - x'| + |u - u'|).$$

Since

(2.1) $$a^m + b^m \le (a + b)^m, \quad \forall a, b \ge 0, \ m \ge 1,$$

we get

$$|L_u(t, x, u)| = \frac{11}{5} \sqrt{(u_1^2)^{6/5} + (u_2^2)^{6/5}} \le \frac{11}{5} |u|^{6/5} \le \frac{11}{5} (|u|^{6/5} + |x|).$$

We now show that there is a constant $C > 0$ such that

(2.2) $$|L_u(t, x, u) - L_u(t, x, u')| \le C(|u - u'|^{6/5} + |u - u'|^{1/5}|u'|).$$

We have

$$|L_u(t, x, u) - L_u(t, x, u')| = \frac{11}{5} \sqrt{(u_1^{6/5} - (u_1')^{6/5})^2 + (u_2^{6/5} - (u_2')^{6/5})^2}.$$

We will prove that there exists $C_1 > 0$ such that

(2.3) $$\left| \sqrt[5]{u_1^6} - \sqrt[5]{(u_1')^6} \right| \le C_1 \left[\sqrt[5]{|u_1 - u_1'|^6} + \sqrt[5]{|u_1 - u_1'||u_1'|} \right].$$

Put $u_1' = a$, $u_1 = a + h$, then (2.3) is equivalent to

(2.4) $$\left| \sqrt[5]{(a + h)^6} - \sqrt[5]{a^6} \right| \le C_1 \left[\sqrt[5]{h^6} + \sqrt[5]{|h||a|} \right].$$

Since

$$a^5 - b^5 = (a - b)(a^4 + a^3 b + a^2 b^2 + a b^3 + b^4)$$
$$= (a - b)\big((a^2 - b^2)^2 + \big(\sqrt{a^3 b} + \sqrt{a b^3}\big)^2 + a^2 b^2\big), \quad \forall a, b \in \mathbb{R},$$

we have

$$|a^3 b^2 - a^2 b^3| = |a^2 b^2 (a - b)| \le |a^5 - b^5|,$$
$$|a^4 b - a b^4| = |ab(a^3 - b^3)| = |(a - b)(a^3 b + a^2 b^2 + a b^3)| \le |a^5 - b^5|.$$

So,

(2.5) $$|a - b|^5 \le |a^5 - b^5| + 5|a^4 b - a b^4| + 10|a^3 b^2 - a^2 b^3| \le 16|a^5 - b^5|.$$

This implies that

$$\left| \sqrt[5]{(a+h)^6} - \sqrt[5]{a^6} \right|^5 \leq 16|(a+h)^6 - a^6|$$

$$\leq 16\big[|h|^6 + 6|h|^5|a| + 15|h|^4|a|^2 + 20|h|^3|a|^3 + 15|h|^2|a|^4 + 6|h||a|^5\big]$$

$$\leq 320\big[|h|^6 + |h|^5|a| + |h|^4|a|^2 + |h|^3|a|^3 + |h|^2|a|^4 + |h||a|^5\big]$$

$$\leq 320\big[\sqrt[5]{|h|^6} + \sqrt[5]{|h|}|a|\big]^5.$$

Therefore, the inequality (2.4) is valid with $C_1 = \sqrt[5]{320}$, and so is (2.3). Thus,

(2.6)

$$|L_u(t,x,u) - L_u(t,x,u')|$$

$$= \frac{11}{5}\sqrt{(u_1^{6/5} - (u_1')^{6/5})^2 + (u_2^{6/5} - (u_2')^{6/5})^2}$$

$$\leq \frac{11}{5}\sqrt{C_1}\sqrt{\left(\sqrt[5]{|u_1 - u_1'|^6} + \sqrt[5]{|u_1 - u_1'|}|u_1'|\right)^2 + \left(\sqrt[5]{|u_2 - u_2'|^6} + \sqrt[5]{|u_2 - u_2'|}|u_2'|\right)^2}$$

$$\leq \frac{11}{5}\sqrt{2C_1}\sqrt{\left(\sqrt[5]{|u_1 - u_1'|^6}\right)^2 + \left(\sqrt[5]{|u_1 - u_1'|}|u_1'|\right)^2 + \left(\sqrt[5]{|u_2 - u_2'|^6}\right)^2 + \left(\sqrt[5]{|u_2 - u_2'|}|u_2'|\right)^2}$$

$$\leq \frac{11}{5}\sqrt{2C_1}\sqrt{\left(\sqrt[5]{|u_1 - u_1'|^6}\right)^2 + \left(\sqrt[5]{|u_2 - u_2'|^6}\right)^2}$$

$$+ \sqrt{\left(\sqrt[5]{|u_1 - u_1'|}|u_1'|\right)^2 + \left(\sqrt[5]{|u_2 - u_2'|}|u_2'|\right)^2}.$$

Since (2.1) holds, we have

(2.7)

$$\sqrt{\left(\sqrt[5]{|u_1 - u_1'|^6}\right)^2 + \left(\sqrt[5]{|u_2 - u_2'|^6}\right)^2} \leq \left(\sqrt{|u_1 - u_1'|^2 + |u_2 - u_2'|^2}\right)^{6/5}$$

$$= |u - u'|^{6/5}.$$

From the inequalities

$$a^8 c^2 \leq (a^2 + c^2)^5, \quad a^6 c^4 \leq (a^2 + c^2)^5,$$

we have

$$5b^8 d^2 (a^{8/5} c^{2/5}) \leq 5b^8 d^2 (a^2 + c^2), \quad 10b^6 d^4 (a^{6/5} c^{4/5}) \leq 10b^6 d^4 (a^2 + c^2).$$

So,

$$(a^{2/5} b^2 + c^{2/5} d^2)^5 \leq (a^2 + c^2)(b^2 + d^2)^5.$$

Hence,

(2.8)

$$\sqrt{\left(\sqrt[5]{|u_1 - u_1'|}|u_1'|\right)^2 + \left(\sqrt[5]{|u_2 - u_2'|}|u_2'|\right)^2}$$

$$\leq \left(\sqrt{|u_1 - u_1'|^2 + |u_2 - u_2'|^2}\right)^{1/5} \sqrt{(u_1')^2 + (u_2')^2} = |u - u'|^{1/5}|u'|.$$

Combining (2.6), (2.7) and (2.8), we obtain (2.2) with $C = \frac{11}{5}\sqrt{2C_1}$. We also have

$$|L_{xx}(t, x, u) - L_{xx}(t, x', u')| = 0, \quad |L_{xu}(t, x, u) - L_{xu}(t, x', u')| = 0$$

and

$$|L_{uu}(t, x, u) - L_{uu}(t, x, u')| \leq \frac{66}{25}\big(|u_1^{1/5} - (u_1')^{1/5}| + |u_2^{1/5} - (u_2')^{1/5}|\big).$$

Since $f(y) = y^{1/5}$ is Hölder continuous with order $\delta = 1/5$, there exists a constant k such that

$$|L_{uu}(t, x, u) - L_{uu}(t, x, u')| \leq \frac{66\sqrt{2}}{25}\sqrt{|u_1^{1/5} - (u_1')^{1/5}|^2 + |u_2^{1/5} - (u_2')^{1/5}|^2}$$

$$\leq \frac{66\sqrt{2}k}{25}\sqrt{|u_1 - u_1'|^{2/5} + |u_2 - (u_2')|^{2/5}}.$$

From (2.5), we have

$$|L_{uu}(t, x, u) - L_{uu}(t, x, u')| \leq \frac{66\sqrt{2}k\sqrt[10]{16}}{25}\left(\sqrt{|u_1 - u_1'|^2 + |u_2 - (u_2')|^2}\right)^{1/5}$$

$$= \frac{66\sqrt{2}k\sqrt[10]{16}}{25}|u - u'|^{1/5}.$$

Thus, the function $L(t, x, u)$ satisfies the assumption (H1).

Example 2.2. Let $p = 3$ and $L : [0, 1] \times \mathbb{R}^2 \times \mathbb{R}^2 \to \overline{\mathbb{R}}$ is defined by

$$L(t, x, u) = u_1^3 - \sqrt{(1 + u_1^6)} + u_2^3, \quad \forall\, x = (x_1, x_2) \in \mathbb{R}^2,\ u = (u_1, u_2) \in \mathbb{R}^2.$$

Then, $L(t, x, u)$ also satisfies the assumption (H1).

Indeed, we have

$$L_u(t, x, u) = 3\left(u_1^2 - \frac{u_1^5}{\sqrt{1 + u_1^6}}, u_2^2\right),$$

$$L_x(t, x, u) = 0, \quad L_{xu}(t, x, u) = L_{ux}(t, x, u) = L_{xx}(t, x, u) = 0$$

and

$$L_{uu}(t, x, u) = 3\begin{bmatrix} 2u_1 - \frac{2u_1^{10} + 5u_1^4}{(1 + u_1^6)\sqrt{1 + u_1^6}} & 0 \\ 0 & 2u_2 \end{bmatrix}.$$

So,

$$|L_u(t, x, u)| \leq 3\sqrt{\left(u_1^2 - \frac{u_1^5}{\sqrt{1 + u_1^6}}\right)^2 + u_2^4} \leq 3\sqrt{2}\sqrt{u_1^4 + u_2^4 + \frac{u_1^{10}}{1 + u_1^6}}$$

$$\leq 6\sqrt{u_1^4 + u_2^4} = 6|u|^2 \leq 6(|u|^2 + |x|).$$

We also have

$$|L_u(t, x, u) - L_u(t, x, u')|$$

$$= 3\sqrt{\left(u_1^2 - (u_1')^2 - \frac{u_1^5}{\sqrt{1 + u_1^6}} + \frac{(u_1')^5}{\sqrt{1 + (u_1')^6}}\right)^2 + (u_2^2 - (u_2')^2)^2}$$

$$\leq 3\left(|u_1^2 - (u_1')^2| + \left|\frac{u_1^5}{\sqrt{1 + u_1^6}} - \frac{(u_1')^5}{\sqrt{1 + (u_1')^6}}\right| + |u_2^2 - (u_2')^2|\right).$$

Put

(2.9)
$$g(y) = \frac{2y^{10} + 5y^4}{(1 + y^6)\sqrt{1 + y^6}}.$$

By Lagrange theorem and

$$|g(y)| = \left|\frac{2y^{10} + 5y^4}{(1 + y^6)\sqrt{1 + y^6}}\right| \leq \left|\frac{2y^{10}}{(1 + y^6)\sqrt{1 + y^6}}\right| + \left|\frac{5y^4}{(1 + y^6)\sqrt{1 + y^6}}\right|$$

$$\leq \left|\frac{2y^{10}}{y^9}\right| + \left|\frac{5y^4}{y^3}\right| = 7|y|,$$

there exist $\theta_1, \theta_2 \in (0, 1)$ such that

$$|u_1^2 - (u_1')^2| \leq 2|u_1' + \theta_1(u_1 - u_1')||u_1 - u_1'| \leq 2(|u_1'||u_1 - u_1'| + |u_1 - u_1'|^2),$$
$$|u_2^2 - (u_2')^2| \leq 2(|u_2'||u_2 - u_2'| + |u_2 - u_2'|^2)$$

and

$$\left|\frac{u_1^5}{\sqrt{1 + u_1^6}} - \frac{(u_1')^5}{\sqrt{1 + (u_1')^6}}\right| = |g(u_1' + \theta_2(u_1 - u_1'))||u_1 - u_1'|$$

$$\leq 7|u_1' + \theta_2(u_1 - u_1')||u_1 - u_1'|$$

$$\leq 7|u_1 - u_1'||u_1'| + |u_1 - u_1'|^2.$$

Hence,

$$|L_u(t, x, u) - L_u(t, x, u')|$$

$$= 3\left(|u_1^2 - (u_1')^2| + |u_2^2 - (u_2')^2| + \left|\frac{u_1^5}{\sqrt{1 + u_1^6}} - \frac{(u_1')^5}{\sqrt{1 + (u_1')^6}}\right|\right)$$

$$\leq 27\left(|u_1 - u_1'||u_1'| + |u_1 - u_1'|^2 + |u_2 - u_2'||u_2'| + |u_2 - u_2'|^2\right)$$

$$= 27\left(|u_1 - u_1'||u_1'| + |u_2 - u_2'||u_2'| + (|u_1 - u_1'|^2 + |u_2 - u_2'|^2)\right)$$

$$\leq 27\left(\sqrt{|u_1 - u_1'|^2 + |u_2 - u_2'|^2}\sqrt{(u_1')^2 + (u_2')^2} + (|u_1 - u_1'|^2 + |u_2 - u_2'|^2)\right)$$

$$= 27\left(|u - u'||u'| + |u - u'|^2\right).$$

From (2.9), we have

$$|g'(y)| = \left| \frac{(20y^9 + 20y^3)(1 + y^6)^{3/2} - 9y^5\sqrt{1 + y^6}(2y^{10} + 5y^4)}{(1 + y^6)^3} \right|$$

$$\leq 20\frac{|y|^9}{(1 + y^6)\sqrt{1 + y^6}} + 20\frac{|y|^3}{(1 + y^6)\sqrt{1 + y^6}}$$

$$+ 18\frac{|y|^{15}}{(1 + y^6)^2\sqrt{1 + y^6}} + 45\frac{|y|^9}{(1 + y^6)^2\sqrt{1 + y^6}}$$

$$\leq 103.$$

Also by Lagrange theorem, there exists $\theta_3 \in (0, 1)$ such that

$$|L_{uu}(t, x, u) - L_{uu}(t, x, u')|$$

$$\leq 3\left| 2(u_1 - u_1') - \frac{2u_1^{10} + 5u_1^4}{(1 + u_1^6)\sqrt{1 + u_1^6}} + \frac{2(u_1')^{10} + 5(u_1')^4}{(1 + (u_1')^6)\sqrt{1 + (u_1')^6}} \right| + 2|u_2 - u_2'|$$

$$\leq 6(|u_1 - u_1'| + |u_2 - u_2'|) + 3\left| \frac{2u_1^{10} + 5u_1^4}{(1 + u_1^6)\sqrt{1 + u_1^6}} - \frac{2(u_1')^{10} + 5(u_1')^4}{(1 + (u_1')^6)\sqrt{1 + (u_1')^6}} \right|$$

$$\leq 6\sqrt{2}\sqrt{|u_1 - u_1'|^2 + |u_2 - u_2'|^2} + 3|g'(u_1' + \theta_3(u_1 - u_1'))||u_1 - u_1'|$$

$$\leq 6\sqrt{2}\sqrt{|u_1 - u_1'|^2 + |u_2 - u_2'|^2} + 309|u_1 - u_1'|$$

$$\leq 618\sqrt{|u_1 - u_1'|^2 + |u_2 - u_2'|^2} = 618|u - u'|.$$

Thus, the function $L(t, x, u)$ also satisfies the assumption (H1).

A pair $z = (x, u) \in X \times U$ is said to be a *critical direction* for the problem (1.1)–(1.4) at $\overline{z} = (\overline{x}, \overline{u})$, if and only if the following conditions hold:

(C1) $\displaystyle\int_0^1 \left[\int_t^1 L_x(s, \overline{x}(s), \overline{u}(s))\, ds \right] \dot{x}(t)\, dt + \int_0^1 L_u(t, \overline{x}(t), \overline{u}(t))u(t)\, dt = 0;$

(C2) $\begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t) \quad \text{a.e. } t \in [0, 1], \\ x(0) = 0; \end{cases}$

(C3) $\overline{g}_x(t)x(t) + bu(t) \in T((-\infty, 0]; \overline{g}(t) + b\overline{u}(t))$ a.e. $t \in [0, 1]$.

We denote by $\Theta(\overline{x}, \overline{u})$ the set of all critical directions to the problem (1.1)–(1.4) at $(\overline{x}, \overline{u})$. Clearly, $\Theta(\overline{x}, \overline{u})$ is a convex cone, which contains $(0, 0)$.

We now state our main result.

Theorem 2.3. *Suppose that $(\overline{x}, \overline{u})$ is a locally optimal solution of the problem (1.1)– (1.4) and the assumptions (H1)–(H4) are satisfied. Then, there exist an unique $w^* \in L^q([0, 1], \mathbb{R}^m)$, and an unique $v \in W^{1,q}([0, 1], \mathbb{R}^n)$, $v(1) = 0$ such that the following conditions are fulfilled:*

(a) *Adjoint equation:*

$$\overline{L}_x(t) = \dot{v}(t) + A^T(t)v(t) - w^*(t)\overline{g}_x(t) \quad a.e. \ t \in [0,1];$$

(b) *Stationarity condition with respect to u:*

$$\overline{L}_u(t) - B^T(t)v(t) + bw^*(t) = 0 \quad a.e. \ t \in [0,1];$$

(c) *Non-negative second-order condition:*

$$\int_0^1 \left[\overline{L}_{xx}(t)x^2(t) + 2\overline{L}_{xu}(t)x(t)u(t) + \overline{L}_{uu}(t)u^2(t)\right] dt + \int_0^1 w^*(t)\overline{g}_{xx}(t)x^2(t)\, dt \geq 0$$

for all $(x,u) \in \Theta(\overline{x}, \overline{u})$;

(d) *Complementarity condition:*

$$w^*(t) \geq 0 \quad and \quad w^*(t)(\overline{g}(t) + b\overline{u}(t)) = 0 \quad a.e. \ t \in [0,1].$$

Here, A^T *stands for the transpose of* A *and* q *is the conjugate number of* p, *that is,* $1 < q < +\infty$ *and* $1/p + 1/q = 1$.

3. Basic definitions and preliminaries

In this section, we recall some notations and facts from variational analysis and generalized differentiation, which will be used in the sequel. These notations and facts can be found in [9, 13, 20, 25, 29, 30, 37].

Let Y_1 and Y_2 be Banach spaces and $F: Y_1 \to 2^{Y_2}$ be a set-valued map. The effective domain, denoted by $\mathrm{dom}\, F$, and the graph of F, denoted by $\mathrm{gph}\, F$, are defined as

$$\mathrm{dom}\, F := \{z \in Y_1 : F(z) \neq \emptyset\}$$

and

$$\mathrm{gph}\, F := \{(z,v) \in Y_1 \times Y_2 : v \in F(z)\}.$$

Let Y be a Banach space, D be a nonempty closed and convex subset of Y, and $\overline{z} \in D$. We define

$$D(\overline{z}) = \mathrm{cone}(D - \overline{z}) = \{\lambda(d - \overline{z}) : d \in D, \lambda > 0\}.$$

The set

$$T(D; \overline{z}) = \liminf_{t \downarrow 0} \frac{D - \overline{z}}{t} = \{h \in Y : \forall t_n \downarrow 0, \exists h_n \to h, \overline{z} + t_n h_n \in D\}$$

is called the *tangent cone* to D at $\bar{z}$. It is well known that

$$T(D; \bar{z}) = \mathrm{cl}(D(\bar{z})) = \mathrm{cl}(\mathrm{cone}(D - \bar{z})).$$

The *second-order tangent set* to D at $\bar{z}$ in the direction $v \in E$ is defined by

$$T^2(D; \bar{z}, v) = \liminf_{t \downarrow 0} \frac{D - \bar{z} - tv}{t^2/2}$$
$$= \left\{ w : \forall t_n \downarrow 0, \exists w_n \to w, \bar{z} + t_n v + \frac{t_n^2}{2} w_n \in D \right\}.$$

When $v \in D(\bar{z}) = \mathrm{cone}(D - \bar{z})$, then there exists $\lambda > 0$ such that $v = \lambda(z - \bar{z})$ for some $z \in D$. By the convexity of D, for any $t_n \downarrow 0$, we have

$$t_n v = t_n \lambda z + (1 - t_n \lambda)\bar{z} - \bar{z} \in D - \bar{z}.$$

This implies that $\bar{z} + t_n v \in D$, and so, $0 \in T^2(D; \bar{z}, v)$. By [9, Proposition 3.1], we have

$$T^2(D; \bar{z}, v) = T(T(D; \bar{z}); v).$$

The set

$$N(D; \bar{z}) = \{ z^* \in Y^* : \langle z^*, z \rangle \le 0, \forall z \in T(D; \bar{z}) \}$$

is call *normal cone* to D at $\bar{z}$. We are familiar with

$$N(D; \bar{z}) = \{ z^* \in Y^* : \langle z^*, z - \bar{z} \rangle \le 0, \forall z \in D \}.$$

Obviously, $T^2(D; \bar{z}, 0) = T(D; \bar{z})$. It is clear that if $T^2(D; \bar{z}, z) \ne \emptyset$ then $z \in T(D; \bar{z})$. However, if we let $D = \{(z_1, z_2) \in \mathbb{R}^2 : z_2 \ge |z_1|^{3/2}\}$ and $\bar{z} = (0, 0)$, $v = (1, 0)$ then $v \in T(D, \bar{z}) = \{(z_1, z_2) \in \mathbb{R}^2 : z_2 \ge 0\}$, but $T^2(D, \bar{z}, v) = \emptyset$ by [5, Example 3.29].

4. The optimal control problem as a programming problem

By using some results from [9] on second-order necessary optimality conditions for a mathematical programming problem under inclusion constraints, this section establishes the theorem, which is the main tool for our subsequent investigations on the optimal control problem.

We suppose that X, Y, and Z are Banach spaces with the dual spaces X^*, Y^*, and Z^*, respectively. Assume that $f \colon Z \to \mathbb{R}$, $\mathcal{F} \colon Z \to Y$ are continuous mappings, $H \colon Z \to X$ is a continuous linear mapping, and $\mathcal{D} \subset Y$ is a closed convex set. Let $H^* \colon X^* \to Z^*$ be an adjoint mapping of H. Fixing $c \in X$, we put

$$\mathcal{E} = \{ z \in Z : H(z) = c \}.$$

Consider the programming problem

(4.1)   Minimize $\{f(z) : z \in \mathcal{E} \text{ and } \mathcal{F}(z) \in \mathcal{D}\}$.

By $\Phi = \mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D})$, we denote the *feasible set* of the problem (4.1), that is,

$$\Phi = \{z \in \mathcal{E} : \mathcal{F}(z) \in \mathcal{D}\}.$$

Let $\mathcal{Q}$ be a subset of Z. The *usual support function* $\sigma(\,\cdot\,, \mathcal{Q}) \colon Z^* \to \mathbb{R}$ of the set $\mathcal{Q}$ is defined by

$$\sigma(z^*, \mathcal{Q}) := \sup_{z \in \mathcal{Q}} \langle z^*, z \rangle.$$

In this section, we assume that f and $\mathcal{F}$ are second-order Fréchet differentiable around $\bar{z}$. In the sequel, we shall need the following lemmas.

Lemma 4.1. [9, Theorem 3.1] *Assume that the following regularity condition is satisfied:*

$$\nabla \mathcal{F}(\bar{z})(\mathcal{E}(\bar{z})) - \mathcal{D}(\mathcal{F}(\bar{z})) = Y.$$

Then, the following formulas are valid:

(i) $T(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \bar{z}) = T(\mathcal{E}; \bar{z}) \cap \nabla \mathcal{F}(\bar{z})^{-1}[T(\mathcal{D}; \mathcal{F}(\bar{z}))]$;

(ii) $T^2(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \bar{z}, z) = T^2(\mathcal{E}; \bar{z}, z) \cap \left[\nabla \mathcal{F}(\bar{z})^{-1}(T^2(\mathcal{D}; \mathcal{F}(\bar{z}), \nabla \mathcal{F}(\bar{z})z) - \nabla^2 \mathcal{F}(\bar{z})zz) \right]$.

Lemma 4.2. [9, Lemma 3] *Let $L \colon Z \to Y$ be a continuous linear mapping, and let $\mathcal{P} \subset Z$, $\mathcal{Q} \subset Y$ be two closed convex sets. If $Y = L(\mathcal{P}) - \mathcal{Q}$, then for all $z^* \in \mathrm{dom}[\sigma(\,\cdot\,, \mathcal{P} \cap L^{-1}(\mathcal{Q}))]$ we have*

$$\sigma(z^*, \mathcal{P} \cap L^{-1}(\mathcal{Q})) = \inf_{y^* \in Y^*} \{\sigma(z^* - L^*y^*, \mathcal{P}) + \sigma(y^*, \mathcal{Q})\},$$

the infimum being in fact attained.

Given a vector $\bar{z} \in Z$, the set

$$\mathcal{C}(\bar{z}) = \{z \in Z : \langle \nabla f(\bar{z}), z \rangle = 0, H(z) = 0, \nabla \mathcal{F}(\bar{z})z \in T(\mathcal{D}; \mathcal{F}(\bar{z}))\}$$

is called the set of *critical directions* for the problem (4.1) at $\bar{z}$, and the set

$$\mathcal{C}_0(\bar{z}) = \{z \in Z : \langle \nabla f(\bar{z}), z \rangle = 0, H(z) = 0, \nabla \mathcal{F}(\bar{z})z \in \mathrm{cone}(\mathcal{D} - \mathcal{F}(\bar{z}))\}$$

is called the set of *strictly critical directions* for the problem (4.1) at $\bar{z}$.

The following theorem is a shaper version of Commineti, which gives second-order necessary optimality conditions for the mathematical programming problem (4.1).

Theorem 4.3. *Suppose that $\bar{z}$ is a local minimum for (4.1) at which the following regularity condition is satisfied:*

$$\nabla \mathcal{F}(\bar{z})(\mathcal{E}(\bar{z})) - \mathcal{D}(\mathcal{F}(\bar{z})) = Y.$$

Assume that the mapping H is surjective. Then for each $z \in \mathcal{C}(\bar{z})$, there exist $w^* \in N(\mathcal{D}; \mathcal{F}(\bar{z}))$ and $x^* \in X^*$ such that the following conditions are fulfilled:

(i) *(Adjoint equation)*
$$\nabla f(\bar{z}) + \nabla \mathcal{F}(\bar{z})^* w^* + H^*(x^*) = 0;$$

(ii) *(Non-negative second-order condition)*
$$\nabla^2 f(\bar{z})zz + \langle w^*, \nabla^2 \mathcal{F}(\bar{z})zz \rangle \geq \sigma(w^*, T^2(\mathcal{D}; \mathcal{F}(\bar{z}), \nabla \mathcal{F}(\bar{z})z));$$

(iii) $\langle w^*, \nabla \mathcal{F}(\bar{z})z \rangle = 0.$

When $\mathcal{D}$ is in fact a cone, then we also have

(iv) *(Complementarity condition)*
$$\langle w^*, \mathcal{F}(\bar{z}) \rangle = 0; \quad w^* \in N(\mathcal{D}; 0).$$

Proof. We first claim that

(4.2) $$N(\mathcal{E}; z^1) = \{H^*(x^*) : x^* \in X^*\}$$

for each $z^1 \in \mathcal{E}$. Indeed, we have

$$\mathcal{E} = \{z \in Z : Hz = c\} = H^{-1}(c).$$

Note that H is a continuous linear mapping, so its derivative $\nabla H(z) = H$ for all $z \in Z$. Hence, [29, Corallary 1.15] implies that

$$N(\mathcal{E}; z^1) = N(H^{-1}(c); z^1) = H^* N(\{c\}; H(z^1)) = H^*(X^*).$$

By the definition of tangent cone and $N(\mathcal{E}; \bar{z}) = H^*(X^*)$, we have

$$\begin{aligned} T(\mathcal{E}; \bar{z}) &= \{z \in Z : \langle y^*, z \rangle \leq 0, \forall\, y^* \in N(\mathcal{E}; \bar{z})\} \\ &= \{z \in Z : \langle H^*(x^*), z \rangle \leq 0, \forall\, x^* \in X^*\} \\ &= \{z \in Z : \langle x^*, Hz \rangle \leq 0, \forall\, x^* \in X^*\} \\ &= \{z \in Z : Hz = 0\} = H^{-1}(0) := \mathcal{K}. \end{aligned}$$

Fixing any $z \in \mathcal{C}(\overline{z})$, we get $H(z + \overline{z}) = c$. Hence,

$$z \in \mathcal{E} - \overline{z} \subset \mathcal{E}(\overline{z}) \subset \mathrm{cl}(\mathcal{E}(\overline{z})) = T(\mathcal{E}; \overline{z}).$$

Therefore,

$$T^2(\mathcal{E}; \overline{z}, z) = T(T(\mathcal{E}; \overline{z}); z) = T(\mathcal{K}; z) = \mathcal{K} \quad \text{and} \quad 0 \in T^2(\mathcal{E}; \overline{z}, z).$$

We now consider the following two cases:

Case 1: $T^2(\mathcal{D}; \mathcal{F}(\overline{z}), \nabla \mathcal{F}(\overline{z})z) = \emptyset$. In this case, the non-negative second-order condition is automatically fulfilled because

$$\sigma(w^*, T^2(\mathcal{D}; \mathcal{F}(\overline{z}), \nabla \mathcal{F}(\overline{z})z)) = -\infty.$$

To obtain the assertions (i) and (iii), we shall separate the sets Ω and $T(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \overline{z})$. Here,

$$\Omega = \{v \in Z : \langle \nabla f(\overline{z})v < 0\}.$$

From the regularity condition, we obtain

(4.3) $$Y = \nabla \mathcal{F}(\overline{z})T(\mathcal{E}; \overline{z}) - T(\mathcal{D}; \mathcal{F}(\overline{z})).$$

So, we can find $w \in T(\mathcal{E}; \overline{z})$ such that

$$\nabla \mathcal{F}(\overline{z})w \in T(\mathcal{D}; \mathcal{F}(\overline{z})).$$

By Lemma 4.1, $w \in T(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \overline{z})$. Now, if $\nabla f(\overline{z}) = 0$, we may just take $w^* = 0$ and $x^* = 0$, so let us assume the contrary, in which case $\Omega \neq \emptyset$. We note that

$$\Omega \cap T(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \overline{z}) = \emptyset.$$

Indeed, if $w \in T(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \overline{z})$ we may choose $w_t \to w$ so that for $t > 0$ small enough such that

$$\overline{z} + tw_t \in \mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D})$$

and

$$f(\overline{z}) \leq f(\overline{z} + tw_t) = f(\overline{z}) + t\langle \nabla f(\overline{z}), w_t \rangle + o(t).$$

So

$$\langle \nabla f(\overline{z}), w \rangle \geq 0,$$

which is equivalent to $w \notin \Omega$. Thus, sets Ω and $T(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \overline{z})$ are nonvoid, convex, open and closed, respectively. The strict separation theorem implies that there exist a nonzero functional $z^* \in Z^*$ and a real $r \in \mathbb{R}$ such that

$$\langle z^*, v \rangle < r \leq \langle z^*, z \rangle, \quad \forall v \in \Omega, \ z \in T(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \overline{z}),$$

or equivalently

(4.4) $$\sigma(z^*, \Omega) + \sigma(-z^*, T(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \bar{z})) \leq 0.$$

So, we have

(4.5) $$\sigma(z^*, \Omega) < +\infty.$$

We will prove that $z^* = \lambda \nabla f(\bar{z})$ for some positive λ. Indeed, suppose that $z^* \notin \{\lambda \nabla f(\bar{z}) : \lambda > 0\}$. It follows from the strict separation theorem that there exists $z_1 \neq 0$ such that

$$\langle \lambda \nabla f(\bar{z}), z_1 \rangle \leq 0 < \langle z^*, z_1 \rangle, \quad \forall \lambda \geq 0.$$

Hence, $\nabla f(\bar{z}) z_1 \leq 0$. Let $z_2 \in \Omega$ then

$$\langle \nabla f(\bar{z}), z_2 + \alpha z_1 \rangle \leq \langle \nabla f(\bar{z}), z_2 \rangle < 0, \quad \forall \alpha > 0.$$

Therefore, $z_2 + \alpha z_1 \in \Omega$ for all $\alpha > 0$. On the other hand, $\langle z^*, z_2 + \alpha z_1 \rangle \to +\infty$ as $\alpha \to +\infty$, this implies that $\sigma(z^*, \Omega) = +\infty$, which contradicts (4.5). By eventually dividing by this λ, we may assume that $z^* = \nabla f(\bar{z})$ and then a direct calculation gives us

(4.6) $$\sigma(z^*, \Omega) = 0.$$

Concerning the second term in (4.4), we notice that Lemma 4.1 implies that

$$T(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \bar{z}) = \mathcal{P} \cap L^{-1}(\mathcal{Q}),$$

where

$$\mathcal{P} = T(\mathcal{E}; \bar{z}), \quad \mathcal{Q} = T(\mathcal{D}, \mathcal{F}(\bar{z})) \quad \text{and} \quad L = \nabla \mathcal{F}(\bar{z}).$$

Moreover, (4.3) gives us $Y = L(\mathcal{P}) - \mathcal{Q}$, so that we may use Lemma 4.2 in order to find $w^* \in Y^*$ such that

$$\sigma(-z^*, T(\mathcal{E} \cap \mathcal{F}^{-1}(\mathcal{D}); \bar{z})) = \sigma(-\nabla f(\bar{z}) - w^* \circ \nabla \mathcal{F}(\bar{z}), T(\mathcal{E}; \bar{z})) \\ + \sigma(w^*, T(\mathcal{D}; \mathcal{F}(\bar{z}))).$$

Combining (4.4) and (4.6), we have

(4.7) $$\langle \nabla f(\bar{z}) + w^* \circ \nabla \mathcal{F}(\bar{z}), w \rangle \geq \sigma(w^*, T(\mathcal{D}; \mathcal{F}(\bar{z}))), \quad \forall w \in T(\mathcal{E}; \bar{z}).$$

Choosing $w = 0 \in T(\mathcal{E}; \bar{z})$, we get

$$\langle w^*, z \rangle \leq 0, \quad \forall z \in T(\mathcal{D}; \mathcal{F}(\bar{z})).$$

So, $w^* \in N(\mathcal{D}; \mathcal{F}(\overline{z}))$. Since $0 \in T(\mathcal{D}; \mathcal{F}(\overline{z}))$ and (4.7), we obtain

$$\langle -\nabla f(\overline{z}) - w^* \circ \nabla \mathcal{F}(\overline{z}), w \rangle \leq 0, \quad \forall\, w \in T(\mathcal{E}; \overline{z}).$$

Hence, $-\nabla f(\overline{z}) - w^* \circ \nabla \mathcal{F}(\overline{z}) \in N(\mathcal{E}; \overline{z})$. Therefore, there exists $x^* \in X^*$ such that

$$\nabla f(\overline{z}) + w^* \circ \nabla \mathcal{F}(\overline{z}) + H^*(x^*) = 0,$$

this is the adjoint equation. From $z \in T(\mathcal{E}; \overline{z})$ and $-\nabla f(\overline{z}) - w^* \circ \nabla \mathcal{F}(\overline{z}) \in N(\mathcal{E}; \overline{z})$, we have

(4.8) $$\langle \nabla f(\overline{z}) + w^* \circ \nabla \mathcal{F}(\overline{z}), z \rangle \geq 0.$$

Besides,

$$\langle \nabla f(\overline{z}) + w^* \circ \nabla \mathcal{F}(\overline{z}), z \rangle = \langle \nabla f(\overline{z}), z \rangle + \langle w^* \circ \nabla \mathcal{F}(\overline{z}), z \rangle = \langle w^*, \nabla \mathcal{F}(\overline{z})z \rangle.$$

Since $\nabla \mathcal{F}(\overline{z})z \in T(\mathcal{D}; \mathcal{F}(\overline{z}))$ and $w^* \in N(\mathcal{D}; \mathcal{F}(\overline{z}))$, we get $\langle w^*, \nabla \mathcal{F}(\overline{z})z \rangle \leq 0$. Hence,

(4.9) $$\langle \nabla f(\overline{z}) + w^* \circ \nabla \mathcal{F}(\overline{z}), z \rangle \leq 0.$$

Combining (4.8) and (4.9), we obtain

$$\langle w^* \circ \nabla \mathcal{F}(\overline{z}), z \rangle = 0,$$

this is the assertion (iii).

Case 2: $T^2(\mathcal{D}; \mathcal{F}(\overline{z}), \nabla \mathcal{F}(\overline{z})z) \neq \emptyset$. By [9, Theorem 4.2], there exists $w^* \in N(\mathcal{D}; \mathcal{F}(\overline{z}))$ such that the lagrangian function $\mathcal{L} = f + w^*\mathcal{F}$ satisfies:

(4.10) $$-\nabla \mathcal{L}(\overline{z}) \in N(\mathcal{E}; \overline{z}),$$

(4.11) $$\langle \nabla \mathcal{L}(\overline{z}), z \rangle = 0,$$

(4.12) $$\langle \nabla \mathcal{L}(\overline{z}), w \rangle + \langle \nabla^2 \mathcal{L}(\overline{z})z, z \rangle \geq \sigma(w^*, T^2(\mathcal{D}; \mathcal{F}(\overline{z}), \nabla \mathcal{F}(\overline{z})z)), \quad \forall\, w \in T^2(\mathcal{E}; \overline{z}, z).$$

From (4.2), (4.10) and $\nabla L(\overline{z}) = \nabla f(\overline{z}) + \nabla \mathcal{F}(\overline{z})^*w^*$, we have the adjoint equation. Since (4.12) and $0 \in T^2(\mathcal{E}; \overline{z}, z)$, we get

$$\langle \nabla^2 \mathcal{L}(\overline{z})z, z \rangle \geq \sigma(w^*, T^2(\mathcal{D}; \mathcal{F}(\overline{z}), \nabla \mathcal{F}(\overline{z})z)),$$

this is the non-negative second-order condition. From (4.11) and $z \in \mathcal{C}(\overline{z})$, we obtain $\langle w^*, \nabla \mathcal{F}(\overline{z})z \rangle = 0$. The proof of theorem is complete. $\qquad \square$

Let $\overline{z}$ be a local minimum for the problem (4.1). The set of Lagrange multipliers denoted by $\Lambda(\overline{z})$ is defined as

$$\Lambda(\overline{z}) = \big\{ (w^*, x^*) \in Y^* \times X^* : \nabla f(\overline{z}) + \nabla \mathcal{F}(\overline{z})^*w^* + H^*x^* = 0, w^* \in N(\mathcal{D}; \mathcal{F}(\overline{z})) \big\}.$$

The following lemma gives the property of the set of Lagrange multipliers.

Lemma 4.4. *Suppose $\overline{z}$ is a local minimum for* (4.1) *at which the following regularity condition is satisfied:*

$$\nabla\mathcal{F}(\overline{z})(\mathcal{E}(\overline{z})) - \mathcal{D}(\mathcal{F}(\overline{z})) = Y.$$

Assume that the function H is surjective. Then, the set of Lagrange multipliers $\Lambda(\overline{z})$ is non-empty and bounded.

Proof. By the adjoint equation of Theorem 4.3, the set of Lagrange multipliers $\Lambda(\overline{z})$ is non-empty. We will prove that the set $\Lambda(\overline{z})$ is bounded. Fixing $(w_0^*, x_0^*) \in \Lambda(\overline{z})$, and let (w^*, x^*) be an any element of $\Lambda(\overline{z})$. Since the definition of $\Lambda(\overline{z})$, we have

$$(4.13) \qquad \nabla\mathcal{F}(\overline{z})^* w_0^* = -\nabla f(\overline{z}) - H^*(x_0^*) = 0, \quad w_0^* \in N(\mathcal{D}; \mathcal{F}(\overline{z}))$$

and

$$(4.14) \qquad \nabla\mathcal{F}(\overline{z})^* w^* = -\nabla f(\overline{z}) - H^*(x^*) = 0, \quad w^* \in N(\mathcal{D}; \mathcal{F}(\overline{z})).$$

By [46, Theorem 2.1], there exists a constant $\rho > 0$ such that

$$B_Y(0, \rho) \subset \nabla\mathcal{F}(\overline{z})[\mathcal{E}(\overline{z}) \cap B_Z] - [(\mathcal{D} - \mathcal{F}(\overline{z})) \cap B_Y],$$

where $B_Y(0, \rho)$ is a ball around zero with radius ρ, B_Z and B_Y are unit balls in Z and Y, respectively. Thus, for each $y \in B_Y(0, \rho)$, there exists $z \in \mathcal{E}(\overline{z}) \cap B_Z$ and $v \in (\mathcal{D} - \mathcal{F}(\overline{z})) \cap B_Y$ such that $y = \nabla\mathcal{F}(\overline{z})z - v$. From (4.13) and (4.14), we have

$$\begin{aligned}
\langle w_0^* - w^*, y \rangle &= \langle w_0^* - w^*, \nabla\mathcal{F}(\overline{z})z - v \rangle \\
&= \langle \nabla\mathcal{F}(\overline{z})^*(w_0^* - w^*), z \rangle + \langle w^* - w_0^*, v \rangle \\
&= \langle H^*(x^* - x_0^*), z \rangle + \langle w^* - w_0^*, v \rangle \\
&= \langle x_0^* - x^*, Hz \rangle + \langle w^*, v \rangle - \langle w_0^*, v \rangle \\
&= \langle w^*, v \rangle - \langle w_0^*, v \rangle.
\end{aligned}$$

We note that $w^* \in N(\mathcal{D}; \mathcal{F}(\overline{z}))$ and $v \in \mathcal{D} - \mathcal{F}(\overline{z}) \subset T(\mathcal{D}; \mathcal{F}(\overline{z}))$. So, $\langle w^*, v \rangle \leq 0$. Hence,

$$\langle w_0^* - w^*, y \rangle \leq -\langle w_0^*, v \rangle \leq \|w_0^*\| \|v\| \leq \|w_0^*\|.$$

Moreover, $-y$ also belongs to $B_Y(0, \rho)$, one has

$$\langle w_0^* - w^*, y \rangle = -\langle w_0^* - w^*, -y \rangle \geq \|w_0^*\|.$$

So,

$$\|w_0^* - w^*\| = \sup_{y \in B_Y(0, \rho)} \left| \left\langle w_0^* - w^*, \frac{1}{\rho}y \right\rangle \right| \leq \frac{1}{\rho} \|w_0^*\|.$$

Therefore,

$$(4.15) \qquad \|w^*\| \le \left(1 + \frac{1}{\rho}\right)\|w_0^*\|.$$

Since $\nabla H(\bar{z}) = H$ is surjective, there exists $\gamma > 0$ such that $B_X(0, \gamma) \subset H(B_Z)$. Thus, for any $x \in B_X(0, \gamma)$, there is $z \in B_Z$ satisfying $x = H(z)$. Then, one has

$$\begin{aligned}
\langle x^*, x \rangle = \langle x^*, H(z) \rangle &= \langle H^*(x^*), z \rangle \\
&= \langle -\nabla f(\bar{z}) - \nabla \mathcal{F}(\bar{z})w^*, z \rangle \\
&\le \|\nabla f(\bar{z}) + \nabla \mathcal{F}(\bar{z})w^*\|\|z\| \\
&\le \|\nabla f(\bar{z})\| + \|\nabla \mathcal{F}(\bar{z})\|\|w^*\|.
\end{aligned}$$

Thus for all $x \in B_X$, we have

$$\langle x^*, x \rangle \le \frac{1}{\gamma}(\|\nabla f(\bar{z})\| + \|\nabla \mathcal{F}(\bar{z})\|\|w^*\|),$$

which implies that

$$\|x^*\| \le \frac{1}{\gamma}(\|\nabla f(\bar{z})\| + \|\nabla \mathcal{F}(\bar{z})\|\|w^*\|).$$

Therefore, we have

$$\|x^*\| \le \frac{1}{\gamma}\left(\|\nabla f(\bar{z})\| + \|\nabla \mathcal{F}(\bar{z})\|\left(1 + \frac{1}{\rho}\right)\|w_0^*\|\right).$$

From this and (4.15), (w^*, x^*) is bounded, and this implies the boundedness of the set of Lagrange multipliers $\Lambda(\bar{z})$. The proof of the lemma is complete. $\qquad \square$

The set $\mathcal{D}$ is said to be *polyhedric* at $\bar{y} \in \mathcal{D}$ if for all $q^* \in N(\mathcal{D}; \bar{z})$, then

$$T(\mathcal{D}; \bar{y}) \cap (q^*)^\perp = \mathrm{cl}[\mathrm{cone}(\mathcal{D} - \bar{y}) \cap (q^*)^\perp],$$

where $(q^*)^\perp = \{q \in Y : \langle q^*, q \rangle = 0\}$. We say that the problem (4.1) satisfies the *strong extended polyhedricity condition* at a feasible point $\bar{z}$ if $C_0(\bar{z})$ is a dense subset of $C(\bar{z})$.

When the problem (4.1) satisfies the strong extended polyhedricity condition, we have the following result.

Corollary 4.5. *Suppose that the problem* (4.1) *satisfies the strong extended polyhedricity condition at $\bar{z}$ and $\bar{z}$ is a local minimum for* (4.1) *at which the following regularity condition is satisfied:*

$$\nabla \mathcal{F}(\bar{z})(\mathcal{E}(\bar{z})) - \mathcal{D}(\mathcal{F}(\bar{z})) = Y.$$

Assume that the mapping H is surjective. Then for each $z \in \mathcal{C}(\bar{z})$, there exist $w^* \in N(\mathcal{D}; \mathcal{F}(\bar{z}))$ and $x^* \in X^*$ such that the following conditions are fulfilled:

(i) *(Adjoint equation)*
$$\nabla f(\bar{z}) + \nabla \mathcal{F}(\bar{z})^* w^* + H^*(x^*) = 0;$$

(ii) *(Non-negative second-order condition)*
$$\nabla^2 f(\bar{z}) zz + \langle w^*, \nabla^2 \mathcal{F}(\bar{z}) zz \rangle \geq 0;$$

(iii) $\langle w^*, \nabla \mathcal{F}(\bar{z}) z \rangle = 0.$

When $\mathcal{D}$ is in fact a cone, then we also have

(iv) *(Complementarity condition)*
$$\langle w^*, \mathcal{F}(\bar{z}) \rangle = 0; \quad w^* \in N(\mathcal{D}; 0).$$

Proof. Take any $z \in \mathcal{C}(\bar{z})$. Since the strong extended polyhedricity condition of (4.1), there exists a sequence $z_n \in \mathcal{C}_0(\bar{z})$, $z_n \to z$ as $n \to \infty$. For each $n \in \mathbb{N}$, we have

$$\langle \nabla f(\bar{z}), z_n \rangle = 0, \quad z_n \in T(\mathcal{A}; \bar{z}), \quad \nabla \mathcal{F}(\bar{z}) z_n \in \mathrm{cone}(\mathcal{D} - \mathcal{F}(\bar{z})).$$

Since $\nabla \mathcal{F}(\bar{z}) z_n \in \mathrm{cone}(\mathcal{D} - \mathcal{F}(\bar{z}))$, we have $0 \in T^2(\mathcal{D}; \mathcal{F}(\bar{z}), \nabla \mathcal{F}(\bar{z}) z_n)$. From $z_n \in \mathcal{C}_0(\bar{z}) \subset \mathcal{C}(\bar{z})$ and Theorem 4.3, there exist $x_n^* \in X^*$ and $w_n^* \in N(\mathcal{D}; \mathcal{F}(\bar{z}))$ such that the following conditions are fulfilled:

(i') (Adjoint equation)
$$\nabla f(\bar{z}) + \nabla \mathcal{F}(\bar{z})^* w_n^* + H^* x_n^* = 0;$$

(ii') (Nonnegativeness second-order condition)
$$\nabla^2 f(\bar{z}) z_n z_n + \langle w_n^*, \nabla^2 \mathcal{F}(\bar{z}) z_n z_n \rangle \geq \sigma(w_n^*, T^2(\mathcal{D}; \mathcal{F}(\bar{z}), \nabla \mathcal{F}(\bar{z}) z_n))$$
$$\geq \langle w_n^*, 0 \rangle = 0;$$

(iii') $\langle w_n^*, \nabla \mathcal{F}(\bar{z}) z_n \rangle = 0.$

When $\mathcal{D}$ is in fact a cone, then we also have

(iv') (Complementarity condition)
$$\langle w_n^*, \mathcal{F}(\bar{z}) \rangle = 0; \quad w_n^* \in N(\mathcal{D}; 0).$$

By Lemma 4.4, w_n^* and x_n^* are bounded for all $n \in \mathbb{N}$. So, w_n^* weakly* converges to w^* and x_n^* weakly* converges to x^* as $n \to \infty$. Letting $n \to \infty$, we obtain the adjoint equation, the nonnegativeness second-order condition and the assertion (iii) of corollary from (i'), (ii') and (iii'). The proof of the corollary is complete. $\square$

5. Proof of the main result

The next step is to apply Theorem 4.3 to the problem (1.1)–(1.4). In order to use this theorem, we first define the following linear mappings: $\mathcal{A} \colon X \to X$ defined by

$$\mathcal{A}x := x - \int_0^{(\cdot)} A(\tau)x(\tau)\,d\tau,$$

$\mathcal{B} \colon U \to X$ defined by

$$\mathcal{B}u := - \int_0^{(\cdot)} B(\tau)u(\tau)\,d\tau$$

and $\mathcal{M} \colon X \times U \to X$ defined by

$$\mathcal{M}(x,u) := \mathcal{A}x + \mathcal{B}u.$$

Under the hypothesis (H2), (1.5) can be written in the form

$$E = \left\{ (x,u) \in X \times U : x = c + \int_0^{(\cdot)} Ax\,d\tau + \int_0^{(\cdot)} Bu\,d\tau \right\}$$

$$= \left\{ (x,u) \in X \times U : x - \int_0^{(\cdot)} Ax\,d\tau - \int_0^{(\cdot)} Bu\,d\tau = c \right\}$$

$$= \{ (x,u) \in X \times U : \mathcal{M}(x,u) = c \}.$$

Recall that for $1 < p < \infty$, we have $L^p([0,1], \mathbb{R}^n)^* = L^q([0,1], \mathbb{R}^n)$, where

$$1 < q < +\infty, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Besides, $L^p([0,1], \mathbb{R}^n)$ is pared with $L^q([0,1], \mathbb{R}^n)$ by the formula

$$\langle x^*, x \rangle = \int_0^1 x^*(t)x(t)\,dt$$

for all $x^* \in L^q([0,1], \mathbb{R}^n)$ and $x \in L^p([0,1], \mathbb{R}^n)$.

Also, we have $W^{1,p}([0,1], \mathbb{R}^n)^* = \mathbb{R}^n \times L^q([0,1], \mathbb{R}^n)$ and $W^{1,p}([0,1], \mathbb{R}^n)$ is pared with $\mathbb{R}^n \times L^q([0,1], \mathbb{R}^n)$ by the formula

$$\langle (a,u), x \rangle = \langle a, x(0) \rangle + \int_0^1 u(t)\dot{x}(t)\,dt$$

for all $(a,u) \in \mathbb{R}^n \times L^q([0,1], \mathbb{R}^n)$ and $x \in W^{1,p}([0,1], \mathbb{R}^n)$ (see [18, p. 21]).

In the case of $p = 2$, $W^{1,2}([0,1], \mathbb{R}^n)$ becomes a Hilbert space with the inner product given by

$$\langle x, y \rangle = \langle x(0), y(0) \rangle + \int_0^1 \dot{x}(t)\dot{y}(t)\,dt$$

for all $x, y \in W^{1,2}([0,1], \mathbb{R}^n)$.

In the sequel, we shall need the following lemma for formulas to define the adjoint mappings $\mathcal{M}^*$, $\mathcal{A}^*$ and $\mathcal{B}^*$.

Lemma 5.1. [44, Lemma 2.3] *Suppose that $\mathcal{M}^*$ is the adjoint mapping of $\mathcal{M}$. Then, the following assertions are valid.*

(i) *The mapping $\mathcal{M}$ is continuous;*

(ii) *$\mathcal{M}^*(a, u) = (\mathcal{A}^*(a, u), \mathcal{B}^*(a, u))$, where $\mathcal{B}^*(a, u) = -B^T u$, and*

$$\mathcal{A}^*(a, u) = \left(a - \int_0^1 A^T(t) u(t)\, dt; u + \int_0^{(\cdot)} A^T(\tau) u(\tau)\, d\tau - \int_0^1 A^T(t) u(t)\, dt \right)$$

for all $(a, u) \in \mathbb{R}^n \times L^q([0, 1], \mathbb{R}^n)$.

Using a similar technique to the proof of [18, Corollary, p. 52], we obtain the following result.

Lemma 5.2. *Suppose that the assumption* (H2) *is satisfied and $\widetilde{x} \in X = W^{1,p}([0, 1], \mathbb{R}^n)$. Then, the equation*

$$x = \int_0^{(\cdot)} A^T(\tau) x(\tau)\, d\tau + \widetilde{x}$$

has an unique solution in X.

By Lemma 5.2, the mapping $\mathcal{A}$ is surjective and the set E is not empty.

The following lemma gives the second-order Fréchet differentiability of the objective function.

Lemma 5.3. *Suppose that the assumption* (H1) *is satisfied. Then, the functional J is second-order Fréchet differentiable around $(\overline{x}, \overline{u})$ and $\nabla J(\overline{z})$, $\nabla^2 J(\overline{z})$ are given by*

(i) *$\nabla J(\overline{z}) = \nabla J(\overline{x}, \overline{u}) = (J_x(\overline{x}, \overline{u}), J_u(\overline{x}, \overline{u}))$ with*

$$J_u(\overline{x}, \overline{u}) = L_u(\,\cdot\,, \overline{x}, \overline{u})$$

and

$$J_x(\overline{x}, \overline{u}) = \left(\int_0^1 L_x(t, \overline{x}(t), \overline{u}(t))\, dt, \int_{(\cdot)}^1 L_x(s, \overline{x}(s), \overline{u}(s))\, ds \right);$$

(ii)

$$\nabla^2 J(\overline{z}) z z$$
$$= \int_0^1 \left(L_{xx}(t, \overline{x}(t), \overline{u}(t)) x^2(t) + 2 L_{xu}(t, \overline{x}(t), \overline{u}(t)) x(t) u(t) + L_{uu}(t, \overline{x}(t), \overline{u}(t)) u^2(t) \right) dt$$

for any $z = (x, u) \in X \times U = Z$.

Proof. By the assumption (H1), the functional J is second-order Fréchet differentiable around $(\overline{x}, \overline{u})$ and $\nabla^2 J(\overline{z})$ is given by

$$\nabla^2 J(\overline{z})zz$$
$$= \int_0^1 \left(L_{xx}(t, \overline{x}(t), \overline{u}(t))x^2(t) + 2L_{xu}(t, \overline{x}(t), \overline{u}(t))x(t)u(t) + L_{uu}(t, \overline{x}(t), \overline{u}(t))u^2(t) \right) dt$$

for any $z = (x, u) \in X \times U = Z$ (see [21, Proposition A]). We now compute $J_u(\overline{x}, \overline{u})$ and $J_x(\overline{x}, \overline{u})$. It is easy to verify that

$$J_u(\overline{x}, \overline{u}) = L_u(\,\cdot\,, \overline{x}(\cdot), \overline{u}(\cdot)).$$

Taking any $h \in W^{1,p}([0, 1], \mathbb{R}^n)$, we have

$$J_x(\overline{x}, \overline{u})h = \int_0^1 L_x(t, \overline{x}(t), \overline{u}(t))h(t)\, dt.$$

Integrating by part yields

$$J_x(\overline{x}, \overline{u})h = h(1) \int_0^1 L_x(t, \overline{x}(t), \overline{u}(t))\, dt - \int_0^1 \left[\int_0^t L_x(s, \overline{x}(s), \overline{u}(s))\, ds \right] \dot{h}(t)\, dt$$

$$= \int_0^1 L_x(t, \overline{x}(t), \overline{u}(t))\, dt \left[h(0) + \int_0^1 \dot{h}(t)\, dt \right]$$

$$- \int_0^1 \left[\int_0^t L_x(s, \overline{x}(s), \overline{u}(s))\, ds \right] \dot{h}(t)\, dt$$

$$= \left(\int_0^1 L_x(t, \overline{x}(t), \overline{u}(t))\, dt \right) h(0)$$

$$+ \int_0^1 \left[\int_0^1 L_x(t, \overline{x}(t), \overline{u}(t))\, dt - \int_0^t L_x(s, \overline{x}(s), \overline{u}(s))\, ds \right] \dot{h}(t)\, dt.$$

Hence,

$$J_x(\overline{x}, \overline{u}) = \left(\int_0^1 L_x(t, \overline{x}(t), \overline{u}(t))\, dt, \int_0^1 L_x(t, \overline{x}(t), \overline{u}(t))\, dt - \int_0^{(\cdot)} L_x(s, \overline{x}(s), \overline{u}(s))\, ds \right)$$

$$= \left(\int_0^1 L_x(t, \overline{x}(t), \overline{u}(t))\, dt, \int_{(\cdot)}^1 L_x(s, \overline{x}(s), \overline{u}(s))\, ds \right).$$

Thus, the proof of the lemma is complete. $\qquad\square$

Given $x \in X$, we put $M = \|x\|_0 = \max_{t \in [0,1]} |x(t)|$. By the assumption (H3), there exists a constant $l_{g_M} > 0$ such that $|g_x(t, x)| \leq l_{g_M}$ for all $x \in \mathbb{R}^n$ with $|x| \leq M$. By the Taylor expansion, we get

$$|g(t, x(t))| \leq |g(t, x(t)) - g(t, 0)| + |g(t, 0)|$$
$$= |g_x(t, \theta(t)x(t))x(t)| + |g(t, 0)|$$
$$\leq l_{g_M} M + |g(t, 0)|.$$

This implies that $g(\,\cdot\,, x) \in L^p([0, 1], \mathbb{R}^m)$.

The following lemma gives the second-order Fréchet differentiability of the constraint function.

Lemma 5.4. *Suppose that the assumption* (H3) *is satisfied. Then, the functional G is second-order Fréchet differentiable around $\overline{z} = (\overline{x}, \overline{u})$ and*

(i) $\nabla^2 G(\overline{z})zz = \overline{g}_{xx}(\cdot)x^2$, $\forall z = (x, u) \in X \times U = Z$;

(ii) $w^* \circ \nabla G(\overline{z}) = w^* \circ \nabla G(\overline{x}, \overline{u}) = (w^* \circ G_x(\overline{x}, \overline{u}), w^* \circ G_u(\overline{x}, \overline{u}))$ *is given by*

$$w^* \circ G_u(\overline{x}, \overline{u}) = bw^*,$$

$$w^* \circ G_x(\overline{x}, \overline{u}) = \left(\int_0^1 w^*(t)\overline{g}_x(t)\, dt, \int_{(\cdot)}^1 w^*(\tau)\overline{g}_x(\tau)\, d\tau \right)$$

for any $w^* \in L^q([0, 1], \mathbb{R}^m)$.

Proof. We first prove that the functional G is Fréchet differentiable around $\overline{z} = (\overline{x}, \overline{u})$ and

$$\nabla G(\overline{x}, \overline{u}) = (\overline{g}_x(\cdot), b).$$

Fixing any $h \in X$, $v \in U$, we choose a positive $M > \|h\|_0 + \|\overline{x}\|_0$. We have

$$\|G(\overline{x} + h, \overline{u} + v) - G(\overline{x}, \overline{u}) - \overline{g}_x(\cdot)h - bv\|^p$$
$$= \int_0^1 |g(t, \overline{x} + h) - g(t, \overline{x}) - g_x(t, \overline{x})h|^p\, dt.$$

By the mean value theorem,

$$\|G(\overline{x} + h, \overline{u} + v) - G(\overline{x}, \overline{u}) - \overline{g}_x(\cdot)h - bv\|^p$$
$$= \int_0^1 \left| \int_0^1 g_x(t, \overline{x} + \theta h)h\, d\theta - g_x(t, \overline{x})h \right|^p dt$$
$$\leq \int_0^1 \left(\int_0^1 |g_x(t, \overline{x} + \theta h)h - g_x(t, \overline{x})h|\, d\theta \right)^p dt$$
$$\leq \int_0^1 \int_0^1 |g_x(t, \overline{x} + \theta h)h - g_x(t, \overline{x})h|^p\, d\theta dt.$$

By the assumption (H3), there exists $l_{gM} > 0$ such that inequalities in (H3) are fulfilled. This implies that

$$\|G(\overline{x} + h, \overline{u} + v) - G(\overline{x}, \overline{u}) - \overline{g}_x(\cdot)h - bv\|^p \leq \int_0^1 \int_0^1 l_{gM}^p |\theta h|^p |h|^p\, d\theta dt.$$

For any $t \in [0, 1]$, we have

$$h(t) = h(0) + \int_0^t \dot{h}(\tau)\, d\tau.$$

So,

$$|h(t)| \leq |h(0)| + \left| \int_0^t \dot{h}(\tau)\, d\tau \right| \leq |h(0)| + \left| \int_0^1 \dot{h}(t)\, dt \right| \leq |h(0)| + \|\dot{h}\|_p = \|h\|_{1,p}.$$

Hence,

$$\|h\|_0 \leq \|h\|_{1,p}.$$

Since $|\theta| \leq 1$, $|h|^p \leq \|h\|_0^p$ and $\|h\|_0 \leq \|h\|_{1,p}$, we get

$$\|G(\overline{x}+h, \overline{u}+v) - G(\overline{x}, \overline{u}) - \overline{g}_x(\cdot)h - bv\|^p \leq l_{gM}^p \|h\|_p^p \|h\|_0^p$$
$$\leq l_{gM}^p \|h\|_p^p (\|h\|_{1,p} + \|v\|_p)^p.$$

So,

$$\frac{\|G(\overline{x}+h, \overline{u}+v) - G(\overline{x}, \overline{u}) - \overline{g}_x(\cdot)h - bv\|}{\|h\|_{1,p} + \|v\|_p} \leq l_{gM} \|h\|_p.$$

Hence,

$$\frac{\|G(\overline{x}+h, \overline{u}+v) - G(\overline{x}, \overline{u}) - \overline{g}_x(\cdot)h - bv\|}{\|h\|_{1,p} + \|v\|_p} \to 0 \quad \text{as } \|h\|_{1,p} + \|v\|_p \to 0.$$

Therefore, the functional G is Fréchet differentiable and

$$\nabla G(\overline{x}, \overline{u}) = (\overline{g}_x(\cdot), b).$$

We now show that ∇G is continuous on a neighborhood of $(\overline{x}, \overline{u})$. In fact, for each (x_0, u_0) in a neighborhood of $(\overline{x}, \overline{u})$, for any sequence $z_n = (x_n, u_n) \to z_0 = (x_0, u_0)$ in $X \times U$ and for any $z = (x, u) \in Z$ with $\|z\|_Z \leq 1$, we have

$$\|\nabla G(z_n)z - \nabla G(z_0)z\|_p^p = \int_0^1 |(g_x(t, x_n) - g_x(t, x_0))x|^p \, dt.$$

Using the assumption (H3) and the similar technique to the above, we can show that

$$\|\nabla G(z_n)z - \nabla G(z_0)z\|_p^p \to 0 \quad \text{as } n \to \infty.$$

This implies that ∇G is continuous at z_0. Also by the assumption (H3), we can prove that

$$H(x, u) = g_x(\cdot, x, u) \quad \text{and} \quad K(x, u) = b$$

are Fréchet differentiable and H, K are of class C^1 around $(\overline{x}, \overline{u})$. So, the functional G is second-order Fréchet differentiable around $\overline{z} = (\overline{x}, \overline{u})$, which is given by

$$\nabla^2 G(\overline{z})zz = \overline{g}_{xx}(\cdot)x^2, \quad \forall z = (x, u) \in X \times U = Z;$$

this is the assertion (i) of the lemma. For any $w^* \in L^q([0,1], \mathbb{R}^m)$, we now compute $w^* \circ \nabla G(\overline{x}, \overline{u})$. We have

$$w^* \circ \nabla G(\overline{x}, \overline{u}) = (w^* \circ \overline{g}_x(\cdot), bw^*).$$

Taking any $h \in X = W^{1,p}([0,1], \mathbb{R}^n)$, we have

$$\langle w^* \circ \overline{g}_x(\cdot), h \rangle = \int_0^1 w^*(t) \overline{g}_x(t) h(t) \, dt.$$

Integrating by part yields

$$
\begin{aligned}
&\langle w^* \circ \overline{g}_x(\cdot), h \rangle \\
&= h(1) \int_0^1 w^*(t) \overline{g}_x(t) \, dt - \int_0^1 \left(\int_0^t w^*(\tau) \overline{g}_x(\tau) \, d\tau \right) \dot{h}(t) \, dt \\
&= \int_0^1 w^*(t) \overline{g}_x(t) \, dt \left(h(0) + \int_0^1 \dot{h}(t) \, dt \right) - \int_0^1 \left(\int_0^t w^*(\tau) \overline{g}_x(\tau) \, d\tau \right) \dot{h}(t) \, dt \\
&= \left(\int_0^1 w^*(t) \overline{g}_x(t) \, dt \right) h(0) + \int_0^1 \left(\int_0^1 w^*(t) \overline{g}_x(t) \, dt - \int_0^t w^*(\tau) \overline{g}_x(\tau) \, d\tau \right) \dot{h}(t) \, dt \\
&= \left(\int_0^1 w^*(t) \overline{g}_x(t) \, dt \right) h(0) + \int_0^1 \left(\int_t^1 w^*(\tau) \overline{g}_x(\tau) \, d\tau \right) \dot{h}(t) \, dt.
\end{aligned}
$$

Hence,

$$w^* \circ \overline{g}_x(\cdot) = \left(\int_0^1 w^*(t) \overline{g}_x(t) \, dt, \int_{(\cdot)}^1 w^*(\tau) \overline{g}_x(\tau) \, d\tau \right).$$

Thus, the proof of the lemma is complete. $\qquad\qquad\square$

.

We have the following result on the regularity condition for the problem (1.1)–(1.4).

Lemma 5.5. *Suppose that the assumptions* (H2)–(H4) *are satisfied. Then, the regularity condition is fulfilled, that is,*

$$\nabla G(\overline{z})(E(\overline{z})) - Q(G(\overline{z})) = U.$$

Proof. We have

$$Q(G(\overline{z})) = \mathrm{cone}(Q - G(\overline{z})) = Q(\overline{g}(\cdot) + b\overline{u}).$$

So, the proof will be complete if we can show that

$$\nabla G(\overline{z})(E(\overline{z})) - Q(\overline{g}(\cdot) + b\overline{u}) = U.$$

Taking any $v \in U$, we now represent

$$v = v^1 - v^2, \quad \text{where } v^1, v^2 \in Q.$$

By the assumption (H4), there exists $u \in L^p([0,1], \mathbb{R}^m)$ such that

$$\overline{g}_x(\cdot)(x - \overline{x}) + bu + \overline{g}(\cdot) - v^1 \in Q,$$

where x and u are satisfied the state equation (1.2) and the initial condition (1.3). Define

$$\begin{aligned}
d &= \overline{g}_x(\cdot)(x - \overline{x}) + b(u - \overline{u}) + \overline{g}(\cdot) + b\overline{u} - v \\
&= \overline{g}_x(\cdot)(x - \overline{x}) + bu + \overline{g}(\cdot) - v^1 + v^2.
\end{aligned}$$

We see that $d \in Q$ and

$$\overline{g}_x(\cdot)(x - \overline{x}) + b(u - \overline{u}) - (d - (\overline{g}(\cdot) + b\overline{u})) = v.$$

We note that

$$\begin{aligned}
z - \overline{z} &= (x, u) - (\overline{x}, \overline{u}) = (x - \overline{x}, u - \overline{u}) \in E(\overline{z}), \\
\overline{g}_x(\cdot)(x - \overline{x}) + b(u - \overline{u}) &\in \nabla G(\overline{z})(E(\overline{z}))
\end{aligned}$$

and

$$d - (\overline{g}(\cdot) + b\overline{u}) \in Q(\overline{g}(\cdot) + b\overline{u}).$$

Hence, the proof of the lemma is complete. $\qquad\square$

We have the following result on the polyhedricity of Q.

Lemma 5.6. *The following formula holds:*

$$T(Q, G(\overline{z})) \cap (q^*)^\perp = \mathrm{cl}[\mathrm{cone}(Q - G(\overline{z})) \cap (q^*)^\perp], \quad \forall\, q^* \in N(Q; G(\overline{z})).$$

Proof. The inclusion $\mathrm{cl}[\mathrm{cone}(Q - G(\overline{z})) \cap (q^*)^\perp] \subset T(Q, G(\overline{z})) \cap (q^*)^\perp$ being trivial. We have to prove other one. Note that $\mathrm{cone}(Q - G(\overline{z})) = \{v \in U : \gamma v + G(\overline{z}) \in Q, \forall\, \gamma > 0\}$. Now fix $q^* \in N(Q; G(\overline{z}))$ and let $v_0 \in T(Q, G(\overline{z})) \cap (q^*)^\perp$. For $\epsilon > 0$, we define

$$v_\epsilon(t) = \begin{cases} v_0(t) & \text{if } G(\overline{z})(t) + \epsilon v_0(t) \leq 0, \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that $v_\epsilon \in \mathrm{cone}(Q - G(\overline{z})) \cap (q^*)^\perp$. Letting $\epsilon \to 0$, we have $v_\epsilon(t) \to v_0(t)$ a.e. $t \in [0,1]$. The Dominated Convergence Theorem gives that $v_\epsilon \to v_0$ in $U = L^p([0,1], \mathbb{R}^m)$. Hence, the proof of the lemma is complete. $\qquad\square$

Proof of the main result. It is easy to prove that J, G are continuous. From Lemma 5.2 to Lemma 5.5, we see that all assumptions of Theorem 4.3 are fulfilled. Take any $(x, u) \in \Theta(\overline{x}, \overline{u})$. By Lemma 5.3, $\nabla J(\overline{z}) = \nabla J(\overline{x}, \overline{u}) = (J_x(\overline{x}, \overline{u}), J_u(\overline{x}, \overline{u}))$, where

$$J_u(\overline{x}, \overline{u}) = \overline{L}_u(\cdot) \quad \text{and} \quad J_x(\overline{x}, \overline{u}) = \left(\int_0^1 \overline{L}_x(t)\, dt, \int_{(\cdot)}^1 \overline{L}_x(\tau)\, d\tau \right).$$

So,

(5.1)
$$\nabla J(\bar{z})z = \nabla J(\bar{x}, \bar{u})(x, u) = J_x(\bar{x}, \bar{u})x + J_u(\bar{x}, \bar{u})u$$
$$= x(0) \int_0^1 \overline{L}_x(t)\, dt + \int_0^1 \left[\int_t^1 \overline{L}_x(\tau)\, d\tau\right] \dot{x}(t)\, dt + \int_0^1 \overline{L}_u(t)u(t)\, dt.$$

By the condition (C1), $x(0) = 0$ and (5.1), we get

(5.2)
$$\nabla J(\bar{z})z = 0.$$

From the condition (C2), we have

$$x - \int_0^{(\cdot)} Ax\, d\tau - \int_0^{(\cdot)} Bu\, d\tau = 0.$$

This is equivalent to

(5.3)
$$\mathcal{M}z = 0.$$

By the condition (C3) and [1, Theorem 8.5.1], we have

(5.4)
$$\nabla G(\bar{z})z = \bar{g}_x x + bu \in T(Q; G(\bar{x}, \bar{u})).$$

Combining (5.2), (5.3) and (5.4), the condition $z \in \mathcal{C}(\bar{z})$ of Theorem 4.3 is fulfilled. According to Theorem 4.3, there exist $w^* \in U^* = L^q([0,1], \mathbb{R}^m)$ and $x^* = (a, v) \in \mathbb{R}^n \times L^q([0,1], \mathbb{R}^n)$ such that the following conditions are fulfilled:

(a1) (Adjoint equation)
$$\nabla J(\bar{z}) + \nabla G(\bar{z})^* w^* + \mathcal{M}^*(a, v) = 0;$$

(a2) (Non-negative second-order condition)

$$\nabla^2 J(\bar{z})zz + \langle w^*, \nabla^2 G(\bar{z})zz \rangle \geq \sigma(w^*, T^2(Q; G(\bar{z}), \nabla G(\bar{z})z));$$

(a3) (Complementarity condition)

$$\langle w^*, G(\bar{z}) \rangle = 0; \quad w^* \in N(Q; 0).$$

By the complementarity condition, we have

$$\langle w^*, w \rangle \leq 0, \quad \forall w \in Q.$$

It is equivalent to

(5.5)
$$\int_0^1 w^*(t)w(t)\, dt \leq 0, \quad \forall w \in Q.$$

We will show that
$$w^*(t)v \leq 0, \quad \forall\, v \leq 0, \text{ a.e. } t \in [0, 1].$$

On the contrary, suppose that there exist a set $D \subset [0, 1]$, $\mu(D) > 0$ and $v_0 \leq 0$ such that

$$w^*(t)v_0 > 0, \quad \forall\, t \in D.$$

Choosing
$$v(t) = \begin{cases} v_0 & \text{if } t \in D, \\ 0 & \text{if } t \in [0, 1] \setminus D. \end{cases}$$

Then, $v \in Q$ and
$$\int_0^1 w^*(t)v(t)\, dt = \int_D w^*(t)v_0\, dt > 0,$$

which contradicts (5.5). Thus,

$$w^*(t)v \leq 0, \quad \forall\, v \leq 0, \text{ a.e. } t \in [0, 1].$$

This is equivalent to

(5.6) $$\qquad\qquad w^*(t) \geq 0 \quad \text{a.e. } t \in [0, 1].$$

By the complementarity condition, we get

$$\int_0^1 w^*(t)(\overline{g}(t) + b\overline{u})\, dt = w^* \circ G(\overline{z}) = 0.$$

We also have $w^*(t)(\overline{g}(t) + b\overline{u}) \leq 0$ a.e. $t \in [0, 1]$. Hence,

(5.7) $$\qquad\qquad w^*(t)(\overline{g}(t) + b\overline{u}) = 0 \quad \text{a.e. } t \in [0, 1].$$

From (5.6) and (5.7), we obtain the complementarity condition of Theorem 2.3. We have

$$\nabla G(\overline{z})^* w^* = w^* \circ \nabla G(\overline{z}) = w^* \circ \nabla G(\overline{x}, \overline{u}) = (w^* \circ G_x(\overline{x}, \overline{u}), w^* \circ G_u(\overline{x}, \overline{u})).$$

From Lemmas 5.1, 5.3 and 5.4, the adjoint equation (a1) is equivalent to

$$\begin{cases} \int_0^1 \overline{L}_x(t)\, dt + \int_0^1 w^*(t)\overline{g}_x(t)\, dt + a - \int_0^1 A^T(t)v(t)\, dt = 0, \\ \int_{(\cdot)}^1 \overline{L}_x(\tau)\, d\tau + \int_{(\cdot)}^1 w^*(\tau)\overline{g}_x(\tau)\, d\tau + v(\cdot) - \int_{(\cdot)}^1 A^T(\tau)v(\tau)\, d\tau = 0, \\ \overline{L}_u(\cdot) + bw^*(\cdot) - B^T(\cdot)v(\cdot) = 0 \end{cases}$$

$$\Longleftrightarrow \begin{cases} a = v(0), \quad v(1) = 0, \\ \int_{(\cdot)}^1 \overline{L}_x(\tau)\, d\tau = \int_{(\cdot)}^1 A^T(\tau)v(\tau)\, d\tau - v(\cdot) - \int_{(\cdot)}^1 w^*(\tau)\overline{g}_x(\tau)\, d\tau, \\ \overline{L}_u(\cdot) + bw^*(\cdot) - B^T(\cdot)v(\cdot) = 0. \end{cases}$$

This implies that

$$v \in W^{1,q}([0,1], \mathbb{R}^n), \quad v(1) = 0, \quad \overline{L}_x(\cdot) = \dot{v}(\cdot) + A^T(\cdot)v(\cdot) - w^*(\cdot)\overline{g}_x(\cdot)$$

and

$$\overline{L}_u(\cdot) - B^T(\cdot)v(\cdot) + bw^*(\cdot) = 0.$$

By Lemma 5.2, the set of Lagrange multipliers $\Lambda(\overline{z}) = \{(w^*, x^*)\}$ is unique and independent of z. Thus, we obtain the adjoint equation and the stationarity condition with respect to u of Theorem 2.3. From the regularity condition of theorem, the linear mapping $\nabla \mathcal{F}(\overline{z})$ is onto. By Lemma 5.6 and [5, Proposition 3.54], the problem (1.1)–(1.4) satisfies the strong extended polyhedricity condition at $\overline{z}$. By Corollary 4.5, we get

$$(5.8) \qquad\qquad \nabla^2 J(\overline{z})zz + \langle w^*, \nabla^2 G(\overline{z})zz \rangle \geq 0, \quad \forall z \in \Theta(\overline{z}).$$

By Lemmas 5.3 and 5.4, we have

$$(5.8) \quad \Longleftrightarrow \quad \int_0^1 \left[\overline{L}_{xx}(t)x^2(t) + 2\overline{L}_{xu}(t)x(t)u(t) + \overline{L}_{uu}(t)u^2(t) \right] dt$$

$$+ \int_0^1 w^*(t)\overline{g}_{xx}(t)x^2(t)\, dt \geq 0, \quad \forall (x,u) \in \Theta(\overline{x}, \overline{u});$$

this is the non-negative second-order condition of Theorem 2.3. The proof of Theorem 2.3 is complete. $\qquad\square$

6. Some examples

To illustrate Theorem 2.3, we provide the following examples.

Example 6.1. Let

$$X = W^{1,11/5}([0,1], \mathbb{R}^2), \quad U = L^{11/5}([0,1], \mathbb{R}^2).$$

Consider the problem

$$J(x,u) = \int_0^1 (x_1^2 + u_1^{11/5} + u_2^{11/5})\, dt \longrightarrow \inf$$

(6.1)   subject to $\dot{x}_1 = 2x_1$, $\dot{x}_2 = x_2 + u_2$, $x_1(0) = 0$, $x_2(0) = 1$, $x_1 + u_1 \geq 0$, $u_2 \geq 0$.

Suppose that $(\overline{x}, \overline{u})$ is a locally optimal solution of the problem. Then,

$$\overline{x} = (0, e^t); \quad \overline{u} = (0,0).$$

Indeed, by Example 2.1 the function

$$L(t,x,u) = x_1^2 + u_1^{11/5} + u_2^{11/5}$$

is satisfied the assumption (H1) and

$$L_x(t, x, u) = (2x_1, 0), \quad L_u(t, x, u) = \frac{11}{5}(u_1^{6/5}, u_2^{6/5}), \quad L_{xu}(t, x, u) = L_{ux}(t, x, u) = 0$$

and

$$L_{xx}(t, x, u) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad L_{uu}(t, x, u) = \frac{66}{25} \begin{bmatrix} u_1^{1/5} & 0 \\ 0 & u_2^{1/5} \end{bmatrix}.$$

From (6.1), we have

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad b = -1.$$

It is easy to see that the condition (H2) is also satisfied. Moreover, by (6.1)

$$g(t, x) = (-x_1, 0),$$

we have

$$g_x = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \quad g_{xx}(t, x) = \begin{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{bmatrix}.$$

So, the assumption (H3) is valid. For each $v = (v_1, v_2) \in Q \subset L^{11/5}([0, 1], \mathbb{R}^2)$, there exists $u = (u_1, u_2) = (-v_1, -v_2) \in L^{11/5}([0, 1], \mathbb{R}^2)$ such that

$$\overline{g}_x(\cdot)(x - \overline{x}) + bu + \overline{g}(\cdot) - v = (-x_1 + \overline{x}_1, 0) - 1(u_1, u_2) + (-\overline{x}_1, 0) - (v_1, v_2)$$
$$= (-x_1 - u_1 - v_1, -u_2 - v_2) = (-x_1, 0),$$

where $x_1(t)$ is a solution of equations $\dot{x}_1(t) = 2x_1(t)$ and $x_1(0) = 0$. So $x_1(t) = 0$. Hence,

$$\overline{g}_x(\cdot)(x - \overline{x}) + bu + \overline{g}(\cdot) - v = (0, 0) \in Q.$$

Therefore, the assumptions of Theorem 2.3 are satisfied. Assume that $(\overline{x}, \overline{u})$ is a locally optimal solution of the problem. By Theorem 2.3, there exist $w^* = (w_1^*, w_2^*) \in L^{11/6}([0, 1], \mathbb{R}^2)$ and $v = (v_1, v_2) \in W^{1, 11/6}([0, 1], \mathbb{R}^2)$, $v_1(1) = 0$, $v_2(1) = 0$ such that the following conditions are fulfilled:

(a*) Adjoint equation:

$$\dot{v}_1(\cdot) + 2v_1(\cdot) + w_1^*(\cdot) = 2\overline{x}_1(\cdot) = 0,$$

(6.2)
$$\dot{v}_2(\cdot) + v_2(\cdot) = 0;$$

(b*) Stationarity condition with respect to u:

(6.3)
$$\frac{11}{5}\overline{u}_1(\cdot)^{6/5} - w_1^*(\cdot) = 0,$$

(6.4)
$$\frac{11}{5}\overline{u}_2(\cdot)^{6/5} - v_2(\cdot) - w_2^*(\cdot) = 0;$$

(c*) Non-negative second-order condition:

$$(6.5) \quad \int_0^1 \left[2x_1^2(t) + \frac{66}{25}\overline{u}_1^{1/5} u_1^2(t) + \frac{66}{25}\overline{u}_2^{1/5} u_2^2(t) \right] dt \geq 0 \quad \text{for all } (x,u) \in \Theta(\overline{x},\overline{u});$$

(d*) Complementarity condition:

$$w_1^*(\cdot), w_2^*(\cdot) \geq 0 \quad \text{and} \quad w_1^*(\cdot)(\overline{x}_1(\cdot) + \overline{u}_1(\cdot)) = 0, \quad w_2^*(\cdot)\overline{u}_2(\cdot) = 0.$$

Consider the following four cases:

   Case 1: $w_1^*(t) = 0$ and $w_2^*(t) = 0$ a.e. $t \in [0,1]$. Substituting $w_1^*(t) = 0$ into (6.3), we have

$$\frac{11}{5}\overline{u}_1(\cdot)^{6/5} = 0.$$

This implies that $\overline{u}_1(t) = 0$. Thus,

$$\overline{x}_1(t) = 0, \quad \overline{u}_1(t) = 0.$$

By (6.2) and $v_2(1) = 0$, we have $v_2(t) = 0$. Substituting $w_2^*(t) = 0$ and $v_2(t) = 0$ into (6.4), we get $\overline{u}_2(t) = 0$. From the condition

$$\dot{\overline{x}}_2(t) = \overline{x}_2(t) + \overline{u}_2(t),$$

and $\overline{u}_2(t) = 0$, we have $\overline{x}_2(t) = e^{t+c}$. Since $\overline{x}_2(0) = 1$, we obtain $\overline{x}_2(t) = e^t$. Substituting $\overline{u}_1(t) = 0$ and $\overline{u}_2(t) = 0$ into (6.5), we get

$$\int_0^1 x_1^2(t)\, dt \geq 0, \quad \forall (x,u) \in \Theta(\overline{x},\overline{u}),$$

this is always satisfied. Thus, if $(\overline{x},\overline{u})$ is a locally optimal solution of the problem then

$$\overline{x} = (0, e^t); \quad \overline{u} = (0,0).$$

   Case 2: $w_1^*(t) = 0$ and $\overline{u}_2(t) = 0$ a.e. $t \in [0,1]$. Substituting $w_1^*(t) = 0$ into (6.3), we have $\overline{u}_1(t) = 0$. From the condition

$$\dot{\overline{x}}_2(t) = \overline{x}_2(t) + \overline{u}_2(t)$$

and

$$\overline{u}_2(t) = 0, \quad \overline{x}_2(0) = 1,$$

we have $\overline{x}_2(t) = e^t$. In Case 1, we checked that if $(\overline{x},\overline{u})$ is a locally optimal solution of the problem then

$$\overline{x} = (0, e^t); \quad \overline{u} = (0,0).$$

Case 3: $w_2^*(t) = 0$ and $\overline{x}_1(t) + \overline{u}_1(t) = 0$ a.e. $t \in [0,1]$. Similarly as in Case 1, we can check that if $(\overline{x}, \overline{u})$ is a locally optimal solution of the problem then

$$\overline{x} = (0, e^t); \quad \overline{u} = (0,0).$$

Case 4: $\overline{x}_1(t) + \overline{u}_1(t) = 0$ and $\overline{u}_2(t) = 0$ a.e. $t \in [0,1]$. Since $\overline{x}_1(t) = 0$, we have $\overline{u}_1(t) = 0$. In Case 3, we checked that if $(\overline{x}, \overline{u})$ is a locally optimal solution of the problem then

$$\overline{x} = (0, e^t); \quad \overline{u} = (0,0).$$

The following example shows that if the second-order necessary condition is not satisfied, then the admissible couple is not solution even it satisfies first-necessary conditions.

Example 6.2. Let

$$X = W^{1,3}([0,1], \mathbb{R}^2), \quad U = L^3([0,1], \mathbb{R}^2).$$

Consider the problem

$$J(x, u) = \int_0^1 \left(u_1^3 - \sqrt{2(1 + u_1^6)} + u_2^3 \right) dt \longrightarrow \inf$$

(6.6)   subject to $\dot{x}_1 = 2x_1$, $\dot{x}_2 = x_2 + u_2$, $x_1(0) = 1$, $x_2(0) = 1$, $x_1 + u_1 \geq 0$, $u_2 \geq 0$.

Suppose that $(\overline{x}, \overline{u})$ is a locally optimal solution of the problem. Then, by first-order optimality conditions, we obtain

$$\overline{x} = (e^{2t}, e^t); \quad \overline{u} = (1,0),$$

or

$$\overline{x} = (e^{2t}, e^t); \quad \overline{u} = (0,0),$$

or

$$\overline{x} = (e^{2t}, e^t); \quad \overline{u} = (-e^{2t}, 0).$$

If we let

$$\overline{x}^1 = (e^{2t}, e^t); \quad \overline{u}^1 = (1,0),$$

or

$$\overline{x}^2 = (e^{2t}, e^t); \quad \overline{u}^2 = (-e^{2t}, 0)$$

then $(\overline{x}^1, \overline{u}^1)$ and $(\overline{x}^2, \overline{u}^2)$ do not satisfy the second-order optimality conditions. Hence, $(\overline{x}^1, \overline{u}^1)$ and $(\overline{x}^2, \overline{u}^2)$ are not locally optimal solutions of the problem. Thus, if $(\overline{x}; \overline{u})$ is a locally optimal solution of the problem, then

$$\overline{x} = (e^{2t}, e^t); \quad \overline{u} = (0,0).$$

Indeed, by Example 2.2 the function

$$L(t, x, u) = u_1^3 - \sqrt{2(1 + u_1^6)} + u_2^3$$

is satisfied the assumption (H1) and

$$L_u(t, x, u) = 3\left(u_1^2 - \sqrt{2}\frac{u_1^5}{\sqrt{1 + u_1^6}}, u_2^2\right),$$

$$L_x(t, xu) = 0, \quad L_{xu}(t, x, u) = L_{ux}(t, x, u) = L_{xx}(t, x, u) = 0,$$

$$L_{uu}(t, x, u) = 3\begin{bmatrix} 2u_1 - \sqrt{2}\frac{2u_1^{10} + 5u_1^4}{(1 + u_1^6)\sqrt{1 + u_1^6}} & 0 \\ 0 & 2u_2 \end{bmatrix}.$$

From (6.6), we have

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad b = -1.$$

It is easy to see that the condition (H2) is also satisfied. Similarly as in Example 6.1, we can prove that the conditions (H3) and (H4) are also fulfilled. Hence, the assumptions of Theorem 2.3 are satisfied. Assume that $(\overline{x}, \overline{u})$ is a locally optimal solution of the problem. By Theorem 2.3, there exist $w^* = (w_1^*, w_2^*) \in L^{3/2}([0, 1], \mathbb{R}^2)$ and $v = (v_1, v_2) \in W^{1,3/2}([0, 1], \mathbb{R}^2)$, $v_1(1) = 0$, $v_2(1) = 0$ such that the following conditions are fulfilled:

(a_1^*) Adjoint equation:

$$\dot{v}_1(\cdot) + 2v_1(\cdot) + w_1^*(\cdot) = 0,$$

(6.7)
$$\dot{v}_2(\cdot) + v_2(\cdot) = 0;$$

(b_1^*) Stationarity condition with respect to u:

(6.8)
$$3\overline{u}_1^2(\cdot) - \sqrt{2}\frac{3\overline{u}_1^5(\cdot)}{\sqrt{1 + \overline{u}_1^6(\cdot)}} - w_1^*(\cdot) = 0,$$

(6.9)
$$3\overline{u}_2^2(\cdot) - v_2(\cdot) - w_2^*(\cdot) = 0;$$

(c_1^*) Non-negative second-order condition:

(6.10)
$$\int_0^1 \left[\left(6\overline{u}_1(t) - \sqrt{2}\frac{6\overline{u}_1^{10}(t) + 15\overline{u}_1^4(t)}{(1 + \overline{u}_1^6(t))\sqrt{1 + \overline{u}_1^6(t)}}\right)u_1^2(t) + 6\overline{u}_2(t)u_2^2(t)\right] dt \geq 0$$

for all $(x, u) \in \Theta(\overline{x}, \overline{u})$;

(d_1^*) Complementarity condition:

$$w_1^*(\cdot), w_2^*(\cdot) \geq 0 \quad \text{and} \quad w_1^*(\cdot)(\overline{x}_1(\cdot) + \overline{u}_1(\cdot)) = 0, \quad w_2^*(\cdot)\overline{u}_2(\cdot) = 0.$$

Consider the following four cases:

Case 1: $w_1^*(t) = 0$ and $w_2^*(t) = 0$ a.e. $t \in [0, 1]$. Substituting $w_1^*(t) = 0$ into (6.8), we have

$$3\overline{u}_1^2(t) - \sqrt{2}\frac{3\overline{u}_1^5(t)}{\sqrt{1 + \overline{u}_1^6(t)}} = 0.$$

This implies that $\overline{u}_1(t) = 0$, or $\overline{u}_1(t) = 1$. By (6.6),

$$\dot{\overline{x}}_1(t) = 2\overline{x}_1(t) \quad \Longleftrightarrow \quad \overline{x}_1(t) = e^{2t+c}.$$

Combining this with the condition $\overline{x}_1(0) = 1$, we get $c = 0$. So $\overline{x}_1(t) = e^{2t}$. Thus

$$\overline{x}_1(t) = e^{2t}, \quad \overline{u}_1(t) = 0,$$

or

$$\overline{x}_1(t) = e^{2t}, \quad \overline{u}_1(t) = 1.$$

By (6.7) and $v_2(1) = 0$, we have $v_2(t) = 0$. Substituting $w_2^*(t) = 0$ and $v_2(t) = 0$ into (6.9), we get $\overline{u}_2(t) = 0$. From the condition

$$\dot{\overline{x}}_2(t) = \overline{x}_2(t) + \overline{u}_2(t),$$

and $\overline{u}_2(t) = 0$, $x_2(0) = 1$ we have $\overline{x}_2(t) = e^t$. Thus, if $(\overline{x}, \overline{u})$ is a locally optimal solution of the problem, then by first-order optimality conditions, we obtain

$$\overline{x} = (e^{2t}, e^t); \quad \overline{u} = (0, 0),$$

or

$$\overline{x} = (e^{2t}, e^t); \quad \overline{u} = (1, 0).$$

Substituting $\overline{u}_1(t) = 1$ and $\overline{u}_2(t) = 0$ into (6.10), we get

(6.11) $$\int_0^1 -\frac{9}{2}u_1^2(t)\, dt \geq 0.$$

But, (6.11) is not fulfilled if $u_1 = 1$ and $(x, u) \in \Theta(\overline{x}, \overline{u})$. Hence,

$$\overline{x} = (e^{2t}, e^t); \quad \overline{u} = (1, 0)$$

is not a locally optimal solution of the problem. Substituting

$$\overline{u}_1(t) = 0 \quad \text{and} \quad \overline{u}_2(t) = 0$$

into (6.10), we obtain

$$\int_0^1 0\, dt \geq 0,$$

this is always satisfied. Thus, if $(\overline{x}, \overline{u})$ is a locally optimal solution of the problem then

$$\overline{x} = (e^{2t}, e^{t}); \quad \overline{u} = (0, 0).$$

Case 2: $w_1^*(t) = 0$ and $\overline{u}_2(t) = 0$ a.e. $t \in [0, 1]$. Similarly as in Case 1, we can prove that if $(\overline{x}, \overline{u})$ is a locally optimal solution of the problem, then

$$\overline{x} = (e^{2t}, e^{t}); \quad \overline{u} = (0, 0).$$

Case 3: $w_2^*(t) = 0$ and $\overline{x}_1(t) + \overline{u}_1(t) = 0$ a.e. $t \in [0, 1]$. Since $\overline{x}_1(t) = e^{2t}$, we have $\overline{u}_1(t) = -e^{2t}$. By (6.7) and $v_2(1) = 0$, we get $v_2(t) = 0$. Substituting $w_2^*(t) = 0$ and $v_2(t) = 0$ into (6.9), we get $\overline{u}_2(t) = 0$. From the condition

$$\dot{\overline{x}}_2(t) = \overline{x}_2(t) + \overline{u}_2(t),$$

and $\overline{u}_2(t) = 0$, $\overline{x}_2(0) = 1$, we have $\overline{x}_2(t) = e^t$. Substituting $\overline{u}_1(t) = -e^{2t}$ and $\overline{u}_2(t) = 0$ into (6.10), we obtain

(6.12)
$$\int_0^1 \left(-6e^{-2t} - \sqrt{2} \frac{6e^{-20t} + 15e^{-8t}}{(1 + e^{-12t})\sqrt{1 + e^{-12t}}} \right) u_1^2(t) \, dt \geq 0.$$

But, (6.12) is not fulfilled if $u_1 = -1$ and $(x, u) \in \Theta(\overline{x}, \overline{u})$. Hence,

$$\overline{x} = (e^{2t}, e^{t}); \quad \overline{u} = (-e^{2t}, 0)$$

is not a locally optimal solution of the problem.

Case 4: $\overline{x}_1(t) + \overline{u}_1(t) = 0$ and $\overline{u}_2(t) = 0$ a.e. $t \in [0, 1]$. Since $\overline{x}_1(t) = e^{2t}$, we have $\overline{u}_1(t) = -e^{2t}$. Similarly as in Case 3, we can show that

$$\overline{x} = (e^{2t}, e^{t}); \quad \overline{u} = (-e^{2t}, 0)$$

is not a locally optimal solution of the problem.

7. Conclusions

We studied second-order necessary optimality conditions for an optimal control problem with nonconvex cost functions and state-control constraints. In order to achieve these conditions, we first established an abstract result on the second-order necessary optimality conditions for a mathematical programming problem and then we derived the second-order necessary optimality conditions for an optimal control problem. The main result of this paper is illustrated by two examples.

Acknowledgments

References

[1] J.-P. Aubin and H. Frankowska, *Set-valued Analysis*, Systems & Control: Foundations & Applications **2**, Birkhäuser Boston, Boston, MA, 1990.

[2] A. Ben-Tal, *Second-order and related extremality conditions in nonlinear programming*, J. Optim. Theory Appl. **31** (1980), no. 2, 143–165.

[3] J. F. Bonnans, R. Cominetti and A. Shapiro, *Second order optimality conditions based on parabolic second order tangent sets*, SIAM J. Optim. **9** (1999), no. 2, 466–492.

[4] J. F. Bonnans, C. de la Vega and X. Dupuis, *First- and second-order optimality conditions for optimal control problems of state constrained integral equations*, J. Optim. Theory Appl. **159** (2003), no. 1, 1–40.

[5] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research, Springer-Verlag, New York, 2000.

[6] E. Casas, J. C. de los Reyes and F. Tröltzsch, *Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints*, SIAM J. Optim. **19** (2008), no. 2, 616–643.

[7] A. Cernea and H. Frankowska, *A connection between the maximum principle and dynamic programming for constrained control problems*, SIAM J. Control Optim. **44** (2005), no. 2, 673–703.

[8] N. H. Chieu, B. T. Kien and N. T. Toan, *Further results on subgradients of the value function to a parametric optimal control problem*, J. Optim. Theory Appl. **168** (2016), no. 3, 785–801.

[9] R. Cominetti, *Metric regularity, tangent sets, and second-order optimality conditions*, Appl. Math. Optim. **21** (1990), no. 3, 265–287.

[10] A. L. Dontchev, *Perturbations, Approximations and Sensitivity Analysis of Optimal Control Systems*, Lecture Notes in Control and Information Sciences **52**, Springer-Verlag, Berlin, 1983.

[11] H. Frankowska, D. Hoehener and D. Tonon, *A second-order maximum principle in optimal control under state constraints*, Serdica Math. J. **39** (2013), no. 3-4, 233–270.

[12] H. Frankowska and D. Tonon, *Pointwise second-order necessary optimality conditions for the Mayer problem with control constraints*, SIAM J. Control Optim. **51** (2013), no. 5, 3814–3843.

[13] R. Henrion, B. S. Mordukhovich and N. M. Nam, *Second-order analysis of polyhedral systems in finite and infinite dimensions with applications to robust stability of variational inequalities*, SIAM J. Optim. **20** (2010), no. 5, 2199–2227.

[14] R. Hilscher and V. Zeidan, *Second order sufficiency criteria for a discrete optimal control problem*, J. Difference Equ. Appl. **8** (2002), no. 6, 573–602.

[15] _____, *Discrete optimal control: Second order optimality conditions*, J. Difference Equ. Appl. **8** (2002), no. 10, 875–896.

[16] D. Hoehener, *Variational approach to second-order optimality conditions for control problems with pure state constraints*, SIAM J. Control Optim. **50** (2012), no. 3, 1139–1173.

[17] A. D. Ioffe, *Necessary and sufficient conditions for a local minimum III: Second order conditions and augmented duality*, SIAM J. Control Optim. **17** (1979), no. 2, 266–288.

[18] A. D. Ioffe and V. M. Tihomirov, *Theory of Extremal Problems*, Studies in Mathematics and its Applications **6**, North-Holland, Amsterdam-New York, 1979.

[19] H. Kawasaki, *An envelope-like effect of infinitely many inequality constraints on second-order necessary conditions for minimization problems*, Math. Programming **41** (1988), no. 1, (Ser. A), 73–96.

[20] B. T. Kien and V. H. Nhu, *Second-order necessary optimality conditions for a class of semilinear elliptic optimal control problems with mixed pointwise constraints*, SIAM J. Control Optim. **52** (2014), no. 2, 1166–1202.

[21] B. T. Kien, V. H. Nhu and N. H. Son, *Second-order optimality conditions for a semilinear elliptic optimal control problem with mixed pointwise constraints*, Set-Valued Var. Anal. **25** (2017), no. 1, 177–210.

[22] B. T. Kien, N. T. Toan, M. M. Wong and J.-C. Yao, *Lower semicontinuity of the solution set to a parametric optimal control problem*, SIAM J. Control Optim. **50** (2012), no. 5, 2889–2906.

[23] L. Li, Y. Gao and H. Wang, *Second order sufficient optimality conditions for hybrid control problems with state jump*, J. Ind. Manag. Optim. **11** (2015), no. 1, 329–343.

[24] K. Malanowski, H. Maurer and S. Pickenhain, *Second-order sufficient conditions for state-constrained optimal control problems*, J. Optim. Theory Appl. **123** (2004), no. 3, 595–617.

[25] O. L. Mangasarian and T. H. Shiau, *Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems*, SIAM J. Control Optim. **25** (1987), no. 3, 583–595.

[26] B. Marinković, *Optimality conditions for discrete optimal control problems*, Optim. Methods Softw. **22** (2007), no. 6, 959–969.

[27] ———, *Optimality conditions for discrete optimal control problems with equality and inequality type constraints*, Positivity **12** (2008), no. 3, 535–545.

[28] ———, *Second-order optimality conditions in a discrete optimal control problem*, Optimization **57** (2008), no. 4, 539–548.

[29] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation I: Basis Theory*, Grundlehren der Mathematischen Wissenschaften **330**, Springer-Verlag, Berlin, 2006.

[30] ———, *Variational Analysis and Generalized Differentiation II: Applications*, Grundlehren der Mathematischen Wissenschaften **331**, Springer-Verlag, Berlin, 2006.

[31] M. Moussaoui and A. Seeger, *Epsilon-maximum principle of Pontryagin type and perturbation analysis of convex optimal control problems*, SIAM J. Control Optim. **34** (1996), no. 2, 407–427.

[32] Z. Páles and V. M. Zeidan, *Nonsmooth optimum problems with constraints*, SIAM J. Control Optim. **32** (1994), no. 5, 1476–1502.

[33] ———, *Optimum problems with certain lower semicontinuous set-valued constraints*, SIAM J. Control Optim. **8** (1998), no. 3, 707–727.

[34] ———, *Characterization of L^1-closed decomposable sets in L^∞*, J. Math. Anal. Appl. **238** (1999), no. 2, 491–515.

[35] ———, *Optimal control problems with set-valued control and state constraints*, SIAM J. Optim. **14** (2003), no. 2, 334–358.

[36] J.-P. Penot, *Optimality conditions in mathematical programming and composite opti-mization*, Math. Programming **67** (1994), no. 2, 225–245.

[37] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Grundlehren der Mathe-matischen Wissenschaften **317**, Springer-Verlag, Berlin, 1998.

[38] R. T. Rockafellar and P. R. Wolenski, *Convexity in Hamilton-Jacobi theory I: Dy-namics and duality*, SIAM J. Control Optim. **39** (2000), no. 5, 1323–1350.

[39] _____, *Convexity in Hamilton-Jacobi theory II: Envelope representations*, SIAM J. Control Optim. **39** (2000), no. 5, 1351–1372.

[40] L. Q. Thuy and N. T. Toan, *Subgradients of the value function in a parametric convex optimal control problem*, J. Optim. Theory Appl. **170** (2016), no. 1, 43–64.

[41] L. Q. Thuy, B. T. Thanh and N. T. Toan, *On the no-gap second-order optimality conditions for a discrete optimal control problem with mixed constraints*, J. Optim. Theory Appl. **173** (2017), no. 2, 421–442.

[42] N. T. Toan, *Mordukhovich subgradients of the value function in a parametric optimal control problem*, Taiwanese J. Math. **19** (2015), no. 4, 1051–1072.

[43] N. T. Toan, Q. H. Ansari and J.-C. Yao, *Second-order necessary optimality conditions for a discrete optimal control problem*, J. Optim. Theory Appl. **165** (2015), no. 3, 812–836.

[44] N. T. Toan and B. T. Kien, *Subgradients of the value function to a parametric optimal control problem*, Set-Valued Var. Anal. **18** (2010), no. 2, 183–203.

[45] N. T. Toan and L. Q. Thuy, *Second-order necessary optimality conditions for a dis-crete optimal control problem with mixed constraints*, J. Global Optim. **64** (2016), no. 3, 533–562.

[46] J. Zowe and S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim. **5** (1979), no. 1, 49–62.

Toan Nguyen and Thuy Le

School of Applied Mathematics and Informatics Hanoi University of Science and Technology, 1 Dai Co Viet, Hanoi, Vietnam

E-mail address: `toan.nguyenthi@hust.edu.vn`, `thuy.lequang@hust.edu.vn`