TECHNISCHE UNIVERSITÄT WIEN

Operations Research and Control Systems

SWM ORCOS

On the strong subregularity of the optimality mapping in mathematical programming and calculus of variations

Nikolai Osmolovskii, Vladimir Veliov

Research Report 2020-12

August 2020

ISSN 2521-313X

Operations Research and Control Systems
Institute of Statistics and Mathematical Methods in Economics
Vienna University of Technology

Research Unit ORCOS
Wiedner Hauptstraße 8 / E105-04
1040 Vienna, Austria
E-mail: orcos@tuwien.ac.at

On the strong subregularity of the optimality mapping in mathematical programming and calculus of variations*

N.P. Osmolovskii† V.M. Veliov‡

Abstract

The paper presents sufficient conditions for a strong metric subregularity (SMSr) property of the optimality mapping associated with the Pontryagin local maximum principle for a Mayer's type optimal control problem with general initial/terminal constraints for the state variable and unconstrained control. This SMSr property is adapted to the involvement of two norms in the basic assumptions: smoothness, constraint qualification, and strong second order sufficient optimality conditions. The proofs are based on a new abstract result for strong metric subregularity (in a two-norms setting) of the Karush-Kuhn-Tucker optimality mapping for a mathematical programming problem in a Banach space, also presented in the paper.

Keywords: optimization, mathematical programming, optimal control, Mayer's problem, metric subregularity

AMS Classification: 90C48, 49K40

1 Introduction

This paper investigates a subregularity property related to the optimization problem

$$\min\ f_0(x), \quad g(x) = 0, \quad f_i(x) \le 0, \quad i = 1, \dots, k, \tag{1}$$

where $f_0 : X \to \mathbb{R}$, $g : X \to Y$, $f_i : X \to \mathbb{R}$, for $i = 1, \dots, k$ and X and Y are Banach spaces. The following system of equations and inequalities is known as Karush-Kuhn-Tucker (KKT) system associated with problem (1):

$$\begin{aligned}
&f_0'(x) + \sum_{i=1}^{k} \alpha_i f_i'(x) + (g'(x))^* y^* = 0,\\
&g(x) = 0,\\
&\alpha_i f_i(x) = 0, \quad i = 1, \dots, k,\\
&f_i(x) \le 0, \quad \alpha_i \ge 0, \quad i = 1, \dots, k,
\end{aligned}$$

where $x \in X$, $y^* \in Y^*$ (Y^* denotes the dual space to Y), and $\alpha := (\alpha_1, \dots, \alpha_k) \in \mathbb{R}^k$. Moreover, "primes" indicate Fréchet derivatives (assuming that these exist), and $(g'(x))^* : Y^* \to X^*$ is the adjoint of the continuous linear operator $g'(x) : X \to Y$.

*This research is supported by the Austrian Science Foundation (FWF) under grant No P31400-N32.

†Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland osmolov@ibspan.waw.pl

‡Institute of Statistics and Mathematical Methods in Economics, TU Wien, Austria, vladimir.veliov@tuwien.ac.at

Under an additional condition called *Mangasarian-Fromovitz constraint qualification*, the existence of a pair $(y^*, \alpha) \in Y^* \times \mathbb{R}^k$, such that the KKT system is fulfilled, is a necessary condition for $x \in X$ to be a local solution of problem (1). The relations in the last two lines of the KKT system can be equivalently rewritten as

$$f(x) \in N_{\mathbb{R}^k_+}(\alpha),$$

where $f = (f_1, \dots, f_k)$, $\mathbb{R}^k_+$ is the set of all elements of $\mathbb{R}^k$ with non-negative components, and the normal cone to the set $\mathbb{R}^k_+$ is defined as usual:

$$N_{\mathbb{R}^k_+}(\alpha) := \begin{cases} \{\lambda \in \mathbb{R}^k \, : \, \langle \lambda, \beta - \alpha \rangle \le 0 \text{ for all } \beta \in \mathbb{R}^k_+\} & \text{if } \alpha \in \mathbb{R}^k_+, \\ \emptyset & \text{if } \alpha \notin \mathbb{R}^k_+, \end{cases}$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbb{R}^k$. Consequently, one can reformulate the KKT system as

$$F(x, y^*, \alpha) := \begin{pmatrix} f_0'(x) + \sum_{i=1}^k \alpha_i f_i'(x) + (g'(x))^* y^* \\ g(x) \\ f(x) - N_{\mathbb{R}^k_+}(\alpha) \end{pmatrix} \ni 0. \tag{2}$$

Therefore, the set-valued mapping $F : X \times Y^* \times \mathbb{R}^k \rightrightarrows X^* \times Y \times \mathbb{R}^k$ is called *KKT optimality mapping*, see e.g. [9] and [11, p. 134].

The property generally called *Strong Metric Subegularity* (SMSr) (see e.g. [11, Chapter 3.9] and [4]) of the mapping F is of key importance in the qualitative and numerical analysis of optimization problems admitting a formulation as (1). In particular, versions of this property (sometimes appearing under different names) are widely used for obtaining error estimates for numerical methods for variational inequalities and optimal control problems, such as gradient methods, Newton-type methods, etc. (see e.g. [2], [26, Section 5], [4, Subsection 7.3], [1], [18] among others). The SMSr property of the KKT mapping for finite-dimensional mathematical programming problems is characterized in [9, Theorem 2.6] and [4, Theorem 7.1]. The characterization involves *strict Mangasarian-Fromovitz condition* and *strong second order sufficient optimality condition*, the latter also called *coercivity condition*, and requires appropriate differentiability properties for the data f_0, f, g.

However, the targeted application in this paper is in calculus of variations and optimal control. It is well known (after the works [23, 21, 22]) that in the optimal control context the norm in which the coercivity condition has to be posed (usually L^2 for the controls) is so weak that the differentiability assumption in this norm is rather restrictive (the so called *norm discrepancy*). For this reason, the *two norm approach* was developed in [21, 22] and subsequent publications, in which differentiability is assumed with respect to a stronger norm than the one in which the coercivity is required. This approach is used for studying Lipschitz dependence of the solutions of optimal control problems on a parameter. Among the large literature on this subject we mention the path-breaking papers [21, 7, 22, 8]. We mention that the Lipschitz dependence of the solution on a parameter is related to the property of strong metric regularity, which is a subject of a huge number of publications (cf. [11, 15]).

In this paper we present a sufficient condition for strong metric subregularity of the KKT mapping in (2) using the two norm approach. The property of strong metric subregularity is weaker than that of stronger metric regularity, therefore the sufficient conditions for SMSr are weaker that those in [21]. In addition we mention that although state constraints are considered in [21], their specific form does not cover the case of initial and terminal constraints which are in the focus of the present paper.

The obtained in this paper abstract subregularity result is applied to the following optimal control problem of Mayer's type on a fixed time interval $[t_0, t_1]$:

$$\text{minimize} \quad \varphi_0(x(t_0), x(t_1)) \tag{3}$$

subject to

$$\dot{x}(t) = h(x(t), u(t)), \qquad t \in [t_0, t_1], \tag{4}$$

$$\psi_j(x(t_0), x(t_1)) = 0, \quad j = 1, \ldots, s, \tag{5}$$

$$\varphi_i(x(t_0), x(t_1)) \le 0, \quad i = 1, \ldots, k, \tag{6}$$

where $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $\psi_j : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $j = 1, \ldots, s$, $\varphi_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $i = 0, \ldots, k$.

We stress that control constraints are not involved, therefore we actually deal with a *general problem of calculus of variations*. The strict Mangasarian-Fromovitz condition and the coercivity condition take specific forms, in which two norms are used similarly as in [21]. The main novelty of the strong subregularity result is that the coercivity condition is posed on a *critical cone* with initial/terminal constraints for the state variable.

The paper is organized as follows. In Section 2 we present a basic theorem about stability of the SMSr property adapted to the consideration of two norms. Needed properties of systems of linear inequalities and equations in Banach spaces are also provided. Sufficient conditions for SMSr of the KKT mapping for a mathematical programming problem in a Banach space setting (with equality constraints and a finite number of inequality constraints) are presented in Section 3. Section 4 deals with the optimal control problem (3)–(6). Here we formulate and prove the main results in the paper: the sufficient conditions of the SMSr property (in the two-norm setting) for two optimality mappings associated with problem (3)–(6): first, for the optimality mapping in Lagrangian format, then for the optimality mapping arising in the Pontryagin local maximum principle.

2 Preliminaries

2.1 Strong metric subregularity

Let $\mathcal{X}$ and $\mathcal{Z}$ be two normed linear spaces with norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{Z}}$. Let also a second norm be defined in $\mathcal{X}$, denoted by $\|\cdot\|'_{\mathcal{X}}$, such that $\|s\|'_{\mathcal{X}} \le \|s\|_{\mathcal{X}}$ for every $s \in \mathcal{X}$. Denote by $\mathbb{B}_{\mathcal{Z}}(z; r)$ the ball of radius r in $\mathcal{Z}$, centered at z. Let $\mathcal{L} : \mathcal{X} \rightrightarrows \mathcal{Z}$ be a set-valued mapping. As usual, $\mathrm{gr}(\mathcal{L}) := \{(s, z) \in \mathcal{X} \times \mathcal{Z} \,:\, z \in \mathcal{L}(s)\}$ is the graph of $\mathcal{L}$. The following property is a modification of the strong metric subregularity property, which will be abbreviated as SMSr2.

Definition 2.1 *The mapping $\mathcal{L}$ has the property SMSr2 at $\hat{s}$ for $\hat{z}$ if $(\hat{s}, \hat{z}) \in \mathrm{gr}(\mathcal{L})$ and there exist neighborhoods $\mathcal{O}_{\hat{s}} \ni \hat{s}$ (in the norm $\|\cdot\|_{\mathcal{X}}$), $\mathcal{O}_{\hat{z}} \ni \hat{z}$ and a number κ such that the relations*

$$s \in \mathcal{O}_{\hat{s}}, \quad z \in \mathcal{O}_{\hat{z}}, \quad z \in \mathcal{L}(s)$$

imply that

$$\|s - \hat{s}\|'_{\mathcal{X}} \le \kappa \|z - \hat{z}\|_{\mathcal{Z}}.$$

The property SMSr2 has the good feature that it is preserved under perturbations with a sufficiently small Lipschitz constant at $\hat{s}$ (more precisely, sufficiently small calmness constant at $\hat{s}$, see [10, Theorem 3I.7]). For strongly metrically subregular mappings this property appeared first in [6, Theorem 3.2], see also [4, Theorem 2.1]. We give a proof of a corresponding to Definition 2.1 modification of this stability property.

Theorem 2.1 *Let $\mathcal{L}$ has the property SMSr2 at $\hat{s}$ for $\hat{z}$ with neighborhoods $\mathcal{O}_{\hat{s}}$ (in the norm $\|\cdot\|_{\mathcal{X}}$), $\mathcal{O}_{\hat{z}}$ and parameter $\hat{\kappa}$. Let the numbers $b > 0$, $\mu \geq 0$ and $\kappa \geq 0$ satisfy the relations*

$$\hat{\kappa}\mu < 1, \quad \kappa \geq \frac{\hat{\kappa}}{1 - \hat{\kappa}\mu}, \quad I\!\!B_{\mathcal{Z}}(\hat{z}; 2b) \subset \mathcal{O}_{\hat{z}}. \tag{7}$$

Then for every function $\varphi : \mathcal{X} \to \mathcal{Z}$ such that

$$\|\varphi(s) - \varphi(\hat{s})\|_{\mathcal{Z}} \leq b \quad \text{and} \quad \|\varphi(s) - \varphi(\hat{s})\|_{\mathcal{Z}} \leq \mu \|s - \hat{s}\|'_{\mathcal{X}} \qquad \forall\, s \in \mathcal{O}_{\hat{s}},$$

the mapping $\varphi+\mathcal{L}$ has the property SMSr2 at $\hat{s}$ for $\hat{z}+\varphi(\hat{s})$ with neighborhoods $\mathcal{O}_{\hat{s}}$, $I\!\!B_{\mathcal{Z}}(\hat{z}+\varphi(\hat{s})); b)$ and constant κ.

Proof. Let $s \in \mathcal{O}_{\hat{s}}$, $z \in I\!\!B_{\mathcal{Z}}(\hat{z} + \varphi(\hat{s}); b)$ and $z \in \varphi(s) + \mathcal{L}(s)$. Then

$$\|(z - \varphi(s)) - \hat{z}\|_{\mathcal{Z}} \leq \|z - (\hat{z} + \varphi(\hat{s}))\|_{\mathcal{Z}} + \|\varphi(s) - \varphi(\hat{s})\|_{\mathcal{Z}} \leq b + b.$$

Thus $z - \varphi(s) \in \mathcal{O}_{\hat{z}}$ and, moreover, $z - \varphi(s) \in \mathcal{L}(s)$. According to the SMSr2 property of $\mathcal{L}$,

$$\begin{aligned} \|s - \hat{s}\|'_{\mathcal{X}} &\leq \hat{\kappa}\|(z - \varphi(s)) - \hat{z}\|_{\mathcal{Z}} \leq \hat{\kappa}\|z - (\hat{z} + \varphi(\hat{s}))\|_{\mathcal{Z}} + \hat{\kappa}\|\varphi(s) - \varphi(\hat{s})\|_{\mathcal{Z}} \\ &\leq \hat{\kappa}\|z - (\hat{z} + \varphi(\hat{s}))\|_{\mathcal{Z}} + \hat{\kappa}\mu\|s - \hat{s}\|'_{\mathcal{X}}, \end{aligned}$$

hence

$$\|s - \hat{s}\|'_{\mathcal{X}} \leq \frac{\hat{\kappa}}{1 - \hat{\kappa}\mu}\|z - (\hat{z} + \varphi(\hat{s}))\|_{\mathcal{Z}}.$$

□

Remark 2.1 Observe that the norm $\|\cdot\|_{\mathcal{X}}$ is involved in Theorem 2.1 only for restricting the solutions s appearing there to be close to the reference solution $\hat{s}$ in this bigger norm. This is needed for the application of the main result in this paper in optimal control.

2.2 Systems of equations and inequalities.

Let X be a Banach space and let X^* be its dual space. The norm in each of these spaces will be denoted by $\|\cdot\|$, but sometimes for more clarity we use $\|\cdot\|_X$, $\|\cdot\|_Y$, $\|\cdot\|_{X^*}$, etc. The value of $l \in X^*$ applied to $x \in X$ will be denoted by $l(x)$ or just by lx.

We also consider another Banach space Y and a linear continuous operator $A : X \to Y$. As usual, we denote by $A^* : Y^* \to X^*$ its dual operator, so that $(A^*y^*)(x) = y^*(Ax)$, therefore we use the notation y^*A for A^*y^*. We also use the notation $\mathbb{R}^k_+ = \{\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathbb{R}^k : \ \alpha_i \geq 0 \; \forall\, i = 1, \ldots, k\}$.

Definition 2.2 The functionals $l_1, \ldots, l_k \in X^*$ are *positively independent on* a subspace $L \subset X$ if the conditions $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathbb{R}^k_+$ and $\sum_{i=1}^k \alpha_i l_i(x) = 0 \;\forall\, x \in L$ imply that $\alpha = 0$.

Denote by L^* the set of functionals $x^* \in X^*$ vanishing on a subspace $L \subset X$. If $AX = Y$, then, as known, $(\ker A)^* = A^* Y^*$, see e.g. [5, Lemma 3.6].

Lemma 2.1 *The following two properties are equivalent:*
(i) $AX = Y$, and the the functionals $l_1, \ldots, l_k \in X^*$ are positively independent on $\ker A$*;*
(ii) AX is closed, and the relations

$$y^* \in Y^*, \;\; \alpha \in \mathbb{R}^k_+ \;\; \textit{and} \;\; y^* A + \sum_{i=1}^k \alpha_i l_i = 0 \tag{8}$$

imply that $y^* = 0$ and $\alpha = 0$.

Proof. 1. If (i) holds, then AX is closed. If (8) is fulfilled then for every $x \in \ker A$ we have $\sum_{i=1}^k \alpha_i l_i(x) = 0$. From the second part of (i) we obtain that $\alpha = 0$. The equalities $y^* A = 0$ and $AX = Y$ imply that $y^* = 0$.

2. If (ii) holds and A is not surjective, then due to the closedness of AX, there exists a non-zero $y^* \in Y^*$ such that $y^* A = 0$. The second part of (ii) with $\alpha = 0$ leads to the contradiction $y^* = 0$. Thus $AX = Y$. Now let $\alpha \in \mathbb{R}^k_+$ and $\sum_{i=1}^k \alpha_i l_i = 0$ on $\ker A$. Then $\sum_{i=1}^k \alpha_i l_i \in (\ker A)^*$ and since $AX = Y$, there exists $y^* \in Y^*$ such that $\sum_{i=1}^k \alpha_i l_i = -y^* A$. Then from the second part of (ii) we obtain that $y^* = 0$ and $\alpha = 0$. $\square$

Proposition 2.1 *If the linear and continuous operator $A : X \to Y$ is surjective and the functionals $l_1, \ldots, l_k \in X^*$ are positively independent on* $\ker A$*, then there exists a constant $c > 0$ such that*

$$\|y^* A + \sum_{i=1}^k \alpha_i l_i\| \geq c \left(\|y^*\| + \sum_{i=1}^k \alpha_i \right) \qquad \forall\, y^* \in Y^* \;\; \textit{and} \;\; \forall\, \alpha \in \mathbb{R}^k_+. \tag{9}$$

Proof. Since the inequality in (9) is positively homogeneous, it suffices to prove it for pairs $(y^*, \alpha) \in Y^* \times \mathbb{R}^k_+$ such that $\|y^*\| + \sum_{i=1}^k \alpha_i = 1$. Suppose that the proposition is not true. Then there is a sequence $(y^*_n, \alpha^n) \in Y^* \times \mathbb{R}^k_+$, where $\alpha^n = (\alpha_{n1}, \ldots, \alpha_{nk})$, such that $\|y^*_n\| + \sum_{i=1}^k \alpha_{ni} = 1$ and $\|y^*_n A + \sum_{i=1}^k \alpha_{ni} l_i\| \to 0$ with $n \to \infty$. Without loss of generality we can assume that $\alpha^n \to \alpha \in \mathbb{R}^k_+$. Then $\|y^*_n A + \sum_{i=1}^k \alpha_i l_i\| \to 0$ and $\sum_{i=1}^k \alpha_i l_i \in (\ker A)^*$, since $(\ker A)^*$ is a closed subspace in X^*. Consequently $y^*_n A$ converges in X^* to $x^* = -\sum_{i=1}^k \alpha_i l_i \in (\ker A)^* = A^* Y^*$. Then, there is (a unique) $y^* \in Y^*$ such that $y^* A = x^*$ and, by the Banach open mapping theorem applied to A^*, $\|y^* - y^*_n\| \leq \delta \|(y^* - y^*_n) A\|$ with some constant δ. Since $\|(y^* - y^*_n) A\| \to \|y^* A - x^*\| = 0$, we obtain that $\|y^* - y^*_n\| \to 0$. Consequently, $y^* A + \sum_{i=1}^k \alpha_i l_i = 0$. In view of Lemma 2.1, this contradicts the assumption of the proposition because $\|y^*\| + \sum_{i=1}^k \alpha_i = 1$. $\square$

As a consequence we obtain the following extension.

Proposition 2.2 *Suppose that A and l_i, $i = 1, \dots, k$, be as in Proposition 2.1 and c be the constant in (9). Let the functionals $\tilde{l}_i \in X^*$, $i = 1, \dots, k$, and operator $\tilde{A} : X \to Y$ satisfy*

$$\|\tilde{l}_i - l_i\| < \varepsilon, \quad i = 1, \dots, k, \quad \|\tilde{A} - A\| < \varepsilon, \quad 0 < \varepsilon < \hat{c}.$$

Then, for the system $\tilde{l}_1, \dots, \tilde{l}_k$, $\tilde{A}$ inequality (9) holds with $\tilde{c} := c - \varepsilon$.

Proof. Let $\alpha \in \mathbb{R}^k_+$, $y^* \in Y^*$. Then

$$\begin{aligned} \|\sum_{i=1}^k \alpha_i \tilde{l}_i + y^* \tilde{A}\| &\geq \|\sum_{i=1}^k \alpha_i l_i + y^* A\| - \|\sum_{i=1}^k \alpha_i (\tilde{l}_i - l_i)\| - \|y^*(\tilde{A} - A)\| \\ &\geq c(\sum_{i=1}^k \alpha_i + \|y^*\|) - \varepsilon(\sum_{i=1}^k \alpha_i + \|y^*\|). \end{aligned}$$

□

The following lemma is a reformulation of the extension of Hoffman's one [13] to a Banach space, obtained in [14, Theorem 3].

Lemma 2.2 *Let $A : X \to Y$ be a linear continuous operator with closed range AX, and let $l_i \in X^*$, $i = 1, \dots, k$. Then there is a constant $C_H > 0$ such that, for any $\xi = (\xi_1, \dots, \xi_k) \in \mathbb{R}^k$, $\eta \in Y$ and $x_0 \in X$ satisfying*

$$l_i\, x_0 \leq \xi_i, \quad Ax_0 = \eta, \tag{10}$$

there is a solution x' of the system

$$l_i\, (x_0 + x') \leq 0, \quad A(x_0 + x') = 0, \tag{11}$$

such that

$$\|x'\| \leq C_H \left(\max\{\xi_1^+, \dots, \xi_k^+\} + \|\eta\| \right), \tag{12}$$

where $\xi_i^+ = \max\{\xi_i, 0\}$.

3 Strong subregularity of the KKT system in mathematical programming: two norms approach

In this section we return to consider the optimization problem (1) as stated in the introduction. The main purpose will be to obtain sufficient conditions for appropriate metric subregularity property of the optimality (KKT) map associate with this problem. As explained in the introduction, due to the targeted application in optimal control we involve two norms in the space X in which the decision variable lives. In the first subsection we begin with some notations and known facts, followed by the assumptions needed for the subregularity result for the optimality map, presented in the second subsection.

3.1 The mathematical programming problem in a Banach space

The inequality constraints in problem (1) will shortly be written as $f(x) \le 0$, the inequality meant component-wise. Let $\hat{x}$ be an admissible point. Denote by I the set of active indices,

$$I = \{i \in \{1, \dots, k\} : f_i(\hat{x}) = 0\}.$$

Assuming that f_i, $i = 0, \dots, k$, and g are (Fréchet) differentiable at $\hat{x}$, we formulate the *Mangasarian-Fromovitz condition (constraint qualification)* (MFCQ) in the following form:
(a) $g'(\hat{x})X = Y$, and (b) the functionals $f_i'(\hat{x})$, $i \in I$, are positively independent on $\ker g'(\hat{x})$.

Remark 3.1 According to Lemma 2.1, MFCQ is equivalent to the following condition: AX is closed, and the relations

$$\alpha_i \ge 0 \text{ for } i \in I, \quad y^* \in Y^*, \quad \sum_{i \in I} \alpha_i f_i'(\hat{x}) + y^* g'(\hat{x}) = 0,$$

imply that $\alpha_i = 0$ for all $i \in I$, and $y^* = 0$.

The following first-order necessary optimality condition for problem (1) is well known, see e.g. [16, Theorem 3, Chapter 1].

Theorem 3.1 *Let $\hat{x}$ be a local minimum in problem (1). Assume that f_i, $i = 0, \dots, k$, and g are continuously differentiable around $\hat{x}$ and that the Mangasarian-Fromovitz constraint qualification holds at $\hat{x}$. Then there exist multipliers $\hat{y}^* \in Y^*$ and $\hat{\alpha} \in \mathbb{R}^k$ such that the triplet $(\hat{x}, \hat{y}^*, \hat{\alpha})$ satisfies the system*

$$\begin{gathered} f_0'(x) + y^* g'(x) + \textstyle\sum_{i=1}^k \alpha_i f_i'(x) = 0, \\ g(x) = 0, \\ \alpha \ge 0, \quad \alpha_i f_i(x) = 0, \quad i = 1, \dots, k, \\ f(x) \le 0. \end{gathered}$$

The above relations are known as Karush-Kuhn-Tucker (KKT) system. It can be formulated in a more compact way as

$$L_x'(x, y^*, \alpha) = 0, \tag{13}$$

$$g(x) = 0, \tag{14}$$

$$f(x) - N_{R_+^k}(\alpha) \ni 0, \tag{15}$$

where $L(x, y^*, \alpha) = f_0(x) + y^* g(x) + \sum_{i=1}^k \alpha_i f_i(x)$ is the Lagrangian, and

$$N_{R_+^k}(\alpha) := \begin{cases} \{\lambda \in \mathbb{R}^k : \langle \lambda, \beta - \alpha \rangle \le 0 \text{ for all } \beta \in \mathbb{R}_+^k\} & \text{if } \alpha \in \mathbb{R}_+^k, \\ \emptyset & \text{if } \alpha \notin \mathbb{R}_+^k, \end{cases}$$

is the normal cone to $\mathbb{R}_+^k$ at $\alpha \in \mathbb{R}^k$. Obviously $N_{R_+^k}(\alpha) = \prod_{i=1}^k N_{R_+}(\alpha_i)$, and (15) incorporates the *complementary slackness* condition $\alpha_i f_i(x) = 0$. The triples $(\hat{x}, \hat{y}^*, \hat{\alpha})$ satisfying the KKT system (13)–(15) are called *KKT points*.

Now we fix a (reference) KKT point $(\hat{x}, \hat{y}^*, \hat{\alpha})$ (such exists if $\hat{x}$ is a solution of the optimization problem (1)) and split the set of active indices I for $\hat{x}$ into two parts:

$$I_0 = \{i \in I : \hat{\alpha}_i = 0\}, \qquad I_1 = \{i \in I : \hat{\alpha}_i > 0\}. \tag{16}$$

Note that $\hat{\alpha}_i = 0$ for all $i \notin I$. The following assumption, known as *strict Mangasarian-Fromovitz condition*, is introduced in [19] in the case of finite-dimensional spaces X and Y.

Assumption 3.1 *For the fixed KKT point $(\hat{x}, \hat{\alpha}, \hat{y}^*)$ the image $g'(\hat{x})X$ is closed and the only pair $(y^*, \alpha) \in Y^* \times \mathbb{R}^k$ that satisfies the relations*

$$y^* g'(\hat{x}) + \sum_{i \in I} \alpha_i f_i'(\hat{x}) = 0, \quad \alpha_i \geq 0 \ (i \in I_0)$$

is $y^* = 0$, $\alpha = 0$.

Remark 3.2 It is proved in [19] (in the finite-dimensional case) that under Assumption 3.1 $(\hat{x}, \hat{\alpha}, \hat{y}^*)$ is the unique KKT point with the fixed $\hat{x}$. The proof is straightforward also in the Banach space setting, but this fact will also follow from Theorem 3.3 below.

In order to perform second order analysis of the optimization problem (1) we introduce, consistently with the material in Subsection 2.1, a second norm in the space X, denoted by $\|\cdot\|'$, which is weaker than $\|\cdot\|$, meaning that $\|x\|' \leq \|x\|$ for all $x \in X$. The dual space of X with respect to this norm will be denoted by X'', thus $X'' = \{l \in X^* : l$ is continuous with respect to $\|\cdot\|'\}$. (In the optimal control application in Section 4, for example, we use $X = L^\infty$, $\|\cdot\| = \|\cdot\|_\infty$, $\|\cdot\|' = \|\cdot\|_2$.) The norm in X'', denoted by $\|\cdot\|''$, is defined as usual:

$$\|l\|'' := \sup_{\|x\|' \leq 1} |lx|, \qquad l \in X''.$$

Obviously, $\|l\|'' \geq \|l\|$ for every $l \in X''$. Below we use the notation $\theta(t)$ for any function $(0, \infty) \to \mathbb{R}$ that converges to zero whenever $t \to 0$.

We make the following "two-norm differentiability" assumptions for the functions f_i and g.

Assumption 3.2 There exists a neighborhood $\hat{O}$ of $\hat{x}$ (in the norm $\|\cdot\|$) such that the following conditions are fulfilled for $i = 0, \ldots, k$ and for all Δx such that $\hat{x} + \Delta x \in \hat{O}$:

(i) The operator g and the functions f_i are continuously Fréchet differentiable in $\hat{O}$ in the norm $\|\cdot\|$, the derivatives $f_i'(\hat{x})$ and $g'(\hat{x})$ are continuous functionals/operator w.r.t. $\|\cdot\|'$, and

$$\begin{aligned} g(\hat{x} + \Delta x) &= g(\hat{x}) + g'(\hat{x})\Delta x + r(\Delta x), \\ f_i(\hat{x} + \Delta x) &= f_i(\hat{x}) + f_i'(\hat{x})\Delta x + r_i(\Delta x), \end{aligned}$$

with

$$\|r(\Delta x)\|_Y \leq \theta(\|\Delta x\|)\|\Delta x\|', \qquad |r_i(\Delta x)| \leq \theta(\|\Delta x\|)\|\Delta x\|'.$$

(ii) There exist bilinear mappings $Q : X \times X \to Y$ and $Q_i : X \times X \to \mathbb{R}$, $i = 0, \ldots, k$, such that

$$\begin{aligned} g'(\hat{x} + \Delta x) &= g'(\hat{x}) + Q(\Delta x, \cdot) + \bar{r}(\Delta x), \\ f_i'(\hat{x} + \Delta x) &= f_i'(\hat{x}) + Q_i(\Delta x, \cdot) + \bar{r}_i(\Delta x), \end{aligned}$$

$$\|Q(x_1,x_2)\|_Y \le C\|x_1\|'\|x_2\|', \quad |Q_i(x_1,x_2)| \le C\|x_1\|'\|x_2\|' \ \ \forall x_1, x_2 \in X, \tag{17}$$

where C is a constant and $\bar{r}$ and $\bar{r}_i$ satisfy

$$\sup_{\|x\|'\le 1} \|\bar{r}(\Delta x)(x)\|_Y \le \theta(\|\Delta x\|)\|\Delta x\|', \qquad \|\bar{r}_i(\Delta x)\|'' \le \theta(\|\Delta x\|)\|\Delta x\|'.$$

It is an easy exercise to show that under Assumption 3.2 one can represent

$$\begin{aligned} g(\hat{x}+\Delta x) &= g(\hat{x}) + g'(\hat{x})\Delta x + \frac{1}{2}Q(\Delta x,\Delta x) + \hat{r}(\Delta x), \\ f_i(\hat{x}+\Delta x) &= f_i(\hat{x}) + f_i'(\hat{x})\Delta x + \frac{1}{2}Q_i(\Delta x,\Delta x) + \hat{r}_i(\Delta x), \quad i = 0,\dots,k, \end{aligned}$$

where

$$\|\hat{r}(\Delta x)\|_Y \le \theta(\|\Delta x\|)(\|\Delta x\|')^2, \qquad |\hat{r}_i(\Delta x)| \le \theta(\|\Delta x\|)(\|\Delta x\|')^2.$$

Define the quadratic functional $\Omega : X \to \mathbb{R}$ as

$$\Omega(x) := Q_0(x,x) + \sum_{i\in I_1} \hat{\alpha}_i Q_i(x,x) + \hat{y}^*\, Q(x,x). \tag{18}$$

The following lemma is also quite obvious.

Lemma 3.1 *The following representation holds:*

$$L(\hat{x}+\Delta x, \hat{\alpha}, \hat{y}^*) = f_0(\hat{x}) + \frac{1}{2}\Omega(\Delta x) + r_L(\Delta x), \tag{19}$$

where $|r_L(\Delta x)| \le \theta(\|\Delta x\|)\,(\|\Delta x\|')^2$.

Define the so-called *critical cone* for the KKT point $(\hat{x}, \hat{\alpha}, \hat{y}^*)$ as

$$K := \{\delta x \in X \ : \ g'(\hat{x})\delta x = 0, \ \ f_i'(\hat{x})\delta x \le 0 \text{ for } i \in I \cup \{0\}\}.$$

It is easy to verify that the above definition is equivalent to the following one:

$$K := \{\delta x \in X \ : \ g'(\hat{x})\delta x = 0, \ \ f_i'(\hat{x})\delta x \le 0 \text{ for } i \in I_0, \ \ f_i'(\hat{x})\delta x = 0 \text{ for } \ i \in I_1\}.$$

Assumption 3.3 There exists a constant $c_0 > 0$ such that

$$\Omega(\delta x) \ge c_0\,(\|\delta x\|')^2 \quad \forall\, \delta x \in K.$$

We formulate the following second-order sufficiency theorem, which is similar to that in the single-norm case, see e.g. [20, pp. 146-148]). The proof is given for completeness.

Theorem 3.2 *Let Assumption 3.2 be fulfilled for the triple* $\hat{s} := (\hat{x}, \hat{y}^*, \hat{\alpha}) \in X \times Y^* \times \mathbb{R}^k$, *let* $\hat{s}$ *be a KKT point, and let Assumptions 3.1 and 3.3 be fulfilled for this point. Then the following quadratic growth condition for the objective function* f_0 *holds at the point* $\hat{x}$*: there exist* $c > 0$ *and* $\varepsilon > 0$ *such that*

$$f_0(x) - f_0(\hat{x}) \ge c\,(\|x-\hat{x}\|')^2$$

for all admissible x *such that* $\|x - \hat{x}\| < \varepsilon$. *Hence,* $\hat{x}$ *is a strict local minimizer in the problem.*

Proof. Let $x = \hat{x} + \Delta x$ be an admissible point with $\|\Delta x\| \leq \varepsilon$, where $\varepsilon > 0$ will be fixed later as sufficiently small. Using Lemma 3.1 we have

$$\begin{aligned} \gamma(x) := f_0(x) - f_0(\hat{x}) &= L(x, \hat{y}^*, \hat{\alpha}) - \hat{y}^* g(x) - \sum_{i=1}^{k} \hat{\alpha}_i f_i(x) - f_0(\hat{x}) \\ &\geq L(x, \hat{y}^*, \hat{\alpha}) - f_0(\hat{x}) = \frac{1}{2}\Omega(\Delta x) + r_L(\Delta x). \end{aligned} \tag{20}$$

From Assumption 3.2(i) we obtain that

$$\begin{aligned} &g'(\hat{x})\Delta x = -r(\Delta x), \\ &f_0'(\hat{x})\Delta x = \gamma(x) - r_0(\Delta x), \qquad f_i'(\hat{x})\Delta x \leq -r_i(\Delta x), \quad i \in I. \end{aligned}$$

According to Lemma 2.2, there exists $\delta x \in X$ such that

$$g'(\hat{x})\delta x = 0, \qquad f_i'(\hat{x})\delta x \leq 0, \quad i \in I \cup \{0\}$$

(which means that $\delta x \in K$) and

$$\begin{aligned} \|\delta x - \Delta x\| &\leq C_H(\|r(\Delta x)\| + \max\{|\gamma(x)| + |r_0(\Delta x)|, |r_i(\Delta x)|, i \in I\}) \\ &\leq \theta(\|\Delta x\|)\, \|\Delta x\|' + |\gamma(x)|. \end{aligned}$$

From the inequality $\|a\|^2 \geq \|b\|^2 - 2\|b\|\|a - b\|$, we get

$$\begin{aligned} (\|\delta x\|')^2 &\geq (\|\Delta x\|')^2 - 2\|\Delta x\|'\|\delta x - \Delta x\| \\ &\geq (\|\Delta x\|')^2 - 2\theta(\|\Delta x\|)(\|\Delta x\|')^2 - 2|\gamma(x)|\|\Delta x\|'. \end{aligned}$$

Using Assumption 3.2(ii) one can easily estimate

$$|\Omega(\delta x) - \Omega(\Delta x)| \leq c_1 \|\delta x - \Delta x\|'\, \|\delta x + \Delta x\|'.$$

Then (20) leads to the inequality

$$\gamma(x) \geq \frac{1}{2}\Omega(\delta x) - c_2\Big(\theta(\|\Delta x\|)\, (\|\Delta x\|')^2 + \|\Delta x\|'|\gamma(x)| + (\gamma(x))^2\Big),$$

where c_1 and c_2 are constants. Finally, using Assumption 3.3 we obtain that

$$\Omega(\delta x) \geq c_0(\|\delta x\|')^2 \geq c_0(\|\Delta x\|')^2 - 2c_0\theta(\|\Delta x\|)(\|\Delta x\|')^2 - 2c_0|\gamma(x)|\|\Delta x\|'.$$

Consequently,

$$\gamma(x) \geq \frac{1}{2}c_0(\|\Delta x\|')^2 - c_3\Big(\theta(\|\Delta x\|)\, (\|\Delta x\|')^2 + \|\Delta x\|'|\gamma(x)| + (\gamma(x))^2\Big),$$

with some $c_3 > 0$, which implies the desired estimate, provided that $\varepsilon > 0$ is fixed sufficiently small. □

3.2 Strong Metric subregularity of the KKT system

In this subsection we investigate the subregularity property SMSr2 (see Definition 2.1) of the KKT mapping

$$\mathcal{F}(x, y^*, \alpha) := \begin{pmatrix} L'_x(x, y^*, \alpha) \\ g(x) \\ f(x) - N_{R^k_+}(\alpha) \end{pmatrix} \tag{21}$$

appearing in the KKT system (13)–(15) for problem (1). As in Subsection 3.1, we consider a reference KKT point $\hat{s} = (\hat{x}, \hat{y}^*, \hat{\alpha})$ (that is a point satisfying $0 \in \mathcal{F}(\hat{s})$) for which assumptions 3.1–3.3 are fulfilled.

Define the space $\mathcal{X} := X \times Y^* \times \mathbb{R}^k$ with the following two norms (see Subsection 3.1):

$$\|s\|_{\mathcal{X}} := \|x\| + \|y^*\| + |\alpha| \quad \text{ and } \quad \|s\|'_{\mathcal{X}} := \|x\|' + \|y^*\| + |\alpha|, \quad s = (x, y^*, \alpha) \in \mathcal{X}.$$

Also define $\mathcal{Z} := X'' \times Y \times \mathbb{R}^k$ with the following norm: for $z = (\zeta, \eta, \xi) \in \mathcal{Z}$

$$\|z\|_{\mathcal{Z}} = \|\zeta\|'' + \|\eta\| + |\xi|.$$

Observe that due to Assumption 3.2(i) we have $y^* g'(\hat{x}) \in X''$, and also $f'_i(\hat{x}) \in X''$. Thus $\mathcal{F} : \mathcal{X} \rightrightarrows \mathcal{Z}$.

Our intension is to apply Theorem 2.1, for which we formally "linearize" the single-valued part of the mapping $\mathcal{F}$, using Assumption 3.2, but also include (for convenience) two quadratic terms. For any $s = (x, y^*, \alpha) \in \mathcal{X}$ denote $\Delta x = x - \hat{x}$, $\Delta y^* = y^* - \hat{y}^*$, $\Delta\alpha = \alpha - \hat{\alpha}$, $\Delta s = s - \hat{s}$. To shorten the notation we set

$$A := g'(\hat{x}), \quad l_i := f'_i(\hat{x}), \quad i = 0, \ldots, k, \quad l := (l_1, \ldots, l_k)^\top, \quad f := (f_1, \ldots, f_k)^\top,$$

where the superscript $\top$ means transposition. Define the set-valued mapping

$$\mathcal{L}(s) := \begin{pmatrix} Q_0(\Delta x, \cdot) + \hat{y}^* Q(\Delta x, \cdot) + \Delta y^* A \\ + \sum_{i=1}^k \hat{\alpha}_i Q_i(\Delta x, \cdot) + \sum_{i=1}^k \Delta\alpha_i l_i + D(\Delta s) \\ A\,\Delta x \\ f(\hat{x}) + l\,\Delta x - N_{R^k_+}(\alpha) \end{pmatrix}, \tag{22}$$

where $D(\Delta s) = \Delta y^* Q(\Delta x, \cdot) + \sum_{i=1}^k \Delta\alpha_i Q_i(\Delta x, \cdot) \in X^*$ is a quadratic term. The definition of $\mathcal{L}$ is the result of the formal linearization using Assumption 3.2(ii) for the first component (where also $D(\Delta s)$ is added), and Assumption 3.2(i) for the second and the third component. Due to the same assumptions $\mathcal{L}$ maps $\mathcal{X}$ to (the subsets of) $\mathcal{Z}$.

Lemma 3.2 *Let assumptions 3.1 and 3.3, and the inequalities (17) be fulfilled for A, Q, l_i and Q_i. Then there exist numbers $\hat{a} > 0$, $\hat{b} > 0$ and $\hat{\kappa} \geq 0$ such that the mapping $\mathcal{L}$ has the property SMSr2 (Definition 2.1) at $\hat{s}$ for $\hat{z} = 0$ with neighborhoods $\mathcal{O}_{\hat{s}} = I\!\!B_X(\hat{x}; \hat{a}) \times Y^* \times \mathbb{R}^k$, $\mathcal{O}_0 = I\!\!B_{\mathcal{Z}}(0; \hat{b})$ and constant $\hat{\kappa}$.*

Proof. The numbers $\hat{a}$ and $\hat{b}$ will be adjusted later in the proof. Take arbitrarily $z = (\zeta, \eta, \xi) \in I\!\!B_{\mathcal{Z}}(0; \hat{b})$ and let $s = (x, y^*, \alpha)$, with $x \in I\!\!B(\hat{x}; \hat{a})$, be a solution of the inclusion

$$\mathcal{L}(s) \ni z. \tag{23}$$

Clearly, $\hat{\alpha}_i > 0$ if and only if $i \in I_1$. Let $\hat{a} > 0$ and $\hat{b} > 0$ be fixed so small that for every $i \notin I$ (where $f_i(\hat{x}) < 0$) we have that $f_i(\hat{x}) + l_i \Delta x < \xi_i$. Then from the complementary slackness we obtain that $\alpha_i = 0$ for $i \notin I$. Then $\Delta \alpha_i = 0$ for $i \notin I$. Moreover, for $i \in I$ we have $f_i(\hat{x}) = 0$. Then (23) implies the following relations:

$$
\begin{aligned}
& Q_0(\Delta x, \cdot) + \hat{y}^* Q(\Delta x, \cdot) + \Delta y^* A \\
& + \sum_{i \in I_1} \hat{\alpha}_i Q_i(\Delta x, \cdot) + \sum_{i \in I} \Delta \alpha_i l_i + D(\Delta s) = \zeta, \qquad (24)\\
& A \Delta x = \eta, \qquad (25)\\
& l_i \, \Delta x - \xi_i \in N_{R_+}(\alpha_i), \quad i \in I. \qquad (26)
\end{aligned}
$$

Now we shall estimate $|\Delta \alpha|$ and $\|\Delta y^*\|$ by $\|\Delta x\|$ and $\|\zeta\|$. For that, we present (24) in the form

$$
\begin{aligned}
& \Delta y^* A + \sum_{i \in I_1} \Delta \alpha_i l_i + \sum_{i \in I_0} \Delta \alpha_i l_i \\
& = \zeta - Q_0(\Delta x, \cdot) - \hat{y}^* Q(\Delta x, \cdot) - \sum_{i \in I_1} \hat{\alpha}_i Q_i(\Delta x, \cdot) - D(\Delta s). \qquad (27)
\end{aligned}
$$

Note that $\Delta \alpha_i \geq 0$, $i \in I_0$. Define the map $\bar{A} : X \to Y \times \mathbb{R}^{|I_1|}$ as $\bar{A}x = (Ax, \{l_j x\}_{j \in I_1})$, where $|I_1|$ is the number of elements of I_1. Due to Assumption 3.1, the functionals $\{l_i\}_{i \in I_1}$ are linearly independent on $\ker A$. This, together with the surjectivity of A implies that $\bar{A}$ is surjective. Again from Assumption 3.1, we have that $\{l_i\}_{i \in I_0}$ are positively independent on $\ker \bar{A}$. Applying Proposition 2.1, we obtain that

$$
\Big\| \Delta y^* A + \sum_{i \in I} \Delta \alpha_i l_i \Big\| \geq c \Big(\|\Delta y^*\| + \sum_{i \in I} |\Delta \alpha_i| \Big) =: c\Theta. \qquad (28)
$$

The left-hand side of this inequality can be estimated from (27). For that we use that due to (17)

$$
\|\hat{y}^* Q(\Delta x, \cdot)\|'' = \sup_{\|x\|'=1} |\hat{y}^* Q(\Delta x, x)| \leq \|\hat{y}^*\| \sup_{\|x\|'=1} \|Q(\Delta x, x)\|_Y \leq \|\hat{y}^*\| C \|\Delta x\|',
$$

and similarly, $\|Q_i(\Delta x, \cdot)\|'' \leq C \|\Delta x\|'$. For the quadratic term we have due to (17) that

$$
\begin{aligned}
& \|D(\Delta s)\| \leq \|D(\Delta s)\|'' \\
& \leq \|\Delta y^*\| \, \|Q(\Delta x, \cdot)\|'' + \sum_{i=1}^{k} |\Delta \alpha_i| \, \|Q_i(\Delta x, \cdot)\|'' \leq \Theta C \|\Delta x\|'. \qquad (29)
\end{aligned}
$$

Then (28) and (27) imply that for some constants c_1, c_2,

$$
c\Theta \leq \|\zeta\| + c_1 \|\Delta x\|' + c_2 \Theta \|\Delta x\|'.
$$

Assuming that $\hat{a}$ (hence also $\|\Delta x\|$) is sufficiently small, we deduce that there is a constant c_3 such that

$$
\|\Delta y^*\| + |\Delta \alpha| \leq c_3 (\|\Delta x\|' + \|\zeta\|) \leq c_3 (\|\Delta x\|' + \|\zeta\|''). \qquad (30)
$$

One consequence of this estimate is that $\alpha_i > 0$ for $i \in I_1$, provided that $\hat{a} > 0$ and $\hat{b}$ are chosen sufficiently small. Hence, using the complementary slackness we obtain that $l_i \Delta x - \xi_i = 0$

for $i \in I_1$. For $i \in I_0$ it holds that $\Delta\alpha_i = \alpha_i$. Again from the complementary slackness we have $l_i \Delta x - \xi_i = 0$ if $\alpha_i > 0$. Thus, in all cases $\Delta\alpha_i(l_i\Delta x - \xi_i) = 0$ for $i \in I$. Then we can estimate

$$\Big| \sum_{i\in I} \Delta\alpha_i l_i \Delta x \Big| \le \sum_{i\in I} |\Delta\alpha_i|\,|\xi_i| \le c_3(\|\Delta x\|' + \|\zeta\|'')\,|\xi|. \tag{31}$$

Now we "multiply" (24) by Δx, (25) by Δy^*, and having in mind the definition of Ω in (18), we obtain that

$$\begin{aligned} \Omega(\Delta x) + (\Delta y^* A)\Delta x + \sum_{i\in I} \Delta\alpha_i l_i \Delta x + D(\Delta s)\Delta x &= \zeta\,\Delta x, \\ \Delta y^*(A\Delta x) &= \Delta y^*\,\eta. \end{aligned}$$

Subtracting the second from the first gives

$$\Omega(\Delta x) + \sum_{i\in I} \Delta\alpha_i l_i \Delta x + D(\Delta s)\Delta x = \zeta\,\Delta x - \Delta y^*\,\eta. \tag{32}$$

Inclusion (26) implies that

$$\begin{aligned} l_i\,\Delta x &\le \xi_i \quad \text{for } i \in I_0, \\ l_i\,\Delta x &= \xi_i \quad \text{for } i \in I_1. \end{aligned}$$

We apply Lemma 2.2 to this system extended with (25), and obtain that there exist $\delta x \in X$ such that

$$\|\delta x - \Delta x\| \le C_H(|\xi| + \|\eta\|) \le C_H\|z\| \tag{33}$$

and

$$l_i\,\delta x \le 0 \ \text{ for } i \in I_0, \qquad l_i\,\delta x = 0 \ \text{ for } i \in I_1, \qquad A\,\delta x = 0.$$

This means that $\delta x \in K$. As in the proof of Theorem 3.2, we estimate

$$|\Omega(\Delta x) - \Omega(\delta x)| \le c_1(\|\Delta x + \delta x\|')(\|\Delta x - \delta x\|') \le c_\Omega\,(\|\Delta x\|' + \|\delta x\|')\,\|z\|$$

with $c_\Omega = c_1 C_H$. Then from (32) and (31) we obtain that

$$\begin{aligned} \Omega(\delta x) &\le |\Omega(\Delta x)| + |\Omega(\Delta x) - \Omega(\delta x)| \\ &\le \Big| \sum_{i\in I} \Delta\alpha_i l_i \Delta x + D(\Delta s)\Delta x - \zeta\,\Delta x + \Delta y^*\,\eta \Big| \\ &\quad + c_\Omega\,(\|\Delta x\|' + \|\delta x\|')\,\|z\| \\ &\le c_3(\|\Delta x\|' + \|z\|)\,\|z\| + |D(\Delta s)\Delta x| + \|\zeta\|''\,\|\Delta x\|' \\ &\quad + \|\Delta y^*\|\,\|\eta\| + c_\Omega\,(\|\Delta x\|' + \|\delta x\|')\,\|z\|. \end{aligned} \tag{34}$$

From (29) and (30) we estimate

$$|D(\Delta s)\Delta x| \le C(\|\Delta y^*\| + |\Delta\alpha|)\|\Delta x\|'\,\|\Delta x\| \le C c_3(\|\Delta x\|' + \|z\|)\|\Delta x\|'\,\|\Delta x\|.$$

Using Assumption 3.3, the last inequality, (30), and (33) (which implies $\|\Delta x\|' \le \|\delta x\|' + C_H\|z\|$) we obtain from (34) that there exists constants c_4 and c_5 such that

$$c_0(\|\delta x\|')^2 \le \Omega(\delta x) \le c_4\left(\|\delta x\|'\|z\| + \|z\|^2\right) + c_5(\|\delta x\|')^2\,\|\Delta x\|.$$

Assuming that $\hat{a} > 0$ and $\hat{b}$ are sufficiently small (such that $c_5\hat{a} < c_0$), for an appropriate constant $\bar{c}$ we obtain the inequality

$$(\|\delta x\|')^2 \leq \bar{c}\left(\|\delta x\|'\|z\| + \|z\|^2\right).$$

This implies

$$\|\delta x\|' \leq \hat{c}\|z\| \qquad \text{with } \hat{c} = \frac{\bar{c} + \sqrt{\bar{c}^2 + 4\bar{c}}}{2}.$$

Then from (33) we have

$$\|\Delta x\|' \leq \|\delta x\|' + \|\delta x - \Delta x\|' \leq \hat{c}\|z\| + \|\delta x - \Delta x\| \leq (\hat{c} + C_H)\|z\|.$$

This together with (30) completes the proof. □

The main result in this section follows.

Theorem 3.3 *Let $\hat{s} = (\hat{x}, \hat{y}^*, \hat{\alpha})$ be a KKT point for problem (1), and let assumptions 3.1–3.3 be fulfilled at this point. Then the KKT mapping $\mathcal{F}$ defined in (21) has the property SMSr2 at $\hat{s}$ for zero. More precisely, there exist constants $a > 0$, $b > 0$ and $\kappa \geq 0$ such that for every $z = (\zeta, \eta, \xi) \in I\!B_{\mathcal{Z}}(0; b)$ (i.e. such that $\|\zeta\|'' + \|\eta\| + |\xi| \leq b$) and for every solution $s = (x, y^*, \alpha)$ of the inclusion $z \in \mathcal{F}(s)$ with $\|x - \hat{x}\| \leq a$ it holds that*

$$\|s - \hat{s}\|' = \|x - \hat{x}\|' + \|y^* - \hat{y}^*\| + |\alpha - \hat{\alpha}| \leq \kappa\|z\| = \kappa(\|\zeta\|'' + \|\eta\| + |\xi|).$$

Proof. We shall apply Theorem 2.1 for the mappings $\mathcal{L}$ in (22) and the function φ, which is the difference between the single-valued parts of $\mathcal{F}$ and $\mathcal{L}$, so that $\mathcal{F} = \varphi + \mathcal{L}$. Namely, having in mind Lemma 3.2, we have to ensure existence of numbers $a > 0$, $b > 0$, $\mu \geq 0$ and κ such conditions (7) are satisfied with the specific form of the neighborhood $\mathcal{O}_{\hat{s}}$ in Lemma 3.2. It is enough to fix $\mu = 1/2\hat{\kappa}$, $\kappa = 2\hat{\kappa}$, $a = \hat{a}$, $b = \hat{b}/2$ and to ensure that the following inequalities are fulfilled:

$$\|\varphi(s) - \varphi(\hat{s})\| \leq \mu\|s - \hat{s}\|' \leq b \qquad \forall\, s = (x, y^*, \alpha) \text{ with } x \in I\!B(\hat{x}; \hat{a}).$$

Clearly, the numbers $\hat{a}$ and $\hat{b}$ in Lemma 3.2 can be assumed as small as necessary, because the property SMSr2 remains true when decreasing the neighborhoods in its definition. In particular, $\hat{a}$ has to be so small that $I\!B_X(\hat{x}; \hat{a}) \subset \hat{\mathcal{O}}$ in order to utilize Assumption 3.2.

For the second component of φ we have the following estimate using Assumption 3.2(i):

$$\begin{aligned} \|g(x) - A\Delta x\| &= \|g(x) - (g(\hat{x}) + g'(\hat{x})\Delta x)\| \\ &\leq \|r(\Delta x)\| \leq \theta(\|\Delta x\|)\|\Delta x\|' \leq (\mu/3)\|\Delta x\|', \end{aligned}$$

provided that $\hat{a} > 0$ is sufficiently small. Similarly we treat the third component of φ (note that this third component vanishes for all $i \notin I$). The first one, denoted further by $\varphi^L \in X''$, requires more attention. After substitution of the expressions for $f_i'(\hat{x} + \Delta x)$ and $g'(\hat{x} + \Delta x)$ from Assumption 3.2(ii) in $L_x'(x, y^*, \alpha)$, the somewhat long expression for φ^L reduces to

$$\varphi^L(s) = \bar{r}_0(\Delta x) + y^*\bar{r}(\Delta x) + \sum_{i=1}^{k} \alpha_i \bar{r}_i(\Delta x).$$

In order to estimate φ^L it is enough to prove that y^* and α are uniformly bounded when $\|\Delta x\| \leq \hat{a}$ and $\|\zeta\| \leq b$. Since the non-active constraints for $\hat{x}$ remain non-active for x when $\hat{a}$ is sufficiently small, we have that $\alpha_i = 0$ for $i \notin I$. Then inclusion $z \in \mathcal{F}(s)$ implies

$$f_0'(x) + y^* g'(x) + \sum_{i \in I} \alpha_i f_i'(x) = \zeta.$$

Since the vectors $l_i = f_i'(\hat{x})$ are positively independent on $\ker A$ (see Assumption 3.1), and f_i' and g' are continuous around $\hat{x}$, Proposition 2.2 gives that there is a constant $\bar{C}$ such that

$$\|y^*\| + |\alpha| \leq \bar{C} \quad \text{whenever } x \in I\!B_X(\hat{x}; \hat{a}), \quad \|\zeta\| \leq b, \quad \hat{a},\, b - \text{sufficiently small.}$$

Now all terms in the expression for φ^L can be directly estimated from Assumption 3.2(ii). For example,

$$\|\hat{y}^* \bar{r}(\Delta x)\|' = \sup_{\|x\|' \leq 1} |\hat{y}^* \bar{r}(\Delta x) x| \leq \|\hat{y}^*\| \sup_{\|x\|' \leq 1} |\bar{r}(\Delta x) x| \leq C\theta(\|\Delta x\|)\|\Delta x\|',$$

and $C\theta(\|\Delta x\|)$ can be made sufficiently small by choosing $\hat{a}$ small enough. This completes the proof of the theorem. □

4 Application of Abstract Result: SMsR in the Mayer Problem

In this section we apply the obtained abstract subregularity result (Theorem 3.3) to the Mayer-type optimal control problem (3)–(6). An important feature is that general initial/terminal constraints for the state are involved.

4.1 Subregularity of the Lagrange (KKT) optimality mapping

First, we will directly translate the results in the previous section to the Mayer problem (3)–(6). For now, we assume that the functions h, ψ_j and φ_i, appearing in (3)–(6) are continuously differentiable. We consider this problem for trajectory-control pairs $w(\cdot) = (x(\cdot), u(\cdot))$ with measurable and essentially bounded $u : [t_0, t_1] \to \mathbb{R}^m$ and absolutely continuous $x : [t_0, t_1] \to \mathbb{R}^n$. Thus the admissible points in the problem belong to the space

$$\mathcal{W} := W^{1,1}([t_0, t_1]; \mathbb{R}^n) \times L^\infty([t_0, t_1]; \mathbb{R}^m) =: W^{1,1} \times L^\infty$$

with the norm

$$\|w\| = \|x\|_{1,1} + \|u\|_\infty = |x(t_0)| + \int_{t_0}^{t_1} |\dot{x}(t)| \,\mathrm{d}t + \operatorname*{ess\,sup}_{t \in [t_0.t_1]} |u(t)|.$$

The mapping g and the functions f_i in the abstract problem (1) will be specified as follows. With the notations

$$q := (x(t_0), x(t_1)), \;\; G(x,u) := h(x,u) - \dot{x}, \;\; \psi = (\psi_1, \ldots, \psi_s), \;\; \varphi = (\varphi_1, \ldots, \varphi_k),$$

we set

$$\mathcal{W} \ni (x,u) \mapsto g(x,u) := \big(G(x,u), \psi(q)\big) \in L^1 \times \mathbb{R}^s =: Y, \tag{35}$$
$$\mathcal{W} \ni (x,u) \mapsto f(x,u) := \varphi(q) \in \mathbb{R}^k.$$

Further on we use the similar abbreviations

$$\hat{q} := (\hat{x}(t_0), \hat{x}(t_1)), \quad \delta q := (\delta x(t_0), \delta x(t_1)), \ldots \tag{36}$$

Due to the continuous differentiability of φ_i, the functions f_i, $i = 0, 1, \ldots, k$, are also continuously differentiable and the Fréchet derivative of f_i at a point $w = (x,u)$ is the linear functional

$$\mathcal{W} \ni (\delta x, \delta u) \mapsto f_i'(w)(\delta x, \delta u) = \varphi_i'(q)\delta q \in \mathbb{R}, \quad i = 0, 1, \ldots, k. \tag{37}$$

Notice that when the functionals $f_i'(w)$ and $\varphi_i'(q)$ are considered as vectors, they are treated as vector-rows, while the elements of the "primal" space, δx, δu, etc. are vector-columns. The same applies to the other linear functionals that will appear later. Also note that in order to make (37) consistent with the matrix calculus, $(\delta x, \delta u)$ should be understood as a vector column of dimension $n + m$.

Since the norm for the u-components of the elements of $\mathcal{W}$ is L^∞, the mapping g is also continuously Fréchet differentiable and its derivative $g'(w) : \mathcal{W} \to L^1 \times \mathbb{R}^s$ at $w \in \mathcal{W}$ is the continuous linear operator defined as

$$\begin{aligned} g'(w)\delta w \;&=\; \big(h'(w)\delta w - \delta\dot{x},\, \psi'(q)\delta q\big) \\ &=: \big(G'(w)\delta w, \psi'(q)\delta q\big), \quad \forall\, \delta w = (\delta x, \delta u) \in \mathcal{W}, \end{aligned} \tag{38}$$

where $h'(w)\delta w = h_x(x,u)\delta x + h_u(x,u)\delta u$ and the derivative $G'(w)$ of G at w is defined by $G'(w)\delta w = h'(w)\delta w - \delta\dot{x}$. Since the operator $G'(w) : \mathcal{W} \to L^1$ is surgective, and the operator $w \in \mathcal{W} \to \psi'(q)\delta q \in \mathbb{R}^s$ is finite dimensional, the operator $g'(w) : \mathcal{W} \to L^1 \times \mathbb{R}^s$ has a closed image, see e.g. [5, Corollary 3.3].

The dual space to Y is $Y^* = L^\infty \times \mathbb{R}^s$ with elements $y^* = (p, \beta)$. Then the Lagrange function associated with problem (3)-(6) takes the form

$$L(w, y^*, \alpha) = L(x, u, p, \beta, \alpha) = \varphi_0(q) + p(-\dot{x} + h(x,u)) + \beta\psi(q) + \alpha\varphi(q),$$

where $pv = \int_{t_0}^{t_1} p(t)v(t)\,\mathrm{d}t$ for $v \in L^1$ (see also notational convention (36)). Its derivative $L_w(w, p, \beta, \alpha) \in \mathcal{W}^* = (W^{1,1})^* \times (L^\infty)^*$ acts on $\delta w \in \mathcal{W}$ as

$$\begin{aligned} &L_w(w, p, \beta, \alpha)(\delta w) \\ &= \varphi_0'(q)\delta q + \int_{t_0}^{t_1} p(t)\left(h'(w(t))\delta w(t) - \delta\dot{x}(t)\right)\mathrm{d}t + \beta\psi'(q)\delta q + \alpha\varphi'(q)\delta q. \end{aligned} \tag{39}$$

We remind that $h'(w)\delta w - \delta\dot{x}$ maps $\mathcal{W}$ to L^1 and

$$pG'(w)\delta w = p(h'(w)\delta w - \delta\dot{x}) = \int_{t_0}^{t_1} p(t)\big(h'(w(t))\delta w(t) - \delta\dot{x}(t)\big)\,\mathrm{d}t. \tag{40}$$

Let $\hat{w} = (\hat{x}, \hat{u})$ be an admissible pair. As in Section 3 we denote

$$I := \{i \in \{1, \ldots, k\} \;:\; \varphi_i(\hat{q}) = 0\}.$$

Taking into account that the operator $g'(\hat{w}) : \mathcal{W} \to L^1 \times \mathbb{R}^s$ has a closed image, we translate to the Mayer problem the Mangasarian-Fromovitz condition at $\hat{w}$ in the form as in Remark 3.1: the relations

$$\alpha_i \geq 0 \ \text{ for } i \in I, \quad (p, \beta) \in L^\infty \times \mathbb{R}^s, \quad pG'(\hat{w}) + \beta\psi'(\hat{q}) + \sum_{i \in I} \alpha_i \varphi_i'(\hat{q}) = 0 \tag{41}$$

imply that $\alpha_i = 0$ for $i \in I$, $\beta = 0$, $p = 0$. We remind that local minimum in the norm $\mathcal{W}$ is called *weak local minimum.* Then Theorem 3.1 implies the following well-known first-order optimality condition (see e.g. [24, page 24]).

Theorem 4.1 *Let $\hat{w} = (\hat{x}, \hat{u})$ be a weak local minimum in problem (3)-(6). Assume that all the functions h, ψ_j, φ_i are continuously differentiable and that the Mangasarian-Fromovitz condition holds at $\hat{w}$. Then there exist multipliers $\hat{p} \in L^\infty$, $\hat{\beta} \in \mathbb{R}^s$, $\hat{\alpha} \in \mathbb{R}^k$ such that the tuple $(\hat{w}, \hat{p}, \hat{\beta}, \hat{\alpha})$ satisfies the system:*

$$\begin{gathered} L_w(w, p, \beta, \alpha) = 0, \\ -\dot{x} + h(w) = 0, \\ \psi(q) = 0, \\ \alpha \geq 0, \quad \varphi(q) \leq 0, \quad \alpha\varphi(q) = 0, \end{gathered}$$

where $q := (x(t_0), x(t_1))$.

We remind that the relations in the last exposed line are equivalent to $\varphi(q) \in N_{\mathbb{R}^k_+}(\alpha)$. Then the KKT optimality mapping for the Mayer problem in consideration is

$$\mathcal{F}(x, u, p, \beta, \alpha) := \begin{pmatrix} L_w(x, u, p, \beta, \alpha) \\ -\dot{x} + h(x, u) \\ \psi(q) \\ \varphi(q) - N_{\mathbb{R}^k_+}(\alpha) \end{pmatrix}, \tag{42}$$

and the optimality system of Theorem 4.1 is equivalent to the condition $0 \in \mathcal{F}(x, u, p, \beta, \alpha)$. Introduce the space $\mathcal{X}$ with elements $s = (x, u, p, \beta, \alpha) \in W^{1,1} \times L^\infty \times L^\infty \times \mathbb{R}^s \times \mathbb{R}^k$, in which the mapping $\mathcal{F}$ is defined. Our goal in this subsection is to obtain sufficient conditions for the subregularity property SMSr2 (Definition 2.1) of the mapping $\mathcal{F}$ by applying Theorem 3.3. For that we introduce a second norm in the space $\mathcal{W}$ as

$$\|w\|' := \|x\|_{1,1} + \|u\|_2.$$

Clearly, we have

$$\|w\|' \leq c\|w\| \quad \forall\, w \in \mathcal{W}$$

with an appropriate constant c. Hence, the norm $\|\cdot\|'$ is weaker than the norm $\|\cdot\|$. Correspondingly, the dual space of $\mathcal{W}$ with respect to this norm, denoted by $\mathcal{W}''$ as in the previous section, is $(W^{1,1})^* \times L^2$, with the norm

$$\|(\pi, \rho)\|'' := \sup_{\|x\|_{1,1} \leq 1} \pi x + \|\rho\|_2.$$

The strict MFCQ (Assumption 3.1) takes the following form. For a given admissible reference point $\hat{w} = (\hat{x}, \hat{u})$ and multipliers $\hat{p}$, $\hat{\beta}$, $\hat{\alpha}$ define

$$I_0 = \{i \in I \ : \ \hat{\alpha}_i = 0\}, \qquad I_1 = \{i \in I \ : \ \hat{\alpha}_i > 0\}.$$

Assumption 4.1 The relations

$$\alpha_i \geq 0 \ \text{ for } i \in I_0, \quad (p, \beta) \in L^\infty \times \mathbb{R}^s, \quad pG'(\hat{w}) + \beta\psi'(\hat{q}) + \sum_{i \in I} \alpha_i \varphi_i'(\hat{q}) = 0$$

imply that $\alpha_i = 0$ for $i \in I$, $\beta = 0$, $p = 0$.

Assumption 4.2 The functions h, ψ_j, $j = 1, \ldots, s$ and φ_i, $i = 0, \ldots, k$ are twice continuously differentiable.

Let us check that Assumption 3.2 is fulfilled. Both points (i) and (ii) are obviously fulfilled for $f_i(x, u) = \varphi_i(x(t_0), x(t_1))$ and for the second component $(x, u) \mapsto \psi(x(t_0), x(t_1))$ of g in (35), because of Assumption 4.2 and the inequality

$$|x(t_i) - \tilde{x}(t_i)| \leq \|x - \tilde{x}\|_{1,1} \leq \|w - \tilde{w}\|' \quad \forall\, w = (x, u), \tilde{w} = (\tilde{x}, \tilde{u}) \in \mathcal{W}, \quad i = 0, 1.$$

Let us check point (i) for the mapping G:

$$G(\hat{w} + \Delta w) - G(w) = h(\hat{w} + \Delta w) - h(\hat{w}) - \Delta\dot{x} = G'(\hat{w})\Delta w + r(\Delta w),$$

where the above relations should be understood point-wise (for a.e. t) and

$$\|r(\Delta w)\|_1 \leq \int_{t_0}^{t_1} \theta(|\Delta w(t)|)\, |\Delta w(t)| \, \mathrm{d}t \leq \theta(\|\Delta w\|_\infty)\|\Delta w\|_2 \leq \theta(\|\Delta w\|)\|\Delta w\|'$$

(notice that, according to the convention before Assumption 3.2, θ may change from place to place).

Let us check point (ii) for the mapping $G'(w)$ with $Q(w_1, w_2) = h''(\hat{w})(w_1, w_2)$:

$$\begin{aligned}(G'(\hat{w} + \Delta w) - G'(\hat{w}))(w) &= h'(\hat{w} + \Delta w)(w) - h'(\hat{w})(w) \\ &= h''(\hat{w})(\Delta w, w) + \bar{r}(\Delta w)(w),\end{aligned}$$

where as above $\|\bar{r}(\Delta w)(w)\|_1 \leq \theta(\|\Delta w\|)\|w\|'$ and

$$\|Q(w_1, w_2)\| \leq c\|w_1\|_2 \, \|w_2\|_2 \leq c_1\|w_1\|' \, \|w_2\|'.$$

Thus Assumption 3.2 is fulfilled.

Next, we define the quadratic functional Ω on $\mathcal{W}$. For this it is convenient to introduce the *Humiltonian (Pontryagin function)* and the *endpoint Lagrange function*:

$$H(x, u, p) = p\, h(x, u), \quad l(q, \alpha, \beta) = \varphi_0(q) + \sum_{i \in I_1} \alpha_i \varphi_i(q) + \beta\psi(q). \tag{43}$$

Then, having in mind the expressions for the respective derivatives, the expression in (18) becomes

$$\Omega(\delta w) = \langle l_{qq}(\hat{q}, \hat{\alpha}, \hat{\beta})\delta q, \delta q\rangle + \int_{t_0}^{t_1} \langle H_{ww}(\hat{x}, \hat{u}, \hat{p})(t)\delta w(t), \delta w(t)\rangle \, \mathrm{d}t,$$

where

$$\begin{aligned}&\langle H_{ww}(\hat{x}, \hat{u}, \hat{p})\, \delta w, \delta w\rangle \\ &= \langle H_{xx}(\hat{x}, \hat{u}, \hat{p})\delta x, \delta x\rangle + 2\langle H_{ux}(\hat{x}, \hat{u}, \hat{p})\delta x, \delta u\rangle + \langle H_{uu}(\hat{x}, \hat{u}, \hat{p})\delta u, \delta u\rangle.\end{aligned}$$

The critical cone K at the point $(\hat{x}, \hat{u})$ takes the form

$$K = \{(\delta x, \delta u) \in W^{1,1} \times L^\infty \;\;:\;\; \begin{array}{l} \delta \dot{x} = h_x(\hat{x}, \hat{u})\delta x + h_u(\hat{x}, \hat{u})\delta u, \\ \psi'(\hat{q})\delta q = 0,\; \varphi_i'(\hat{q})\delta q \le 0,\; i \in I \cup \{0\}\}, \end{array}$$

where $\delta q = (\delta x(t_0), \delta x(t_1))$. As in the general case (Subsection 3.1), the critical cone can be equivalently defined as the set of those pairs $(\delta x, \delta u) \in W^{1,1} \times L^\infty$ which satisfy

$$\begin{array}{l} \delta \dot{x} = h_x(\hat{x}, \hat{u})\delta x + h_u(\hat{x}, \hat{u})\delta u, \quad \psi'(\hat{q})\delta q = 0, \\ \varphi_i'(\hat{q})\delta q \le 0,\; i \in I_0, \quad \varphi_i'(\hat{q})\delta q = 0,\; i \in I_1. \end{array}$$

The following condition repeats Assumption 3.3 (the coercivity condition) with the meaning of the notations from the present subsection:

Assumption 4.3 (*Coercivity*) There exists a constant $c_0 > 0$ such that

$$\Omega(\delta w) \ge c_0\, (\|\delta w\|')^2 \quad \forall\, \delta w \in K.$$

Remark 4.1 It is well-known and easy to prove that the coercivity condition is equivalent to the following simpler one: there is a constant $c_0' > 0$ such that

$$\Omega(\delta w) \ge c_0'\, (|\delta x(t_0)|^2 + \|\delta u\|_2^2) \quad \forall\, \delta w \in K.$$

Then Theorem 3.3 directly implies the property SMSr2 of the mapping $\mathcal{F}$ in (42). The elements of the image space, $\mathcal{Z}$ (which are considered as disturbances of $\mathcal{F}$) will be denoted by

$$z = (\zeta, \eta, \mu, \xi) \in \mathcal{Z} := \mathcal{W}'' \times L^1 \times \mathbb{R}^s \times \mathbb{R}^k = (W^{1,1})^* \times L^2) \times L^1 \times \mathbb{R}^s \times \mathbb{R}^k,$$

so that the inclusion $z \in \mathcal{F}(s)$ reads as

$$\begin{pmatrix} \zeta \\ \eta \\ \mu \\ \xi \end{pmatrix} \in \begin{pmatrix} L_w(x, u, p, \beta, \alpha) \\ -\dot{x} + h(x, u) \\ \psi(q) \\ \varphi(q) - N_{\mathbb{R}^k_+}(\alpha) \end{pmatrix}.$$

Theorem 4.2 *Let $\hat{s} = (\hat{x}, \hat{u}, \hat{p}, \hat{\beta}, \hat{\alpha}) \in \mathcal{X}$ be a solution of the inclusion $0 \in \mathcal{F}(s)$, and let assumptions 4.1–4.3 be fulfilled at this point. Then there exist constants $a > 0$, $b > 0$ and $\kappa \ge 0$ such that for every $z = (\zeta, \eta, \mu, \xi) \in \mathcal{Z}$ with $\|z\| := \|\zeta\|'' + \|\eta\|_1 + |\mu| + |\xi| \le b$ and for every solution $s = (x, u, p, \beta, \alpha) \in \mathcal{X}$ of the inclusion $z \in \mathcal{F}(s)$ with $\|x - \hat{x}\|_{1,1} + \|u - \hat{u}\|_\infty \le a$ it holds that*

$$\|x - \hat{x}\|_{1,1} + \|u - \hat{u}\|_2 + \|p - \hat{p}\|_\infty + |\beta - \hat{\beta}| + |\alpha - \hat{\alpha}| \le \kappa \|z\|.$$

4.2 Subregularity of the Pontryagin optimality mapping

It is reasonable to restrict the considerations of disturbances $\zeta \in (W^{1,1})^* \times L^2$ to such having representatives by triplets $(\pi, \nu, \rho) \in L^1 \times \mathbb{R}^{2n} \times L^2$:

$$\zeta\, \delta w = \nu \delta q + \int_{t_0}^{t_1} [\pi(t)\delta x(t) + \rho(t)\delta u(t)]\, \mathrm{d}t, \tag{44}$$

where $\delta q = (\delta x(t_0), \delta x(t_1))$ (we remind the convention (36)) and $\nu = (\nu_0, \nu_1)$. Thus we consider the space of disturbances

$$\hat{\mathcal{Z}} = \big\{(\pi, \nu, \rho, \eta, \mu, \xi) \in L^1 \times \mathbb{R}^{2n} \times L^2 \times L^1 \times \mathbb{R}^s \times \mathbb{R}^k\big\}.$$

Let Assumption 4.2 be fulfilled. The following lemma, extending [25, Lemma 2], plays the key role in this subsection.

Lemma 4.1 *The equation $L_w(x,u,p,\beta,\alpha) = \zeta$, with $(x,u) = w \in \mathcal{W}$, $p \in L^\infty$, $\alpha \in \mathbb{R}^k$, $\beta \in \mathbb{R}^s$, and ζ having the representation (44) with $(\pi, \nu, \rho) \in L^\infty \times \mathbb{R}^{2n} \times L^2$, is equivalent to the system*

$$0 = \dot{p} + p\, h_x(w) - \pi, \tag{45}$$
$$0 = p h_u(w) - \rho, \tag{46}$$
$$0 = (p(t_0), -p(t_1)) + \varphi_0'(q) + \alpha\varphi'(q) + \beta\psi'(q) - \nu, \tag{47}$$

with $p \in W^{1,1}$ and $\rho \in L^\infty$.

Proof. 1. Let $L_w(x,u,p,\beta,\alpha) = \zeta$ as in the formulation of the lemma. Then from (39),

$$l_q(q,\alpha,\beta)\delta q + \int_{t_0}^{t_1} p\left(h'(w)\delta w - \delta\dot{x}\right) \mathrm{d}t = \nu\delta q + \int_{t_0}^{t_1} \left(\pi\delta x + \rho\delta u\right) \mathrm{d}t \quad \forall\, \delta w \in \mathcal{W},$$

where l as in (43). Setting $\delta x = 0$, due to the arbitrariness of $\delta u \in L^\infty$, we obtain (46), and hence $\rho \in L^\infty$.

Let for an arbitrary function $v \in L^1$ the function δx be any solution of the equation

$$\delta\dot{x}(t) = h_x(w(t))\delta x(t) - v(t).$$

Then, also using (46), we have

$$l_q(q,\alpha,\beta)\delta q + \int_{t_0}^{t_1} p\, v\, \mathrm{d}t = \nu\delta q + \int_{t_0}^{t_1} \pi\delta x\, \mathrm{d}t. \tag{48}$$

Let $\bar{p} \in W^{1,1}$ be any solution of the equation

$$-\dot{\bar{p}}(t) = \bar{p}(t)h_x(w(t)) - \pi(t).$$

Integration by parts gives

$$\bar{p}(t_1)\delta x(t_1) - \bar{p}(t_0)\delta x(t_0) = \int_{t_0}^{t_1} \frac{\mathrm{d}}{\mathrm{d}t}(\bar{p}(t)\delta x(t))\, \mathrm{d}t$$
$$= \int_{t_0}^{t_1} (-\bar{p}(t)h_x(w(t)) + \pi(t))\delta x(t)\, \mathrm{d}t + \int_{t_0}^{t_1} \bar{p}(t)(h_x(w(t))\delta x(t) - v(t))\, \mathrm{d}t$$
$$= \int_{t_0}^{t_1} (\pi(t)\delta x(t) - \bar{p}(t)v(t))\, \mathrm{d}t.$$

If $\bar{p}$ and δx are chosen so that

$$\bar{p}(t_1)\delta x(t_1) - \bar{p}(t_0)\delta x(t_0) = l_q(q,\alpha,\beta)\delta q - \nu\delta q \tag{49}$$

(say, $\delta x(t_0)=0$ and $p(t_1)=l_{q_1}(q,\alpha,\beta)-\nu_1$), then

$$l_q(q,\alpha,\beta)\delta q+\int_{t_0}^{t_1}\bar p v\,\mathrm{d}t=\int_{t_0}^{t_1}\pi\delta x\,\mathrm{d}t+\nu\delta q.$$

Subtracting this equality from (48), we obtain that $\int_{t_0}^{t_1}(p-\bar p)v\,\mathrm{d}t=0$. Since $v\in L^1$ is arbitrarily chosen, we obtain that $p=\bar p$. Thus $p\in W^{1,1}$ and satisfies (45).

Let us represent $\nu=(\nu_0,\nu_1)\in\mathbb{R}^n\times\mathbb{R}^n$, $l_q(q,\alpha,\beta)=(l_0,l_1)\in\mathbb{R}^n\times\mathbb{R}^n$, and $(\delta x(t_0),\delta x(t_1))=(q_0,q_1)$. We still have the freedom to choose the initial or the final condition for $\bar p$ and δx so that (49) is satisfied. One way is to choose $\delta x(t_0)=0$ and $\bar p(t_1)=l_1-\nu_1$ (as already said), another way is to choose $\delta x(t_1)=0$ and $\bar p(t_0)=-l_0+\nu_0$. In both cases, the corresponding function $\bar p$ is the same, equal to p. Thus (47) is also satisfied by p.

2. Now let us prove the converse claim. Summing (45) and (46), the first multiplied by $\delta x\in W^{1,1}$ and the second multiplied by $\delta u\in L^\infty$, and then integrating over the segment $[t_0,t_1]$, results in the equality (where $\delta w=(\delta x,\delta u)\in\mathcal{W}$)

$$\int_{t_0}^{t_1}\dot p\delta x\,\mathrm{d}t+\int_{t_0}^{t_1}ph_w(w)\delta w\,\mathrm{d}t-\int_{t_0}^{t_1}(\pi\delta x+\rho\delta u)\,\mathrm{d}t=0.$$

Integrating by parts the first term and using (47), we obtain that

$$\Big(l_q(q,\alpha,\beta)-\nu\Big)\delta q+\int_{t_0}^{t_1}\Big(-p\delta\dot x+p\,h_w(w)\delta w-\pi\delta x-\rho\delta u\Big)\,\mathrm{d}t=0.$$

Having in mind (39),(44) and the arbitrariness of $\delta w\in\mathcal{W}$, this gives
$L_w(x,u,p,\beta,\alpha)=\zeta$. □

The following assumption strengthens Assumption 4.1.

Assumption 4.4 The relations

$$\alpha_i\ge 0\ \text{ for } i\in I_0,\quad (p,\beta)\in W^{1,1}\times\mathbb{R}^s, \tag{50}$$

$$\dot p=-p\,h_x(\hat w),\ ph_u(\hat w)=0,\ (-p(t_0),p(t_1))=\sum_{i\in I}\alpha_i\varphi_i'(\hat q)+\beta\psi'(\hat q) \tag{51}$$

imply that $\alpha_i=0$ for $i\in I$, $\beta=0$, $p=0$.

Let us show that Assumption 4.4 implies Assumption 4.1 for optimal control problem (3)–(6). By Lemma 4.1, system (50)–(51) is equivalent to the equation $L_w(\hat x,\hat u,p,\beta,\alpha)=\zeta$, where ζ is represented by the triple (π,ν,ρ) in the form (44) with $\pi=0$, $\rho=0$ and $\nu=\varphi_0'(\hat q)$. Using also (39) in this equation, we get

$$\int_{t_0}^{t_1}p\,(h'(\hat w)\delta w-\delta\dot x)\,\mathrm{d}t+\beta\psi'(\hat q)\delta q+\alpha\varphi'(\hat q)\delta q=0.$$

According to (40), the latter is equivalent to

$$pG'(\hat w)\delta w+\beta\psi'(\hat q)\delta q+\alpha\varphi'(\hat q)\delta q=0\quad\forall\,\delta w\in\mathcal{W}.$$

In view of (38), this means that $y^* g'(\hat{w}) + \alpha \varphi'(\hat{q}) = 0$, where $y^* = (p, \beta)$. Therefore, indeed Assumption 4.4 implies Assumption 4.1.

In particular, Assumption 4.1 implies the Mangasarian-Fromovitz condition (41), hence also the claim of Theorem 4.1 is valid. Applying Lemma 4.1 with $\zeta = 0$ we obtain that any weak solution $\hat{w} = (\hat{x}, \hat{u})$ of problem (3)–(6), together with the corresponding Lagrange multipliers $(\hat{p}, \hat{\beta}, \hat{\alpha}) \in W^{1.1} \times \mathbb{R}^s \times \mathbb{R}^k$ satisfies (for a.e. $t \in [t_0, t_1]$) the following (local) Pontryagin conditions:

$$\hat{\mathcal{F}}(x,u,p,\beta,\alpha)(t) := \begin{pmatrix} \dot{p}(t) + p(t)\, h_x(w(t)) \\ (p(t_0), -p(t_1)) + \varphi_0'(q) + \sum_{i\in I} \alpha_i \varphi_i'(q) + \beta \psi'(q) \\ p(t) h_u(w(t)) \\ -\dot{x}(t) + h(x(t), u(t)) \\ \psi(q) \\ \varphi(q) - N_{\mathbb{R}^k_+}(\alpha) \end{pmatrix} \ni 0. \tag{52}$$

Introduce the space $\hat{\mathcal{X}}$ with elements $s = (x, u, p, \beta, \alpha) \in W^{1,1} \times L^\infty \times W^{1,1} \times \mathbb{R}^s \times \mathbb{R}^k$, in which the mapping $\hat{\mathcal{F}}$ is defined. Note that the disturbed system (52) has the form $z \in \hat{\mathcal{F}}(s)$ with $s = (x, u, p, \beta, \alpha) \in \hat{\mathcal{X}}$ and $z = (\pi, \nu, \rho, \eta, \mu, \xi) \in \hat{\mathcal{Z}}$. Obviously, it is equivalent to the system (45)-(47) complemented by the conditions $h(x, u) - \dot{x} = \eta$, $\psi(q) = \beta$, and $\varphi(q) - \xi \in N_{\mathbb{R}^k_+}(\alpha)$.

The following theorem claiming the subregularity property SMSr2 of the Pontryagin mapping $\hat{\mathcal{F}}$ with the space $\hat{\mathcal{Z}}$ of disturbances (almost directly) follows from Theorem 4.2.

Theorem 4.3 *Let $\hat{s} = (\hat{x}, \hat{u}, \hat{p}, \hat{\beta}, \hat{\alpha}) \in \hat{\mathcal{X}}$ be a solution of the inclusion $0 \in \hat{\mathcal{F}}(s)$, and let assumptions 4.2–4.4 be fulfilled at this point. Then there exist constants $a > 0$, $b > 0$ and $\kappa \geq 0$ such that for every $z = (\pi, \nu, \rho, \eta, \mu, \xi) \in \hat{\mathcal{Z}}$ with $\|z\|^\# := \|\pi\|_1 + |\nu| + \|\rho\|_2 + \|\eta\|_1 + |\mu| + |\xi| \leq b$ and for every solution $s = (x, u, p, \beta, \alpha) \in \hat{\mathcal{X}}$ of the inclusion $z \in \hat{\mathcal{F}}(s)$ with $\|x - \hat{x}\|_{1,1} + \|u - \hat{u}\|_\infty \leq a$ it holds that*

$$\|x - \hat{x}\|_{1,1} + \|u - \hat{u}\|_2 + \|p - \hat{p}\|_{1,1} + |\beta - \hat{\beta}| + |\alpha - \hat{\alpha}| \leq \kappa \|z\|^\#. \tag{53}$$

Proof. Due to Lemma 4.1, the inclusion $(\pi, \nu, \rho, \eta, \mu, \xi) \in \hat{\mathcal{F}}(s)$ with $s = (x, u, p, \beta, \alpha) \in \hat{\mathcal{X}}$ implies the inclusion $(\zeta, \eta, \mu, \xi) \in \mathcal{F}(s)$ with ζ given by (44). Then from Theorem 4.2 we obtain the estimation

$$\|x - \hat{x}\|_{1,1} + \|u - \hat{u}\|_2 + \|p - \hat{p}\|_\infty + |\beta - \hat{\beta}| + |\alpha - \hat{\alpha}| \leq \kappa(\|\zeta\|'' + \|\eta\|_1 + |\mu| + |\xi|).$$

Using (44), we estimate

$$\begin{aligned} \|\zeta\|'' &= \sup_{\|w\|' \leq 1} |\zeta w| \leq \sup_{\|w\|' \leq 1} (|\nu|\, |q| + \|\pi\|_1\, \|x\|_\infty + \|\rho\|_2\, \|u\|_2) \\ &\leq \sup_{\|w\|' \leq 1} (|\nu| + \|\pi\|_1 + \|\rho\|_2) \|w\|' \leq \|z\|^\#. \end{aligned}$$

Consequently,

$$\|x - \hat{x}\|_{1,1} + \|u - \hat{u}\|_2 + \|p - \hat{p}\|_\infty + |\beta - \hat{\beta}| + |\alpha - \hat{\alpha}| \leq 2\kappa \|z\|^\#, \tag{54}$$

It remains to estimate $\|p-\hat{p}\|_{1,1}$ using the already obtained estimation and (51). With an appropriate constant c_1 we have

$$\begin{aligned}\|p-\hat{p}\|_{1,1} &= |p(t_0)-\hat{p}(t_0)| + \|\dot{p}-\dot{\hat{p}}\|_1 \\ &\le |\varphi_{0q_0}(q)-\varphi_{0q_0}(\hat{q})| + \Big|\sum_{i\in I}\alpha_i\varphi_{iq_0}(q) - \sum_{i\in I}\hat{\alpha}_i\varphi_{iq_0}(\hat{q})\Big| \\ &\quad + \big|\beta\psi_{q_0}(q)-\hat{\beta}\psi_{q_0}(\hat{q})\big| + \|ph_x(w)-\hat{p}h_x(\hat{w})\|_1 \\ &\le c_1(|q-\hat{q}|+|\alpha-\hat{\alpha}|+|\beta-\hat{\beta}| \\ &\quad +\|p-\hat{p}\|_\infty + \|x-\hat{x}\|_\infty + \|u-\hat{u}\|_2).\end{aligned}$$

Then using (54) we complete the proof. □

We mention that Theorem 4.3 is more general than Theorem 4.2, because it allows for more general class of disturbances ζ, not necessarily representable in the form (44). Such disturbances may lead to discontinuous multipliers p. However, their practical relevance is unclear.

4.3 Estimation in the L^∞-norm for the controls

Observe that the estimation for the controls in (53) is with respect to the L^2-norm. The utilization of the space L^2 for the control functions was essential for the formulation of the coercivity condition. On the other hand, this choice of the space allows to involve disturbances ρ which are small in L^2 but not necessarily in L^∞.

However, the regularity results in [7, 8] under coercivity include L^∞ estimation for the control, given that ρ is sufficiently small in L^∞. This is, because, as shown in [7] the L^∞ estimation follows from the strengthened Legendre-Klebsch condition, which in its turn follows from coercitivity. All this concerns problems without initial and terminal constraints for the state. In order to obtain a similar result for problem (3)-(6) we first prove that the strong Legendre-Klebsch condition is fulfilled also for this problem.

Suppose that Assumptions 4.2–4.4 are fulfilled for the reference point $(\hat{x},\hat{u},\hat{p},\hat{\beta},\hat{\alpha})$. Then, by Theorem 3.2, the cost φ_0 satisfies the quadratic growth condition: there exist $c>0$ and $\varepsilon>0$ such that

$$\varphi_0(q)-\varphi_0(\hat{q}) \ge c(\|u-\hat{u}\|_2^2 + \|x-\hat{x}\|_{1,1}^2)$$

for all admissible $w=(x,u)$ such that $\|x-\hat{x}\|_{1,1}<\varepsilon$ and $\|u-\hat{u}\|_\infty<\varepsilon$. This implies that $(\hat{y},\hat{w})$ is a strong minimum in the problem

$$\text{minimize} \quad \varphi_0(q)-(y(t_1)-y(t_0)),$$

subject to

$$\dot{y}=c|u-\hat{u}(t)|^2, \quad \dot{x}=h(x,u), \quad \varphi(q)\le 0, \quad \psi(q)=0, \quad u\in U(t),$$

where $U(t)=\{v\in\mathbb{R}^m : |v-\hat{u}(t)|<\varepsilon\}$ and $\hat{y}=0$. Introduce

$$\begin{aligned}\tilde{H} &= ph(x,u)+p^yc|u-\hat{u}(t)|^2 = H(x,u,p)+p^yc|u-\hat{u}(t)|^2, \\ \tilde{l} &= \alpha_0(\varphi_0(q)-(y_1-y_0))+\alpha\varphi(q)+\beta\psi(q).\end{aligned}$$

Then $(\hat{y}, \hat{w})$ satisfies the conditions of Pontryagin minimum principle: there are $\alpha_0 \in \mathbb{R}_+$, $\alpha \in \mathbb{R}^k_+$, $\beta \in \mathbb{R}^s$, $p \in W^{1,1}$, $p^y \in W^{1,1}$ such that

$$\alpha\varphi(\hat{q}) = 0, \quad \alpha_0 + |\alpha| + |\beta| > 0, \quad -\dot{p} = ph_x(\hat{w}), \quad -\dot{p}^y = 0,$$
$$(-p(t_0), p(t_1)) = \tilde{l}_q, \quad p^y(t_0) = p^y(t_1) = -\alpha_0,$$
$$H(\hat{x}(t), u, p(t)) + p^y c|u - \hat{u}(t)|^2 \geq H(\hat{x}(t), \hat{u}(t), p(t)) \text{ for all } u \in U(t).$$

(The latter holds for a.a. $t \in [t_0, t_1]$.) It follows that $p^y = -\alpha_0$. Assume that $\alpha_0 = 0$. Then we obtain the conditions of the minimum principle subject to the constraint $|u - \hat{u}(t)| < \varepsilon$, which implies the local minimum principle with $\alpha_0 = 0$ in problem (3)-(6). As we know, this is impossible. Consequently, we can take $\alpha_0 = 1$. Then, as we know, $\alpha = \hat{\alpha}$, $\beta = \hat{\beta}$, $p = \hat{p}$ and we get for a.a. $t \in [t_0, t_1]$

$$H(\hat{x}(t), u, \hat{p}(t)) - H(\hat{x}(t), \hat{u}(t), \hat{p}(t)) \geq c|u - \hat{u}(t)|^2$$

for all $u \in \mathbb{R}^m$ such that $|u - \hat{u}(t)| < \varepsilon$. The strengthened Legendre-Klebsch condition follows: for a.a. $t \in [t_0, t_1]$

$$\langle H_{uu}(\hat{x}(t), \hat{u}(t), \hat{p}(t))v, v\rangle \geq c|v|^2 \quad \forall v \in \mathbb{R}^m. \tag{55}$$

Theorem 4.4 *Let $\hat{s} = (\hat{x}, \hat{u}, \hat{p}, \hat{\beta}, \hat{\alpha}) \in \hat{\mathcal{X}}$ be a solution of the inclusion $0 \in \hat{\mathcal{F}}(s)$, and let assumptions 4.2–4.4 be fulfilled at this point. Then there exist constants $a > 0$, $b > 0$ and $\kappa \geq 0$ such that for every $z = (\pi, \nu, \rho, \eta, \mu, \xi) \in \hat{\mathcal{Z}}$ with $\|z\|^\circ := \|\pi\|_1 + |\nu| + \|\rho\|_\infty + \|\eta\|_1 + |\mu| + |\xi| \leq b$ and for every solution $s = (x, u, p, \beta, \alpha) \in \hat{\mathcal{X}}$ of the inclusion $z \in \hat{\mathcal{F}}(s)$ with $\|x - \hat{x}\|_{1,1} + \|u - \hat{u}\|_\infty \leq a$ it holds that*

$$\|x - \hat{x}\|_{1,1} + \|u - \hat{u}\|_\infty + \|p - \hat{p}\|_{1,1} + |\beta - \hat{\beta}| + |\alpha - \hat{\alpha}| \leq \kappa\|z\|^\circ. \tag{56}$$

Proof. Let $\Theta \subset [t_0, t_1]$ be a set of full measure in which (55) is fulfilled. Then the condition number of the matrix $H_{uu}(\hat{x}(t), \hat{u}(t), \hat{p}(t))$ is uniformly bounded for $t \in \Theta$.

Let (x, u, p, β, α) be as in the theorem. For a fixed $t \in \Theta$ we have (after a possible redefinition of ρ on a set of measure zero)

$$\rho(t) = H_u(x(t), u(t), p(t)) = H_u(\hat{x}(t), u(t), \hat{p}(t)) - r(t),$$

where, according to (53),

$$|r(t)| = |H_u(x(t), u(t), p(t)) - H_u(\hat{x}(t), u(t), \hat{p}(t))| \leq c_1\|z\|^{\#}.$$

Here c_1 is a constant independent of t. Then from the equations

$$H_u(\hat{x}(t), \hat{u}(t), \hat{p}(t)) = 0, \qquad H_u(\hat{x}(t), u(t), \hat{p}(t)) = \rho(t) + r(t),$$

the inverse function theorem, and the uniform boundedness of the condition number of $H_{uu}(\hat{x}(t), \hat{u}(t), \hat{p}(t))$ we obtain that for a constant c_2 it holds that

$$|u(t) - \hat{u}(t)| \leq c_2(|\rho(t)| + |r(t)|) \leq c_2\|\rho\|_\infty + c_1c_2\|z\|^{\#}.$$

Recall that this inequality holds on the set Θ of a full measure in $[t_0, t_1]$. This completes the proof. □

References

[1] W. Alt, U. Felgenhauer and M. Seydenschwanz. Euler discretization for a class of nonlinear optimal control problems with control appearing linearly. *Computational Optimization and Applications*, **69**:825–856, 2018.

[2] J.F. Bonnans. Local Analysis of Newton-Type Methods for Variational Inequalities and Nonlinear Programming. *Applied Mathematics and Optimization*, **29**:161–186, 1994.

[3] J.F. Bonnans and A Shapiro. *Perturbation analysis of optimization problems.* Springer, 2000.

[4] R. Cibulka, A.L. Dontchev, and A.Y. Kruger. Strong metric subregularity of mappings in variational analysis and optimization. *Journal of Mathematical Analysis and Application*, **457**:1247–1282, 2018.

[5] A.V. Dmitruk and N.P. Osmolovskii. A General Lagrange Multipliers Theorem and Related Questions, Control Systems and Mathematical Methods in Economics (Feichtinger et al. eds.), Lecture Notes in Economics and Mathematical Systems, **687**:165–194, Springer, 2018.

[6] A. L. Dontchev. Characterizations of Lipschitz stability in optimization. In *Recent developments in well-posed variational varaitional problems*, 95–115, Kluwer, 1995.

[7] A. L. Dontchev, W. W. Hager. Lipschitzian stability in nonlinear control and optimization. *SIAM J. Control Optim.*, **31**:569–603, 1993.

[8] A.D. Dontchev and K. Malanowski. A characterization of Lipschitzian stability in optimal control. In *Calculus of variations and optimal control*, A Ioffe, S. Reich and I. Shafrir, eds. Chapman Hall/CRS Res. Math. 411, pp. 62–76, 2000.

[9] A. L. Dontchev, R. T. Rockafellar. Characterizations of Lipschitz stability in nonlinear programming *Mathematical programming with data perturbations*, 65–82, Lecture Notes in Pure and Appl. Math., 195, Dekker, New York, 1998.

[10] A. L. Dontchev, R. T. Rockafellar. Regularity and conditioning of solution mappings in variational analysis. *Set-Valued Analysis*, **12**:79–109, 2004.

[11] A.L. Dontchev and R.T. Rockafellar. *Implicit Functions and Solution Mappings: A View from Variational Analysis. Second edition.* Springer, New York , 2014.

[12] A. Ya. Dubovitskii, and A.A. Milyutin. Extremum problems in the presence of restrictions, *USSR Comput. Math. and Math. Phys.,* **5**(3):1–80, 1965.

[13] A.J. Hoffman, On approximate solutions of systems of linear inequalities. *J. Res. Nat'l Bur. Standarts* **49**:263–265, 1952.

[14] A.D. Ioffe. Regular points of Lipschitz functions. *Trans. Amer. Math. Sci.*, **251**:61–69, 1979.

[15] A.D. Ioffe. *Variational analysis and regular mappings.* Springer, 2017.

[16] A.D Ioffe and V.M. Tikhomorov. *Theory of extremal problems.* Nauka, Moscow, 1974 (in Russian); English translation: North Holland, 1979.

[17] D. Klatte and B. Kummer. *Nonsmooth equations in optimization.* Kluwer Academic Publisher, 2002.

[18] Nonsmooth Kantorovich–Newton Methods: Hypotheses and Auxiliary Problems. *Vietnam Journal of Mathematics*, **47**(3):639–657, 2019.

[19] J. Kyparisis. On uniqueness of Kuhn-Tucker multiplies in non-linear programming. *Math. Programming*, **32**:242–246, 1985.

[20] E.S. Levitin, A.A. Milyutin and N.P. Osmolovskii. Higher-order local minimum conditions in problems with constraints. *Uspekhi Mat. Nauk*, **33**:8–148, 1978; English translation in *Russian Math. Surveys*, **33**:9–168, 1978.

[21] K. Malanowski. Two norm approach in stability analysis of optimization and optimal control problems. *WCNA '92: Proceedings of the first world congress on World congress of nonlinear analysts*, **3**:2557–2568, May, 1996.

[22] K. Malanowski. Stability and sensitivity of solutions to nonlinear optimal control problems *Applied Mathematics and Optimization*, **32**:111–141, 1995.

[23] H. Maurer. First- and second-order sufficient optimality conditions in mathematical programming and optimal control. *Math. Programming Study*, **14**:163–177, 1981.

[24] A.A. Milyutin and N.P. Osmolovskii. *Calculus of variations and optimal control.* Amer. Math. Soc., 1998.

[25] N.P. Osmolovskii. Necessary second-order conditions for a weak local minimum in a problem with endpoint and control constraints. *J. Math. Anal. Appl.*, **457**(2):1613–1633, 2018.

[26] J. Preininger, T. Scarinci, and V.M. Veliov. Metric regularity properties in bang-bang type linear-quadratic optimal control problems. *Set-Valued and Variational Analysis*, **27**:381–404, 2019.