TECHNISCHE UNIVERSITÄT WIEN

Operations Research and Control Systems

SWM ORCOS

On the regularity of Mayer-type affine optimal control problems

Nikolai P. Osmolovskii, Vladimir M. Veliov

Research Report 2019-07

March 2019

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 / E105-4
1040 Vienna, Austria
E-mail: orcos@tuwien.ac.at

On the regularity of Mayer-type affine optimal control problems*

Nikolai P. Osmolovskii[1,2,3] and Vladimir M. Veliov[4]

[1] Department of Informatics and Mathematics, University of Technology and Humanities in Radom, Radom, Poland
[2] Systems Research Institute, Polish Academy of Sciences, Warszawa, Poland
[3] Department of Applied Mathematics, Moscow State University of Civil Engineering, Moscow, Russia, osmolovski@uph.edu.pl
[4] Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, veliov@tuwien.ac.at

Abstract. The paper presents a sufficient condition for strong metric sub-regularity (SMsR) of the system of first order optimality conditions (optimality system) for a Mayer-type optimal control problem with a dynamics affine with respect to the control. The SMsR property at a reference solution means that any solution of the optimality system, subjected to "small" disturbances, which is close enough to the reference one is at a distance to it, at most proportional to the size of the disturbance. The property is well understood for problems satisfying certain coercivity conditions, which however, are not fulfilled for affine problems.

1 Introduction

In the paper, we analyze the question of strong metric sub-regularity (SMsR) of the system of optimality conditions at a reference point $\hat{u}$ for the following affine optimal control problem, presented in the Mayer form:

$$\min g(x(T)) \tag{1}$$

subject to

$$\dot{x}(t) = a(t, x(t)) + B(t, x(t))u(t), \quad x(0) = x^0, \tag{2}$$

$$u(t) \in U, \quad t \in [0, T], \tag{3}$$

where the state x is a vector in $I\!R^n$, the control u has values $u(t)$ that belong to a given set U in $I\!R^m$ for almost every (a.e.) $t \in [0, T]$. The initial state x^0 and the final time $T > 0$ are fixed. The set of feasible control functions u, denoted in the sequel by $\mathcal{U}$, consists of all Lebesgue measurable and bounded functions $u : [0, T] \to U$. Accordingly, the state trajectories x, that are solutions of (2) for feasible controls, are Lipschitz continuous functions of time $t \in [0, T]$. For brevity we denote $f(t, x, u) := a(t, x) + B(t, x)u$.

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

Assumption (A1). The set U is convex and compact, the components of the functions a, B, g (of dimensions $n \times 1$, $n \times m$ and 1, correspondingly) are two times differentiable with respect to x, the second derivatives are continuous in x locally uniformly in t,[5] a and B and their first and second derivatives in x are measurable and bounded in t.

Here, and in the squeal, we use the following standard notations. The euclidean norm and the scalar product in $I\!R^n$ (the elements of which are regarded as column-vectors) are denoted by $|\cdot|$ and $\langle \cdot, \cdot \rangle$, respectively. The transpose of a matrix (or vector) E is denoted by $E^\top$. For a function $\psi : I\!R^p \to I\!R^r$ of the variable z we denote by $\psi_z(z)$ its derivative (Jacobian), represented by an $(r \times p)$-matrix. If $r = 1$, $\nabla_z \psi(z) = \psi_z(z)^\top$ denotes its gradient (a vector-column of dimension p). Also for $r = 1$, $\psi_{zz}(z)$ denotes the second derivative (Hessian), represented by a $(p \times p)$-matrix. For a function $\psi : I\!R^{p \times q} \to I\!R$ of the variables (z, v), $\psi_{zv}(z, v)$ denotes its mixed second derivative, represented by a $(p \times q)$-matrix. The space $L^k = L^k([0,T], I\!R^r)$, with $k = 1, 2$ or $k = \infty$, consists of all (classes of equivalent) Lebesgue measurable r-dimensional vector-functions defined on the interval $[0,T]$, for which the standard norm $\|\cdot\|_k$ is finite. As usual, $W^{1,k} = W^{1,k}([0,T], I\!R^r)$ denotes the space of absolutely continuous r-dimensional vector-functions on $[0,T]$ for which the first derivative belongs to L^k, with the usual norm, $\|\cdot\|_{1,k}$. Often the specification $([0,T], I\!R^r)$ will be omitted in the notations. In any metric space we denote by $\mathbf{B}_a(x)$ the closed ball of radius a centered at x.

Define the Hamiltonian associated with problem (1)–(3) as usual:

$$H(t, x, u, p) := \langle p, f(t, x, u) \rangle, \quad p \in I\!R^n.$$

Although the feasible controls $u \in \mathcal{U}$ are bounded, we consider the control-trajectory pairs (x, u) as elements of the space $W^{1,1}([0,1], I\!R^n) \times L^1([0,1], I\!R^m)$.

The local form of the Pontryagin maximum (here minimum) principle for problem (1)-(3) can be represented by the following optimality system for (x, u) and an absolutely continuous (here Lipschitz) function $p : [0,T] \to I\!R^n$: for a.e. $t \in [0,T]$

$$0 = -\dot{x}(t) + f(t, x(t), u(t)), \quad x(0) - x^0 = 0, \tag{4}$$

$$0 = \dot{p}(t) + \nabla_x H(t, x(t), u(t), p(t)), \tag{5}$$

$$0 = p(T) - \nabla_x g(x(T)), \tag{6}$$

$$0 \in \nabla_u H(t, x(t), u(t), p(t)) + N_U(u(t)), \tag{7}$$

where the normal cone $N_U(u)$ to the set U at $u \in I\!R^m$ is defined as

$$N_U(u) = \begin{cases} \{y \in I\!R^n \mid \langle y, v - u \rangle \le 0 \text{ for all } v \in U\} & \text{if } u \in U, \\ \emptyset & \text{otherwise.} \end{cases}$$

[5] Applied to a, for example, this means that for every bounded set $S \subset I\!R^n$ there exists a function (called *modulus of continuity*) $\omega : (0, +\infty) \to [0, +\infty)$ with $\lim_{s \to 0} \omega(s) = 0$, such that $|a(t, x') - a(t, x)| \le \omega(|x' - x|)$ for every $t \in [0,T]$ and x, $x' \in S$.

Introduce the spaces

$$\mathcal{Y} := W_0^{1,1} \times \mathcal{U} \times W^{1,1}, \qquad \mathcal{Z} := L^1 \times L^1 \times I\!R^n \times L^\infty,$$

where $W_0^{1,1}$ is the affine space consisting of those $x \in W^{1,1}$ for which $x(0) = x^0$ and $\mathcal{U}$ is endowed with the L^1-metric. Norms in this spaces are also defined as usual: for $y = (x, u, p) \in \mathcal{Y}$ and $z = (\xi, \pi, \nu, \rho) \in \mathcal{Z}$

$$\|y\| = \|x\|_{1,1} + \|u\|_1 + \|p\|_{1,1}, \qquad \|z\| = \|\xi\|_1 + \|\ \pi\|_1 + |\nu| + \|\rho\|_\infty.$$

Then the optimality system (4)–(7) can be recast as the generalized equation

$$0 \in \psi(y) + \Psi(y), \tag{8}$$

where $y = (x, u, p)$ and

$$\mathcal{Y} \ni y \mapsto \psi(y) := \begin{pmatrix} -\dot{x} + f(\cdot, x, u) \\ \dot{p} + \nabla_x H(\cdot, y) \\ p(T) - \nabla_x g(x(T)) \\ \nabla_u H(\cdot, y) \end{pmatrix} \in \mathcal{Z}, \quad \mathcal{Y} \ni y \rightrightarrows \Psi(y) := \begin{pmatrix} 0 \\ 0 \\ 0 \\ N_{\mathcal{U}}(u) \end{pmatrix} \subset \mathcal{Z}.$$

Here $N_{\mathcal{U}}(u) := \{v \in L^\infty : v(t) \in N_U(u(t)) \text{ for a.e. } t \in [0, T]\}$ is the normal cone to the set $\mathcal{U}$ (considered as a subset of L^1) at $u \in \mathcal{U}$. For $u \notin \mathcal{U}$ the normal cone is empty.

Below we remind one of the several equivalent definitions of the notion of strong metric sub-regularity, adapted to our notations, see e.g. [4, p. 202].

Definition 1 *The set-valued mapping $\psi + \Psi : \mathcal{Y} \rightrightarrows \mathcal{Z}$ is* Strongly Metrically sub-Regular *(SMsR) at the point $\hat{y}$ for $\hat{z}$ if $\hat{z} \in \psi(\hat{y}) + \Psi(\hat{y})$ and there exist numbers $\alpha_0 > 0$, $\beta_0 > 0$ and c_0 such that for any $z \in \mathcal{Z}$ with $\|z\| \le \alpha_0$ and for any solution $y \in \mathcal{Y}$ of the inclusion $z \in \psi(y) + \Psi(y)$ with $\|y - \hat{y}\| \le \beta_0$, it holds that $\|y - \hat{y}\| \le c_0 \|z - \hat{z}\|$.*

We will prove the SMsR property of the mapping $\psi + \Psi$ in the optimality system (8) under an additional assumption, given in the next section. At the end of the next section we will also compare our result with the few existing ones. Here we only mention that in contrast to the case of the so-called "coercive" problems, where a Legendre-type condition is satisfied, the investigation of regularity properties for affine control problems started just a few years ago and is still in progress.

2 Main result

Let a solution $\hat{y} = (\hat{x}, \hat{u}, \hat{p}) \in W^{1,1} \times \mathcal{U} \times W^{1,1}$ of the optimality system (4)–(7) be fixed. To shorten the notations we skip arguments with "hat" in functions, shifting the "hat" on the top of the notation of the function, so that $\hat{f}(t) := f(t, \hat{x}(t), \hat{u}(t))$, $\hat{B}(t) := B(t, \hat{x}(t))$, $\hat{H}(t) := H(t, \hat{x}(t), \hat{u}(t), \hat{p}(t))$, $\hat{H}(t, u) := H(t, \hat{x}(t), u, \hat{p}(t))$, etc. Moreover, denote

$$\hat{A}(t) := \hat{f}_x(t) = f_x(t, \hat{x}(t), \hat{u}(t)), \qquad \hat{\sigma}(t) := \nabla_u \hat{H}(t, \hat{x}(t), \hat{u}(t), \hat{p}(t)) = \hat{B}(t)^\top \hat{p}(t).$$

Remark 1 Due to Assumption (A1), and since the solution $\hat{x}$ of (2) with $u = \hat{u}$, exists on $[0,T]$, there exist a number $r > 0$ and a convex compact set $\bar{S} \subset \mathbb{R}^n$ such that for every $u \in \mathcal{U}$ with $\|u - \hat{u}\|_1 \leq r$ the solution x of (2) exists on $[0,T]$ and $\mathbf{B}_1(x(t)) \subset \bar{S}$ for all $t \in [0,T]$. By taking $\bar{S}$ sufficiently large we may also ensure that $\mathbf{B}_1(\hat{p}(t)) \subset \bar{S}$ for all $t \in [0,T]$. Using Assumption (A1), we denote by L a Lipschitz constant with respect to $x \in \bar{S}$ (uniformly with respect to $t \in [0,T]$, $u \in U$, $p \in \bar{S}$) of the functions a, B, g, f, and H and their first derivatives in x. Further, we denote by M a bound of all these functions, their first and second derivatives in x, for $(t,x,u,p) \in [0,T] \times \bar{S} \times U \times \bar{S}$. Finally, we denote by $\bar{\omega}$ a modulus of continuity of the second derivative in x of the functions a, B, g, f, and H, uniformly with respect to $(t,u,p) \in [0,T] \times U \times \bar{S}$ (see Footnote 5). Due to the Grönwal inequality, the following estimation holds for every $u \in \mathcal{U}$ with $\|u - \hat{u}\|_1 \leq r$: $\|x - \hat{x}\|_C \leq c_f \|u - \hat{u}\|_1$ where $c_f = M e^{LT}$.

According to this remark, for any $u \in \mathcal{U}$ with $\|u - \hat{u}\|_1 \leq r$ the value of the objective functional $J(u) := g(x(T))$ is well defined. For any function $\delta u \in \mathcal{U} - \hat{u}$ we introduce the linearized version of equation (2):

$$\delta \dot{x}(t) = \hat{A}(t)\delta x + \hat{B}(t)\delta u(t), \quad \delta x(0) = 0, \quad t \in [0,T]. \tag{9}$$

Denote by Γ the set of all pairs $(\delta x, \delta u) \in W^{1,1} \times L^1$ such that $\delta u \in \mathcal{U} - \hat{u}$ and δx is the solution of the linearized equation (9). Let us introduce the following quadratic functional of $(\delta x, \delta u) \in W^{1,1} \times L^1$:

$$\begin{aligned} \Omega(\delta x, \delta u) := & \frac{1}{2}\langle g_{xx}(\hat{x}(T))\, \delta x(T), \delta x(T)\rangle \\ & + \int_0^T \left[\frac{1}{2}\langle \hat{H}_{xx}(t)\delta x(t), \delta x(t)\rangle + \langle \hat{H}_{ux}(t)\delta x(t), \delta u(t)\rangle\right] \mathrm{d}t. \end{aligned} \tag{10}$$

Assumption (A2). There exists a constant $c_0 > 0$ such that

$$\int_0^T \langle \hat{\sigma}(t), \delta u(t)\rangle \,\mathrm{d}t + 2\,\Omega(\delta x, \delta u) \geq c_0 \|\delta u\|_1^2 \quad \text{for all } (\delta x, \delta u) \in \Gamma. \tag{11}$$

Consider the following "perturbed" version of (8):

$$z \in \psi(y) + \Psi(y), \tag{12}$$

where $z = (\xi, \pi, \nu, \rho) \in \mathcal{Z}$.

Theorem 1 *Let assumptions (A1) and (A2) be fulfilled. Then there exist constants $\alpha_0 > 0$, $\beta_0 > 0$ and c_0 such that for any $z \in \mathcal{Z}$ with $\|z\| \leq \alpha_0$ and for any solution $y = (x,u,p) \in \mathcal{Y}$ of the disturbed optimality system (12), with $\|u - \hat{u}\|_1 \leq \beta_0$, it holds that*

$$\|x - \hat{x}\|_{1,1} + \|u - \hat{u}\|_1 + \|p - \hat{p}\|_{1,1} \leq c_0 \|z\|. \tag{13}$$

Thus, the mapping $\psi + \Psi$, associated with problem (1)-(3), is strongly metrically sub-regular at $\hat{y} = (\hat{x}, \hat{u}, \hat{p})$ for zero.

It is interesting to note that Assumption (A2) guarantees not only strong metric sub-regularity of the mapping $\psi + \Psi$ at the point $(\hat{x}, \hat{u}, \hat{p})$, but also that $\hat{u}$ is a strict L^1-local minimizer. Moreover, it can be proved that even (formally) a weaker condition than (A2) guarantees such a minimum at $\hat{u}$. This weakened condition differs from (A2) only by changing 2Ω with Ω in the inequality (11) (and then it follows from (A2) since $\langle \hat{\sigma}(t), \delta u(t) \rangle \geq 0$ a.e. in $[0, T]$ for all $\delta u \in \mathcal{U} - \hat{u}$).

The following more demanding condition than (A2) appears in [1, Section 5] and is used for error analysis of the Euler discretization scheme:

$$\int_0^T \langle \hat{\sigma}(t), \delta u(t) \rangle \, \mathrm{d}t \geq c_1 \|\delta u\|_1^2 \qquad \text{for all } (\delta x, \delta u) \in \Gamma, \tag{14}$$
$$2\,\Omega(\delta x, \delta u) \geq c_2 \|\delta u\|_1^2 \qquad \text{for all } (\delta x, \delta u) \in \Gamma$$

with $c_1 + c_2 > 0$. Sufficient conditions for (14) are known in the case of a box-like set U; see the forthcoming paper [8] for the relevant bibliography, and also for the case of a general compact convex polyhedral set U.

3 Proof of Theorem 1

We give the proof omitting some details. The positive numbers α_0 and β_0 will be fixed later as depending only on L, M and $\bar{\omega}$ (see Remark 1). Now take an arbitrary $z = (\xi, \pi, \nu, \rho) \in \mathcal{Z}$ with $\|z\| \leq \alpha_0$ and a solution $y = (x, u, p) \in \mathcal{Y}$ of (12) satisfying $\|u - \hat{u}\|_1 \leq \beta_0$. In detail, inclusion (12) reads as

$$\dot{x}(t) = f(t, x(t), u(t)) - \xi(t), \quad x(0) = x^0, \tag{15}$$
$$\dot{p}(t) = -\nabla_x H(t, x(t), u(t), p(t)) + \pi(t), \tag{16}$$
$$p(T) = \nabla_x g(x(T)) + \nu, \tag{17}$$
$$-N_U(u(t)) \ni \nabla_u H(t, x(t), u(t), p(t)) - \rho(t), \tag{18}$$

where differential equations (15), (16) and inclusion (18) have to be fulfilled for a.e. $t \in [0, T]$. Denote $\Delta x(t) = x(t) - \hat{x}(t)$, $\Delta \dot{x}(t) = \dot{x}(t) - \dot{\hat{x}}(t)$, $\Delta u(t) = u(t) - \hat{u}(t) = \delta u(t)$, $\Delta f(t) = f(t, x(t), u(t)) - \hat{f}(t)$, etc. Then

$$\Delta \dot{x}(t) = \Delta f(t) - \xi(t), \quad \Delta x(0) = 0, \tag{19}$$

$$\Delta \dot{p}(t) = -\Delta(\nabla_x H)(t) + \pi(t), \quad \Delta p(T) = \Delta(\nabla_x g)(T) + \nu, \tag{20}$$

$$\Delta \sigma(t) = \Delta(\nabla_u H)(t) - \rho(t). \tag{21}$$

Applying the Grönwall inequality to equation (19) we obtain

$$\|\Delta x\|_C \leq e^{LT}(M\|\delta u\|_1 + \|\xi\|_1) \leq c_f(\|\delta u\|_1 + \|\xi\|_1) \leq c_f(\|\delta u\|_1 + \|z\|), \tag{22}$$

where $c_f = e^{LT} \max\{M, 1\}$. Similarly applying the Grönwall inequality to equation (20) and using (22) we get

$$\|\Delta p\|_C \leq c_H(\|\delta u\|_1 + \|z\|), \tag{23}$$

where $c_H = e^{MT}\max\{L(T+1)c_f + M, 1\}$. The following obvious equality serves as a source for further key estimations:

$$\int_0^T \langle \Delta\dot p(t), \Delta x(t)\rangle \,\mathrm{d}t + \int_0^T \langle \Delta p(t), \Delta\dot x(t)\rangle \,\mathrm{d}t = \langle \Delta p(T), \Delta x(T)\rangle,$$

whence in view of (19) and (20), we get

$$\langle \Delta(\nabla_x g), \Delta x(T)\rangle + \int_0^T \Big(\langle \Delta(\nabla_x H)(t), \Delta x(t)\rangle - \langle \Delta p(t), \Delta f(t)\rangle\Big)\,\mathrm{d}t$$

$$= -\langle \nu, \Delta x(T)\rangle + \int_0^T \Big(\langle \pi(t), \Delta x(t)\rangle - \langle \Delta p(t), \xi(t)\rangle\Big)\,\mathrm{d}t. \qquad (24)$$

Further, simple expansions and transformations show that

$$\begin{aligned}
\langle \Delta(\nabla_x H)(t), \Delta x(t)\rangle &= \langle \hat H_{xu}(t)\delta u(t), \Delta x(t)\rangle + \langle \hat H_{xx}(t)\Delta x(t), \Delta x(t)\rangle \\
&\quad + \langle \hat H_{xp}(t)\Delta p(t), \Delta x(t)\rangle + r_{H_x}, \\
\langle \Delta p(t), \Delta f(t)\rangle &= \langle \hat H_{pu}(t)\delta u(t), \Delta p(t)\rangle + \langle \hat H_{px}(t)\Delta x(t), \Delta p(t)\rangle + r_{H_p}(t), \\
\langle \hat H_{pu}(t)\delta u(t), \Delta p(t)\rangle &= \langle \Delta(\nabla_u H)(t), \delta u(t)\rangle - \langle \hat H_{ux}(t)\Delta x(t), \delta u(t)\rangle + r_\sigma(t),
\end{aligned}$$

where

$$\begin{aligned}
\|r_{H_x}\|_1 &\le \big(T\bar\omega(\|\Delta x\|_C) + MT\|\Delta p\|_C + M\|\delta u\|_1\big)\|\Delta x\|_C^2 + M\|\Delta x\|_C\|\Delta p\|_C\|\delta u\|_1, \\
\|r_{H_p}\|_1 &\le \big(T\bar\omega(\|\Delta x\|_C) + M\|\delta u\|_1\big)\|\Delta x\|_C\|\Delta p\|_C, \\
\|r_\sigma\|_1 &\le \bar\omega(\|\Delta x\|_C)\|\Delta x\|_C\|\delta u\|_1 + M\|\Delta x\|_C\|\delta u\|_1\|\Delta p\|_C.
\end{aligned}$$

Using these formulas together with (22) and (23), we obtain

$$\langle \Delta(\nabla_x g), \Delta x(T)\rangle + \int_0^T \Big(\langle \Delta(\nabla_x H)(t), \Delta x(t)\rangle - \langle \Delta p(t), \Delta f(t)\rangle\Big)\,\mathrm{d}t$$

$$= 2\Omega(\Delta x, \delta u) - \int_0^T \langle \Delta(\nabla_u H)(t), \delta u(t)\rangle \,\mathrm{d}t + r_\Omega, \qquad (25)$$

where

$$|r_\Omega| \le c_{r_\Omega}[\bar\omega(c_f(\|z\| + \|\delta u\|_1)) + \|z\| + \|\delta u\|_1](\|z\| + \|\delta u\|_1)^2, \qquad (26)$$

and $c_{r_\Omega} > 0$ depends only on L, M, T (therefore also on $\hat x, \hat u$).

The next step involves replacing $(\Delta x, \delta u)$ with $(\delta x, \delta u) \in \Gamma$ in the quadratic form Ω. Let $\delta x(t)$ be the solution of linear equation (9). Then it follows from (19) and (9) that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\Delta x(t) - \delta x(t)) = \hat A(t)(\Delta x(t) - \delta x(t)) + r_f(t) - \xi(t),$$

where $|r_f(t)| \le M|\Delta u(t)||\Delta x(t)| + \frac{1}{2}L|\Delta x(t)|^2$. Using the Grönwall inequality and estimate (22) we obtain that

$$\|\Delta x - \delta x\|_C \le e^{MT}(\|r_f\|_1 + \|\xi\|_1) \le c_f e^{MT}(M + \frac{1}{2}LTc_f)(\|\xi\|_1 + \|\delta u\|_1)^2 + e^{MT}\|\xi\|_1.$$

Since $\|\xi\|_1 \le \alpha_0$ and hence $(\|\xi\|_1 + \|\delta u\|_1)^2 \le 2\alpha_0\|\xi\|_1 + 2\|\delta u\|_1^2$, we get

$$\|\varDelta x - \delta x\|_C \le \tilde{d}\|\delta u\|_1^2 + \hat{d}\|\xi\|_1 \le \tilde{d}\|\delta u\|_1^2 + \hat{d}\|z\|_1, \tag{27}$$

where $\tilde{d}$ and $\hat{d}$ depend only on L, M, and T. Now we can estimate the difference $r_{\varDelta\varOmega}(\delta u) := \varOmega(\varDelta x, \delta u) - \varOmega(\delta x, \delta u)$ as follows:

$$|r_{\varDelta\varOmega}(\delta u)| \le M\Big[\frac{1}{2}(1+T)(\|\varDelta x\|_C + \|\delta x\|_C) + \|\delta u\|_1\Big]\|\varDelta x - \delta x\|_C.$$

In view of (9) we have $\|\delta x\|_C \le e^{MT}M\|\delta u\|_1 =: c_M\|\delta u\|_1$. Using this estimate together with (22) and (27), we obtain

$$|r_{\varDelta\varOmega}(\delta u)| \le c_{\varDelta\varOmega}(\|\delta u\|_1 + \|z\|)(\|\delta u\|_1^2 + \|z\|), \tag{28}$$

where $c_{\varDelta\varOmega} := M(\frac{1}{2}(1+T)c_f + c_M)(\tilde{d} + \hat{d})$. It follows from (25) and the definition of $r_{\varDelta\varOmega}$ that

$$\langle \varDelta(\nabla_x g), \varDelta x(T)\rangle + \int_0^T \Big(\langle \varDelta(\nabla_x H)(t), \delta x(t)\rangle - \langle \varDelta p(t), \varDelta f(t)\rangle\Big)\,\mathrm{d}t$$

$$= 2\varOmega(\delta x, \delta u) - \int_0^T \langle \varDelta(\nabla_u H)(t), \delta u(t)\rangle\,\mathrm{d}t + r_\varOmega + 2r_{\varDelta\varOmega}. \tag{29}$$

According to (21) $\sigma(t) - \hat{\sigma}(t) = \varDelta(\nabla_u H)(t) - \rho(t)$. This and the inequality $\langle \sigma(t), \delta u(t)\rangle \le 0$ a.e. in $(0,T)$, following from (18), imply

$$\int_0^T \langle \varDelta(\nabla_u H)(t), \delta u(t)\rangle\,\mathrm{d}t \le -\int_0^T \langle \hat{\sigma}(t), \delta u(t)\rangle\,\mathrm{d}t + \int_0^T \langle \rho(t), \delta u(t)\rangle\,\mathrm{d}t,$$

and then in view of Assumption (A2) we get from (29) that

$$\langle \varDelta(\nabla_x g), \varDelta x(T)\rangle + \int_0^T \Big(\langle \varDelta(\nabla_x H)(t), \delta x(t)\rangle - \langle \varDelta p(t), \varDelta f(t)\rangle\Big)\,\mathrm{d}t$$

$$\ge 2\varOmega(\delta x, \delta u) + \int_0^T \langle \hat{\sigma}(t), \delta u(t)\rangle\,\mathrm{d}t - \int_0^T \langle \rho(t), \delta u(t)\rangle\,\mathrm{d}t + r_\varOmega + 2r_{\varDelta\varOmega}$$

$$\ge c_0\|\delta u\|_1^2 - \int_0^T \langle \rho(t), \delta u(t)\rangle\,\mathrm{d}t + r_\varOmega + 2r_{\varDelta\varOmega}. \tag{30}$$

Combining this with (24) we get

$$c_0\|\delta u\|_1^2 + r_\varOmega + 2r_{\varDelta\varOmega} \le \int_0^T \Big(\langle \pi(t), \varDelta x(t)\rangle\rangle - \langle \varDelta p(t), \xi(t)\rangle + \langle \rho(t), \delta u(t)\rangle\Big)\,\mathrm{d}t$$

$$-\langle \nu, \varDelta x(T)\rangle \le (|\varDelta x(T)| + \|\varDelta x\|_C + \|\varDelta p\|_C + \|\delta u\|_1)\|z\|.$$

Using also (22) and (23) we obtain

$$c_0\|\delta u\|_1^2 + r_\varOmega + 2r_{\varDelta\varOmega} \le \hat{c}(\|\delta u\|_1 + \|z\|)\|z\|, \tag{31}$$

where $\hat{c} = (2c_f + c_H + 1)$. Take α_0 and β_0 such that $\bar{\omega}(c_f(\|z\| + \|\delta u\|_1)) + \|z\| + \|\delta u\|_1 \le \varepsilon$ for all $\|z\| \le \alpha_0$, $\|\delta u\|_1 \le \beta_0$, where $\varepsilon \in (0,1)$ will be defined later. Then in view of (26) and (28), $|r_\Omega| \le 2c_{r_\Omega}\varepsilon\|z\|^2 + 2c_{r_\Omega}\varepsilon\|\delta u\|_1^2$ and $|r_{\Delta\Omega}| \le c_{\Delta\Omega}\varepsilon\|\delta u\|_1^2 + c_{\Delta\Omega}\|z\|(\|\delta u\|_1 + \|z\|)$. Using these estimates and inequality (31) we get

$$(c_0 - 2c_{r_\Omega}\varepsilon - 2c_{\Delta\Omega}\varepsilon)\|\delta u\|_1^2 \le (\hat{c} + 2c_{r_\Omega}\varepsilon + 2c_{\Delta\Omega})(\|\delta u\|_1 + \|z\|)\|z\|.$$

Take $\varepsilon \in (0,1)$ such that $c_0 - 2c_{r_\Omega}\varepsilon - 2c_{\Delta\Omega}\varepsilon \ge \frac{1}{2}c_0$, and set $a = (\hat{c} + 2c_{r_\Omega} + 2c_{\Delta\Omega})/c_0$. Then $\|\delta u\|_1^2 \le 2a(\|\delta u\|_1 + \|z\|)\|z\|$, whence $(\|\delta u\|_1 - a\|z\|)^2 \le (a+1)^2\|z\|^2$, which implies $\|\delta u\|_1 \le (2a+1)\|z\|$. Combined with (22) and (23) this inequality 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 (2018), pp. 825-856.
2. W. Alt, C. Schneider and M. Seydenschwanz, *Regularization and implicit Euler discretization of linear-quadratic optimal control problems with bang-bang solutions*, Appl. Math. Comput., 287 (2016), pp. 104–124.
3. J.-P. Aubin and H. Frankowska. *Set-valued Analysis.* Birkhäuser, Boston, Basel, Berlin, 1990.
4. A.L. Dontchev, R.T. Rockafellar. *Implicit Functions and Solution Mappings: A View from Variational Analysis. Second edition.* Springer, New York, 2014.
5. U. Felgenhauer, *On stability of bang-bang type controls*, SIAM J. Control Optim., 41 (2003), pp.1843–1867.
6. U. Felgenhauer, *Discretization of semilinear bang-singular-bang control problems*, Computational Optimization and Applications, 64 (2016), pp. 295–326.
7. U. Felgenhauer, *A Newton-type method and optimality test for problems with bang-singular-bang optimal control.* Pure and Applied Functional Analysis, 1 (2016), pp. 197–215.
8. N.P. Osmolovskii and V.M. Veliov. Metric sub-regularity in optimal control of affine problems with free end state. Submitted; available as Research Report 2019-04, ORCOS, TU Wien, 2019, at https://orcos.tuwien.ac.at/research/research_reports.
9. A. Pietrus, T. Scarinci, and V.M. Veliov. *High order discrete approximations to Mayer's problems for linear systems.* SIAM J. Control Optim., 56 (2018), pp. 102–119.
10. J. Preininger, T. Scarinci and V.M. Veliov, *Metric regularity properties in bang-bang type linear-quadratic optimal control problems*, to appear in Set-Valued and Variational Analysis, https://doi.org/10.1007/s11228-018-0488-1.
11. M. Quincampoix and V. Veliov, *Metric Regularity and Stability of Optimal Control Problems for Linear Systems*, SIAM J. Control Optim, 51 (2013), pp. 4118–4137.
12. M. Seydenschwanz, *Convergence results for the discrete regularization of linear-quadratic control problems with bang-bang solutions*, Comput. Optim. Appl., 61 (2015) pp. 731–760.