ELSEVIER

Systems & Control Letters 35 (1998) 137–143

SYSTEMS & CONTROL LETTERS

A new approach to Lipschitz continuity in state constrained optimal control[1]

A.L. Dontchev [a], William W. Hager [b,*]

[a] *Mathematical Reviews, Ann Arbor, MI 48107, USA*
[b] *Department of Mathematics, University of Florida, Gainesville, FL 32611, USA*

Received 22 January 1998; received in revised form 19 March 1998

Abstract

For a linear-quadratic state constrained optimal control problem, it is proved in [11] that under an independence condition for the active constraints, the optimal control is Lipschitz continuous. We now give a new proof of this result based on an analysis of the Euler discretization given in [9]. There we exploit the Lipschitz continuity of the control to estimate the error in the Euler discretization. Here we show that the theory developed for the Euler discretization can be used to derive the Lipschitz continuity of the optimal control. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Optimal control; State constraints; Regularity; Discrete approximations

1. Introduction

We consider the following linear-quadratic problem with state constraints:

$$\text{minimize} \quad \frac{1}{2}\int_0^1 (x(t)^\mathsf{T} Qx(t) + u(t)^\mathsf{T} Ru(t))\,\mathrm{d}t \tag{1}$$

subject to

$$\dot{x}(t) = Ax(t) + Bu(t) \quad \text{for a.e. } t \in [0,1], \quad x(0) = a, \tag{2}$$

$$Kx(t) + b \leqslant 0 \quad \text{for all } t \in [0,1], \tag{3}$$

$$u \in L^2, \qquad x \in H^1, \tag{4}$$

where $x(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^m$, the matrices Q, R, A and B have compatible dimensions, K is a $k \times n$ matrix, $a \in \mathbb{R}^n$, and $b \in \mathbb{R}^k$. Throughout, L^p denotes the usual Lebesque space of functions with integrable p-power, $p \geqslant 1$, and H^1 is the space of absolutely continuous functions with derivatives in L^2. We assume that Q is positive semidefinite and R is positive definite, and that there exists a pair (x, u) satisfying the constraints (2)–(4). Under these conditions, there exists a unique optimal solution (x^*, u^*) of problem (1).

Let $\mathscr{A}$ denote the active constraint set defined by

$$\mathscr{A}(t) = \{ j \in \{1, 2, \dots, k\} \mid K_j x^*(t) + b_j = 0 \},$$

where K_j is the jth row of the matrix K. We assume that the constraints satisfy the following regularity condition introduced in [11]:

Independence: $\mathscr{A}(0) = \emptyset$ *and there exists* $\beta > 0$ *such that*

$$\left| \sum_{j \in \mathscr{A}(t)} v_j K_j B \right| \geqslant \beta |v| \tag{5}$$

* Corresponding author. E-mail: hager@math.ufl.edu., http://www.math.ufl.edu/~hager.

[1] This research was supported by the National Science Foundation.

0167-6911/98/$ – see front matter
PII: S0167-6911(98)00043-7

for all $t \in [0,1]$ *where* $\mathscr{A}(t) \neq \emptyset$ *and for all* $v = \{v_j \mid j \in \mathscr{A}(t)\}$.

As shown in ([8], Eq. (43)), this regularity condition can be extended to the set of ε-active constraints; that is, there exists an $\varepsilon > 0$ such that the independence condition holds with the set $\mathscr{A}$ replaced by

$$\mathscr{A}_\varepsilon(t) = \{ j \in \{1,2,\ldots,k\} \mid K_j x^*(t) + b_j \geqslant -\varepsilon \}.$$

The independence condition also implies that there exists a feasible control $\tilde{u}$ such that the corresponding trajectory $x = \tilde{x}$ associated with $u = \tilde{u}$ in Eq. (2) satisfies

$$K\tilde{x}(t) + b < 0 \quad \text{for all } t \in [0,1]. \tag{6}$$

This fact follows from the more general results contained in ([7], Lemma 3) or ([8], Lemma 3.6).

There are various forms available in the literature for the first-order optimality conditions (maximum principle) for the problem (1). For our purposes here, it is convenient to use the following duality based result from Hager and Mitter [12]: If there is a feasible pair $(\tilde{x}, \tilde{u})$ satisfying Eq. (6), then there exists a function v^* of bounded variation which is nondecreasing, right continuous, and with $v^*(1) = 0$, and a function $\psi^* \in H^1$ which satisfy the following relations for almost every $t \in [0,1]$:

$$u^*(t) = R^{-1} B^{\mathsf{T}} (K^{\mathsf{T}} v(t) - \psi(t)), \tag{7}$$

$$\dot{\psi}(t) = -A^{\mathsf{T}} \psi(t) + A^{\mathsf{T}} K^{\mathsf{T}} v(t) - Q x^*(t),$$

$$\psi(1) = 0, \tag{8}$$

$$\int_0^1 (K x^*(t) + b)^{\mathsf{T}} \, \mathrm{d}v(t) = 0. \tag{9}$$

Although u^* only lies in L^2 according to Eq. (4), it follows from Eq. (7) that there exists a member of the equivalence class associated with u^* that has bounded variation. Consequently, we assume henceforth that u^* is a particular element of the equivalence class that has bounded variation. In this paper, we establish Lipschitz continuity properties for the solution and the associated multipliers.

Theorem 1. *The multiplier* v^* *is Lipschitz continuous on* $[0,1)$, *while the equivalence classes associated with* $\dot{x}^*$, u^*, *and* $\dot{\psi}^*$ *contain functions that are Lipschitz continuous on* $[0,1]$.

This theorem is proved in [11] by applying a general homotopy approach involving “compatible data”. Here we present another proof of this result using a discrete approximation approach. Our proof goes along the following lines. We first show that the solution of a discretized problem converges in an appropriate norm to the solution of the original problem. After showing the solution of the discrete problem is Lipschitz continuous in discrete time, with a Lipschitz constant independent of the mesh size, Theorem 1 is obtained through a compactness argument. With the help of additional results from [9], the analysis in this paper can be extended to handle control constraints, time-dependent matrices, and a weaker coercivity condition for the objective function.

In the case $k = 1$, corresponding to one state constraint, a stronger result is obtained in [10]. Namely, under the independence condition, there exists an optimal control which is a piecewise analytic function of time. This follows from the more general result, proved in [10], that in this case there are finitely many points where the state constraint changes from active to nonactive. The generalization of this result to problems with multiple state constraints is an open problem. For recent works on regularity of solutions in optimal control, see [2–4, 13, 14].

Throughout this paper, c is a generic constant which is independent of time t and mesh size h, $\mathrm{Var}(f)$ is the total variation of f on the interval $[0,1]$, $\mathrm{Var}(f)|_\alpha^\beta$ is the total variation of f on the interval $[\alpha, \beta]$, $B_r(x)$ is the closed ball centered at x with radius r, and $\mathrm{O}(h)$ is an expression that is bounded by ch.

2. The discrete problem

Let N be a natural number and define $h = 1/N$. The Euler approximation to problem (1) associated with the grid $t_i = ih$, $i = 0,1,\ldots,N$, is the following:

$$\text{minimize} \quad \frac{1}{2} \sum_{i=0}^{N-1} x_i^{\mathsf{T}} Q x_i + u_i^{\mathsf{T}} R u_i \tag{10}$$

$$\text{subject to} \quad x_i' = A x_i + B u_i, \quad x_0 = a,$$

$$K x_i + b \leqslant 0, \quad i = 0,1,\ldots,N-1, \tag{11}$$

where x_i' is the first divided difference, $x_i' = (x_{i+1} - x_i)/h$.

Utilizing the independence condition, we show below that the problem (10) is feasible for h sufficiently small. Since R is positive definite and Q is semidefinite, Eq. (10) has a unique solution (x^h, u^h). Moreover, by the Karush–Kuhn–Tucker conditions, there exist Lagrange multipliers p^h and μ^h such that the following optimality conditions hold:

$$p'_{i-} = -Qx_i - A^{\mathsf{T}} p_i - K^{\mathsf{T}}\mu_i, \quad p_{N-1} = 0, \tag{12}$$

$$u_i = -R^{-1}B^{\mathsf{T}} p_i, \tag{13}$$

$$Kx_i + b \in \mathcal{N}_{\mathbb{R}^k_+}(\mu_i),$$

for $i = 0, 1, \ldots, N-1$, where $\mathcal{N}_{\mathbb{R}^k_+}(\mu)$ denotes the normal cone to the positive orthant $\mathbb{R}^k_+$ at the point μ and p'_{i-} is the backward divided difference, $p'_{i-} = (p_i - p_{i-1})/h$.

The KKT multipliers p^h and μ^h are connected to the multipliers ψ and v of the continuous problem through the following transformation:

$$v_i = -h\sum_{l=i}^{N} \mu_l, \quad \text{where } \mu_N = 0, \quad \text{and}$$

$$\psi_i = p_i + K^{\mathsf{T}} v_{i+1}. \tag{14}$$

We note that $\mu_i = v'_i$. In terms of these transformed multipliers, the first-order optimality conditions can be stated in the following way: There exist multipliers ψ^h and v^h associated with the solution (x^h, u^h) of Eq. (10) such that (x^h, u^h, ψ^h, v^h) satisfies the following conditions:

$$x'_i = Ax_i + Bu_i, \quad x_0 = a, \tag{15}$$

$$\psi'_{i-} = -A^{\mathsf{T}}\psi_i - Qx_i + A^{\mathsf{T}}K^{\mathsf{T}}v_{i+1}, \quad \psi_{N-1} = 0, \tag{16}$$

$$Ru_i + B^{\mathsf{T}}\psi_i - B^{\mathsf{T}}K^{\mathsf{T}}v_{i+1} = 0, \tag{17}$$

$$Kx_i + b \in \mathcal{N}_{\mathbb{R}^k_+}(v'_i). \tag{18}$$

Our first lemma compares a continuous state trajectory to its discrete counterpart.

Lemma 1. *Let v be a function of bounded variations and let $x = y$ be the solution of the state equation* (2) *associated with $u = v$. Let v^h be a discrete-time control and $x = y^h$ be the solution of* Eq. (11) *associated with $u = v^h$. Then we have*

$$\sup_{0 \leqslant j \leqslant N} |y(t_j) - y^h_j| \leqslant ch\left(\sum_{i=0}^{N-1} |v^h_i - v(t_i)| + \mathrm{Var}(y) + \mathrm{Var}(v)\right). \tag{19}$$

Proof. By the state equation (2), we have

$$\begin{aligned} y(t_{i+1}) - y(t_i) &= \int_{t_i}^{t_{i+1}} \dot{y}(t)\,\mathrm{d}t \\ &= \int_{t_i}^{t_{i+1}} Ay(t) + Bv(t)\,\mathrm{d}t \\ &= hAy(t_i) + hBu(t_i) + r_i, \end{aligned}$$

where

$$r_i = \int_{t_i}^{t_{i+1}} A(y(t) - y(t_i)) + B(v(t) - v(t_i))\,\mathrm{d}t.$$

Combining this with Eq. (15) gives

$$\Delta y_{i+1} = \Delta y_i + hA\Delta y_i + hB(v(t_i) - v^h_i) + r_i,$$

$$\Delta y_0 = 0,$$

where $\Delta y_i = y(t_i) - y^h_i$. It follows that

$$\begin{aligned} |\Delta y_j| &\leqslant c\sum_{i=0}^{j-1} h|v(t_i) - v^h_i| + |r_i| \\ &\leqslant c\sum_{i=0}^{N-1}\left(h|v(t_i) - v^h_i| + \int_{t_i}^{t_{i+1}} |y(t) - y(t_i)| + |v(t) - v(t_i)|\,\mathrm{d}t\right) \\ &\leqslant ch\sum_{i=0}^{N-1}\left(|v(t_i) - v^h_i| + \mathrm{Var}(y)\Big|_{t_i}^{t_{i+1}} + \mathrm{Var}(v)\Big|_{t_i}^{t_{i+1}}\right) \\ &\leqslant ch\sum_{i=0}^{N-1} |v(t_i) - v^h_i| + ch(\mathrm{Var}(y) + \mathrm{Var}(v)). \end{aligned}$$

This completes the proof. □

As a consequence of Lemma 1, we have the existence of feasible points for (11).

Corollary 1. *There exists $\eta > 0$ and $\bar{h} > 0$ with the following property: For all $h \leqslant \bar{h}$, there exists a discrete control $\tilde{u}^h$ for which the state $x = \tilde{x}^h$ associated with $u = \tilde{u}^h$ in* Eq. (11) *satisfies*

$$K\tilde{x}^h_i + b \leqslant -\eta < 0 \quad \textit{for } i = 0, 1, \ldots, N-1.$$

Proof. By ([7], Lemma 3), there exists a control $\tilde{u} \in L^\infty$ for which the state $x = \tilde{x}$ associated with $u = \tilde{u}$

in Eq. (2) satisfies

$$K\tilde{x}(t) + b \leqslant -\eta < 0 \quad \text{for all } t \in [0,1].$$

Since the infinitely differentiable functions C^∞ are dense in L^2, $\tilde{u}$ can be chosen in C^∞. Defining $\tilde{u}_i^h = \tilde{u}(t_i)$, taking h sufficiently small, and applying Lemma 1, the proof is complete. □

The discrete analogues of the L^2, L^∞ and H^1 norms are defined in the following way. For a sequence of n-vectors $z_0, z_1, \ldots, z_{N-1}$, we have

$$\|z\|_{L^2} = \sqrt{\sum_{i=0}^{N-1} h|z_i|^2},$$

$$\|z\|_{L^\infty} = \sup_{0 \leqslant i \leqslant N-1} |z_i|,$$

$$\|z\|_{H^1} = \sqrt{\|z\|_{L^2}^2 + \|z'\|_{L^2}^2}.$$

Defining $H_a^1 = \{x \in H^1 \mid x_0 = a\}$, and $H_N^1 = \{\psi \in H^1 \mid \psi_{N-1} = 0\}$, we consider the following map from $X := H_a^1 \times L^2 \times H_N^1 \times L^2$ to $Y := L^2 \times L^2 \times L^2 \times H^1$:

$$(x, u, \psi, v) = w \mapsto -\mathcal{L}^h(w) + \mathcal{F}^h(w),$$

where

$$\mathcal{L}^h(w)_i = \begin{pmatrix} x_i' - Ax_i - Bu_i \\ \psi_{i-}' + A^\mathsf{T}\psi_i + Qx_i - A^\mathsf{T}K^\mathsf{T}v_{i+1} \\ Ru_i - B^\mathsf{T}\psi_i + B^\mathsf{T}K^\mathsf{T}v_{i+1} \\ Kx_i + b \end{pmatrix},$$

and $\mathcal{F}$ is the set-valued map of the form

$$\mathcal{F}^h(w)_i = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \mathcal{N}_{\mathbb{R}_+^k}(v_i') \end{pmatrix}.$$

Letting $w_i^* = (x_i^*, u_i^*, \psi_i^*, v_i^*)$ denote the discrete function whose ith component is the value of the corresponding continuous functions at t_i, we have

$$\mathcal{L}^h(w^*) + \delta^h \in \mathcal{F}^h(w^*), \tag{20}$$

where δ^h is the residual defined by

$$(\delta^h)_i = -\begin{pmatrix} (x_i^*)' - Ax_i^* - Bu_i^* \\ (\psi_{i-}^*)' + A^\mathsf{T}\psi_i^* + Qx_i^* - A^\mathsf{T}K^\mathsf{T}v_{i+1}^* \\ B^\mathsf{T}K^\mathsf{T}(v_{i+1}^* - v_i^*) \\ \Delta_i \end{pmatrix}, \tag{21}$$

with

$$(\Delta_i)_j = \begin{cases} 0 & \text{if } (Kx^*(t) + b)_j < 0 \\ & \quad \text{for all } t \in (t_i, t_{i+1}), \\ (Kx_i^* + b)_j & \text{otherwise,} \end{cases}$$

$j = 1, 2, \ldots, k$. Our second lemma provides an estimate for the residual.

Lemma 2. *For the residual* δ^h *defined in* Eq. (21), *we have*

$$\|\delta^h\|_Y = \mathrm{O}(\sqrt{h}).$$

Proof. Since v^* and u^* have bounded variation, Eqs. (2) and (9) imply that $\dot{x}^*$ and $\dot{\psi}^*$ have bounded variation when restricted to some set $E \subset [0,1]$ of measure one. Focusing on the first component of δ^h, observe that

$$\begin{aligned} h|(x^*)_i' - Ax_i^* - B^*u_i^*| &= \left| \int_{t_i}^{t_{i+1}} (\dot{x}^*(t) - Ax_i^* - Bu_i^*)\,\mathrm{d}t \right| \\ &\leqslant \int_{t_i}^{t_{i+1}} |\dot{x}^*(t) - \dot{x}^*(t_i)|\,\mathrm{d}t \\ &\leqslant h\,\mathrm{Var}(\dot{x}^*)|_{t_i}^{t_{i+1}}. \end{aligned} \tag{22}$$

Here $\mathrm{Var}(\dot{x}^*)$ means the variation when restricted to E. Let M denote the first integer $\geqslant \mathrm{Var}(\dot{x}^*)$. Since there are at most M choices of the index i for which

$$\mathrm{Var}(\dot{x}^*)|_{t_i}^{t_{i+1}} \geqslant 1,$$

and since $c^2 \leqslant c$ when $0 \leqslant c \leqslant 1$, we have

$$(\mathrm{Var}(\dot{x}^*)|_{t_i}^{t_{i+1}})^2 \leqslant \mathrm{Var}(\dot{x}^*)|_{t_i}^{t_{i+1}} \quad \text{when } \mathrm{Var}(\dot{x}^*)|_{t_i}^{t_{i+1}} \leqslant 1,$$

and by Eq. (22),

$$\begin{aligned} &\sum_{i=0}^{N-1} h|(x^*)_i' - Ax_i^* - B^*u_i^*|^2 \\ &\leqslant h \sum_{i=0}^{N-1} (\mathrm{Var}(\dot{x}^*)|_{t_i}^{t_{i+1}})^2 \\ &\leqslant Mh\,\mathrm{Var}(\dot{x}^*)^2 + \sum_{i=0}^{N-1} h\,\mathrm{Var}(\dot{x}^*)\Big|_{t_i}^{t_{i+1}} \\ &\leqslant Mh\,\mathrm{Var}(\dot{x}^*)^2 + h\,\mathrm{Var}(\dot{x}^*) \\ &= \mathrm{O}(h). \end{aligned} \tag{23}$$

Hence, the first component of δ^h is bounded in L^2 norm by $O(\sqrt{h})$. A similar treatment of the second and the third components of δ^h shows that they are bounded in L^2 norm by $O(\sqrt{h})$ as well.

Now, consider the last component of δ^h for which we need to obtain an estimate in H^1. Observe that if $(Kx^*(s)+b)_j=0$ for some $s\in(0,1)$, then $K_j\dot{x}^*(s^-)\geqslant 0$ and $K_j\dot{x}^*(s^+)\leqslant 0$ (otherwise the constraint $Kx^*+b\leqslant 0$ is violated). For any given $t\in E$, let $s^\pm$ denote either s^+ or s^- where we choose s^+ if $K_j\dot{x}^*(t)>0$ and s^- if $K_j\dot{x}^*(t)<0$. It follows that if t and $s\in(t_i,t_{i+1})$, then

$$|K_j\dot{x}^*(t)|\leqslant|K_j\dot{x}^*(t)-K_j\dot{x}^*(s^\pm)|$$
$$\leqslant c\,\mathrm{Var}(\dot{x}^*)|_{s^\pm}^{t}\leqslant c\,\mathrm{Var}(\dot{x}^*)|_{t_i}^{t_{i+1}}. \tag{24}$$

For any given i and j, either $(\Delta_i)_j=0$ or $(\Delta_i)_j=(Kx_i^*+b)_j$. In this latter case, there exists $s\in(t_i,t_{i+1})$ such that $(Kx^*(s)+b)_j=0$. It follows that

$$\begin{aligned}|(\Delta_i)_j|&=|(Kx_i^*+b)_j|\\&=|(Kx^*(t_i)+b)_j-(Kx^*(s)+b)_j|\\&=\left|\int_{t_i}^{s}K_j\dot{x}^*(t)\,\mathrm{d}t\right|\leqslant ch\,\mathrm{Var}(\dot{x}^*)|_{t_i}^{t_{i+1}},\end{aligned}$$

where we utilize Eq. (24) in the last inequality. Proceeding as in Eq. (23), we conclude that $\|\Delta\|_{L^2}=O(h\sqrt{h})$. Since $\|\Delta'\|_{L^2}\leqslant 2h^{-1}\|\Delta\|_{L^2}$, it follows that $\|\Delta\|_{H^1}=O(\sqrt{h})$. This completes the proof. □

Lemma 3. *There exist positive numbers σ,γ and $\bar{h}$ such that for every $h\leqslant\bar{h}$, the map*

$$y\mapsto(-\mathscr{L}^h+\mathscr{F}^h)^{-1}(y)$$

is single-valued and Lipschitz continuous from the ball $B_\sigma(\delta^h)\subset Y$ to X.

Proof. In ([9], Lemmas 10.1, 10.2, Corollary 10.3) we prove this result in a more general setting, although there it is assumed that the optimal control is Lipschitz continuous, in which case the residual is $O(h)$. The regularity is exploited in the proof when we note that the residual δ^h goes to zero as h goes to zero. In this paper, on the other hand, we show in Lemma 2 that the residual still goes to zero, as h goes to zero, when the optimal control is of bounded variation. Hence, by the analysis of [9], we obtain the Lipschitz continuity result of Lemma 3 when the optimal control is of bounded variation. To summarize the analysis of [9], we first show that Independence implies surjectivity of the gradient of the ε-active state constraints. Using a translation, we remove the parameter y from the constraints, obtaining a related linear quadratic problem with a parameter entering linearly in the cost functional. Exploiting the convexity of the cost function, we show that the solution map is a Lipschitz continuous function of the parameter in the cost function. Finally, translating back to the original problem and again utilizing the Independence condition, we obtain that both the original solution map and the Lagrange multiplier map are single-valued and Lipschitz continuous with respect to y. □

Corollary 2. *For h sufficiently small, we have*

$$\|w^*-w^h\|_X=O(\sqrt{h}), \tag{25}$$

where $w^h=(x^h,u^h,\psi^h,v^h)$.

Proof. Referring to Lemma 2, choose h small enough that $0\in B_\sigma(\delta^h)$. Since $w^h\in(-\mathscr{L}^h+\mathscr{F}^h)^{-1}(0)$ and $w^*\in(-\mathscr{L}^h+\mathscr{F}^h)^{-1}(\delta^h)$ by Eq. (25), the estimate Eq. (25) is an immediate consequence of Lemma 3. □

Lemma 4. *There exists c and $\bar{h}>0$ such that for all $h\leqslant\bar{h}$,*

$$\|u^h\|_{L^\infty}+\|(x^h)'\|_{L^\infty}+\|(\psi^h)'\|_{L^\infty}+\|v^h\|_{L^\infty}<c.$$

Proof. See ([6], Lemma 4.2) or the first part of the proof of ([9], Lemma 11.1). □

3. Lipschitz continuity

We now establish Lipschitz continuity, first in the discrete problem, and then in the continuous problem.

Lemma 5. *There exist c and $\bar{h}>0$ such that for all $0<h<\bar{h}$,*

$$\begin{aligned}&\|(u^h)'\|_{L^\infty}+\|(x^h)''\|_{L^\infty}\\&\quad+\|(\psi^h)''\|_{L^\infty}+\|(v^h)'\|_{L^\infty}<c.\end{aligned}$$

Proof. By Corollary 2, $\|x^h-x^*\|_{H^1}\to 0$ as $h\to 0$. Hence, for any given ε and for h sufficiently small, the constraints active in the discrete problem (10) are ε active in the continuous problem (1). As noted in the introduction, the independence condition also holds for ε active constraints. As a consequence, the independence condition holds, for h sufficiently small, for the solution to the discrete problem. The remainder of the proof is a repetition to the second part of the proof of Lemma 11.1 in [9], applied in our context. For completeness, we present it here.

Henceforth, we omit the superscript "h" on the discrete variables. For any given i, let K_i and b_i denote the sub-matrix of K and the sub-vector of b associated with the active state constraints at time level i. We have

$$K_i x_{i-1} + b_i \leqslant 0, \tag{26}$$

$$K_i x_i + b_i = 0, \tag{27}$$

$$K_i x_{i+1} + b_i \leqslant 0. \tag{28}$$

Subtracting Eq. (27) from the Eq. (28) gives $K_i x_i' \leqslant 0$. Substituting for x_i' from Eq. (11) and u_i from Eq. (13), we obtain

$$K_i A x_i - K_i B R^{-1} B^{\mathsf T} p_i \leqslant 0. \tag{29}$$

Similarly, subtracting Eq. (27) from Eq. (28) and substituting x_{i-}' from Eq. (11) and u_{i-1} from Eq. (13), we get

$$-K_i A x_{i-1} + K_i B R^{-1} B^{\mathsf T} p_{i-1} \leqslant 0. \tag{30}$$

Adding Eq. (30) to Eq. (29) results in

$$K_i A x_{i-}' - K_i B R^{-1} B^{\mathsf T} p_{i-}' \leqslant 0.$$

Substituting for p_{i-}' from Eq. (12) and rearranging yields

$$\begin{aligned} &K_i B R^{-1} B^{\mathsf T} K_i^{\mathsf T} \mu_i^+ \\ &\quad \leqslant -K_i A x_{i-}' - K_i B R^{-1} B^{\mathsf T} A^{\mathsf T} p_i - K_i B R^{-1} B^{\mathsf T} Q x_i, \end{aligned} \tag{31}$$

where μ_i^+ is the subvector of μ_i associated with the active state constraints at time level i (recall that the components of μ_i associated with inactive state constraints are all zero). By the independence condition for the discrete problem, the matrix $K_i B R^{-1} B^{\mathsf T} K_i^{\mathsf T}$ is a symmetric, positive-definite matrix with smallest eigenvalue $\gamma > 0$, independent of i and h. Multiplying both sides of Eq. (31) by the nonnegative vector μ_i^+, and taking into account Eq. (14), we obtain

$$\begin{aligned} &\gamma |\mu_i^+|^2 \\ &\quad \leqslant -(\mu_i^+)^{\mathsf T}(K_i A x_{i-}' + K_i B R^{-1} B^{\mathsf T} A^{\mathsf T}(\psi_i - K^{\mathsf T} v_{i+1}) \\ &\qquad + K_i B R^{-1} B^{\mathsf T} Q x_i). \end{aligned}$$

Since the right-hand side of this inequality is bounded by $c|\mu_i^+|$ by Lemma 4, it follows that $\|\mu\|_{L^\infty}$ is bounded, independent of h, for all h sufficiently small. Since $v_i' = \mu_i$, we conclude that $\|v'\|_{L^\infty}$ is bounded, independent of h. This bound combined with Eqs. (16) and (17) gives us the desired result. □

Proof of Theorem 1. First, let us consider the control. Letting v^h be the continuous piecewise linear interpolant to u^h, v^h is uniformly bounded in L^∞ by Lemma 4, while v^h is Lipschitz continuous on $[0,1]$ with a Lipschitz constant independent of h by Lemma 5. Since the sequence $\{v^h\}$ is uniformly bounded and equicontinuous, it follows from the Ascoli–Arzela theorem ([1], p. 10) that there exists a continuous function v^* such that

$$\|v^h - v^*\|_{L^\infty} \to 0 \quad \text{as } h \to 0.$$

The limit v^* is Lipschitz continuous since the v^h are uniformly Lipschitz continuous. By the triangle inequality,

$$\begin{aligned} &\int_0^1 |v^h(t) - u^*(t)|\,\mathrm{d}t \\ &\quad \leqslant \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} |v^h(t) - u_i^h|\,\mathrm{d}t \\ &\qquad + \sum_{i=0}^{N-1} h|u_i^h - u_i^*| \\ &\qquad + \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} |u^*(t_i) - u^*(t)|\,\mathrm{d}t. \end{aligned}$$

On the right-hand side, the first term is $O(h)$ by Lemma 5, the second term is $O(\sqrt{h})$ by Corollary 2 and the fact that the L^1 norm is bounded by the L^2 norm, and the third term is $O(h)$ since u^* has bounded variation:

$$\begin{aligned} &\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} |u^*(t_i) - u^*(t)|\,\mathrm{d}t \\ &\quad \leqslant \sum_{i=0}^{N-1} h \operatorname{Var}(u^*)\Big|_{t_i}^{t_{i+1}} \leqslant h \operatorname{Var}(u^*). \end{aligned}$$

Hence, u^* is the L^1 limit of v^h. Since v^h converges to both v^* and u^*, we conclude that $v^* = u^*$ almost everywhere. This shows that the equivalence class associated with u^* contains a Lipschitz continuous function. In a similar fashion, v^* is equal almost everywhere to a Lipschitz continuous function on $[0,1]$. Since v^* is right continuous on $[0,1]$, we conclude that v^* is Lipschitz continuous on $[0,1)$. The Lipschitz continuity of $\dot{x}^*$ and $\dot{\psi}^*$ can be deduced from the state equation (2) and the adjoint equation (9). □

Utilizing the regularity established in Theorem 1, it can be shown (see ([9], Lemma 5.1)) that $\|\delta^h\|_Y = \mathrm{O}(h)$. Exploiting this estimate, as in Corollary 1, yields

$$\|w^h - w^*\|_X = \mathrm{O}(h),$$

which is the same estimate for the control error obtained in [5] for a convex problem (see also ([6], Ch. 4)) and in [9] for a general nonlinear problem.

References

[1] R.A. Adams, Sobolev Spaces, Academic Press, New York, 1975.
[2] A.A. Agrachev, On regularity properties of extremal controls, J. Dyn. Control Systems 1 (1995) 319–324.
[3] F.N. Clarke, Y. Ledyaev, R.B. Vinter, Regularity properties of solutions to linear quadratic optimal control problems with state constraints, Systems Control Lett. 30 (1997) 265–272; Addendum, Systems Control Lett. 33 (1998) 139.
[4] V.V. Dikusar, A.A. Milyutin, Qualitative and Numerical Methods in the Maximum Principle, Nauka, Moscow, 1989.
[5] A.L. Dontchev, Error estimates for a discrete approximation to constrained control problems, SIAM J. Numer. Anal. 18 (1981) 500–514.
[6] A.L. Dontchev, Perturbations, Approximations and Sensitivity Analysis of Optimal Control Systems, Lecture Notes in Control and Inf. Sci., vol. 52, Springer, New York, 1983.
[7] A.L. Dontchev, W.W. Hager, Lipschitzian stability in nonlinear control and optimization, SIAM J. Control Optim. 31 (1993) 569–603.
[8] A.L. Dontchev, W.W. Hager, Lipschitzian stability for state constrained nonlinear optimal control, SIAM J. Control Optim. 36 (1998) 696–718.
[9] A.L. Dontchev, W.W. Hager, The Euler approximation in state constrained optimal control, Department of Mathematics, University of Florida, Gainesville, FL 32611, November 7, 1997 (http://www.math.ufl.edu/~hager).
[10] A.L. Dontchev, I. Kolmanovsky, On regularity of optimal control, in: R. Durier, C. Michelot (Eds.), Recent Developments in Optimization, Proc. French-German Conf. on Optimization, Lecture Notes in Economics and Mathematical Systems, vol. 429, Springer, Berlin, 1995, pp. 125–135.
[11] W.W. Hager, Lipschitz continuity for constrained processes, SIAM J. Control Optim. 17 (1979) 321–337.
[12] W.W. Hager, S.K. Mitter, Lagrange duality theory for convex control problems, SIAM J. Control Optim. 14 (1976) 843–856.
[13] I. Lasiecka, R. Triggiani, Further regularity properties in quadratic cost problems for parabolic equations with boundary control and non-smoothing final state penalization, Lecture Notes in Pure and Appl. Math., vol. 160, Dekker, New York, 1994, pp. 173–194.
[14] K. Malanowski, Regularity of solutions in stability analysis of optimization and optimal control problems, Control Cybernet. 23 (1994) 61–86.