Wang and Liu *Journal of Inequalities and Applications* 2012, **2012**:0
http://www.journalofinequalitiesandapplications.com/content/2012/231/0

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

A maximum principle for optimal control system with endpoint constraints

Weifeng Wang and Bin Liu*

*Correspondence:
binliu@mail.hust.edu.cn
School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, Hubei 430074, P.R. China

Abstract

Pontryagin's maximum principle for an optimal control system governed by an ordinary differential equation with endpoint constraints is proved under the assumption that the control domain has no linear structure. We also obtain the variational equation, adjoint equation and Hamilton system for our problem.

MSC: 65K10; 34H05; 93C15

Keywords: Pontryagin's maximum principle; optimal control; Hamilton system; transversality condition; linear structure

1 Introduction

Optimal control problems have been studied for a long time and have a lot of practical applications in the fields such as physics, biology and economics, *etc.* Hamilton systems are derived from Pontryagin's maximum principle, which is known as a necessary condition for optimality. Many results have been obtained both for finite and infinite dimensional control systems such as [1–5]. Regarding the state constraint problems, lots of results are also obtained. For example, readers can refer to [6, 7] and the references therein.

To our best knowledge, to derive the necessary conditions of Pontryagin's maximum principle type for optimal control problems, there are two main perturbation methods. When the control domain is convex, we often use the convex perturbation. When the control domain is non-convex and does not have any linear structure, we usually use the spike perturbation. Many relevant results have been obtained; see [3, 6, 8–10] and the references therein. The two methods have their advantages and disadvantages. The convex variational needs the control domain being convex, but in reality it is not always satisfied. And the spike variational needs more regularity for the coefficients and the solutions to the state equations, especially in the stochastic case.

In 2010, Lou [7] introduced a new method to study the necessary and sufficient conditions of optimal control problems in the absence of linear structure for the deterministic case. The author gave a local linearization of the optimal control problem along the optimal control, and transformed the original problem into a new relaxed control problem. Moreover, he proved the equivalence of the two problems in some sense. Being directly inspired by [7], we are also interested in applying this method to an endpoint constraints optimal control system, which is also in the absence of linear structure. Also, Pontryagin's maximum principle is obtained for our problem.

The rest of this paper is organized as follows. Section 2 begins with a general formulation of our state constraints optimal control problem and the local linearization of the problem is given. In Section 3, we give our main result and its proof. Moreover, we obtain the variational equation, adjoint equation and Hamilton system for our optimal control system.

2 Preliminaries

We consider the controlled ordinary differential equation in R^n

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)), \quad \text{in } [0, T], \\ x(0) = x_0, \end{cases} \tag{2.1}$$

with the cost functional

$$J\bigl(x_0, u(\cdot)\bigr) = \int_0^T l\bigl(t, x(t), u(t)\bigr)\, dt, \tag{2.2}$$

where $T > 0$ is a given constant, $x_0 \in R^n$ is a decision variable and $u(\cdot) \in U[0, T]$ with

$$U[0, T] = \bigl\{v : [0, T] \to V | v(\cdot) \text{ is measurable}\bigr\},$$

where V is a non-convex set in R^k.

Let $U_{\mathrm{ad}}[0, T]$ be the set of all elements $(x_0, u(\cdot)) \in S_1 \times U[0, T]$ satisfying

$$x(T) \in S_2,$$

where S_1 and S_2 are closed convex subsets of R^n. Let $S = S_1 \times S_2$.

Then the optimal control problem can be stated as follows.

Problem 2.1 Find a pair $(\bar{x}_0, \bar{u}(\cdot)) \in U_{\mathrm{ad}}[0, T]$ such that

$$J\bigl(\bar{x}_0, \bar{u}(\cdot)\bigr) = \inf_{(x_0, u(\cdot)) \in U_{\mathrm{ad}}[0,T]} J\bigl(x_0, u(\cdot)\bigr). \tag{2.3}$$

Any $(\bar{x}_0, \bar{u}(\cdot)) \in U_{\mathrm{ad}}[0, T]$ satisfying the above identity is called an optimal control, and the corresponding state $x(\cdot; \bar{x}_0, \bar{u}(\cdot)) \triangleq \bar{x}(\cdot)$ is called an optimal trajectory; $(\bar{x}(\cdot), \bar{u}(\cdot))$ is called an optimal pair.

Let the following hypotheses hold:

(H_1) The metric space (V, d) is separable, and d is the usual metric in R^k.

(H_2) Functions $f = (f^1, f^2, \ldots, f^n)^\top : [0, T] \times R^n \times V \to R^n$ and $l : [0, T] \times R^n \times V \to R$ are measurable in t, continuous in (x, u) and continuously differentiable in x, where $\top$ denotes the transpose of a matrix. Moreover, there exists a constant $L > 0$ such that

$$\begin{cases} |f(t, x_1, u) - f(t, x_2, u)| \le L|x_1 - x_2|, \\ |l(t, x_1, u) - l(t, x_2, u)| \le L|x_1 - x_2|, \\ |f(t, 0, u)| \le L, \qquad |l(t, 0, u)| \le L, \\ \quad \forall (t, x_1, x_2, u) \in [0, T] \times R^n \times R^n \times V, \end{cases} \tag{2.4}$$

where $|\cdot|$ denotes the usual Euclidean norm.

From the above conditions, it is easy to know that the state equation (2.1) has a unique solution $x(\cdot;x_0,u(\cdot))$ for any $(x_0,u(\cdot)) \in U[0,T]$.

Let $(\bar{x}_0,\bar{u}(\cdot)) \in \mathcal{U}_{\mathrm{ad}}[0,T]$ be a minimizer of $J(x_0,u(\cdot))$ over $\mathcal{U}_{\mathrm{ad}}[0,T]$, and we linearize $\mathcal{U}_{\mathrm{ad}}[0,T]$ along $(x_0,\bar{u}(\cdot))$ (for any fixed x_0) in the following manner. Define

$$M_{\mathrm{ad}}[0,T] \triangleq \bigl\{x_0 \times \bigl((1-\alpha)\delta_{\bar{u}(\cdot)} + \alpha\delta_{u(\cdot)}\bigr) \big| \alpha \in [0,1], \bigl(x_0,u(\cdot)\bigr) \in \mathcal{U}_{\mathrm{ad}}[0,T]\bigr\},$$

where $\delta_{u(\cdot)}$ denotes the Dirac measure at $u(\cdot)$ on $U[0,T]$. Let $\sigma(\cdot) = (1-\alpha)\delta_{\bar{u}(\cdot)} + \alpha\delta_{u(\cdot)}$, then it is easy to see that $(x_0,\sigma(\cdot)) \in M_{\mathrm{ad}}[0,T]$. Now, we define

$$f\bigl(t,x,\sigma(t)\bigr) \equiv \int_U f(t,x,v)\sigma(t)\,dv = (1-\alpha)f\bigl(t,x,\bar{u}(t)\bigr) + \alpha f\bigl(t,x,u(t)\bigr), \tag{2.5}$$

and similarly,

$$l\bigl(t,x,\sigma(t)\bigr) \equiv \int_U l(t,x,v)\sigma(t)\,dv = (1-\alpha)l\bigl(t,x,\bar{u}(t)\bigr) + \alpha l\bigl(t,x,u(t)\bigr).$$

Then we define $x(\cdot) = x(\cdot;x_0,\sigma(\cdot))$ as the solution of the following equation:

$$\begin{cases} \dot{x}(t) = f(t,x(t),\sigma(t)), & \text{in } [0,T], \\ x(0) = x_0, \end{cases} \tag{2.6}$$

and the corresponding cost functional is

$$J\bigl(x_0,\sigma(\cdot)\bigr) = \int_0^T l\bigl(t,x\bigl(t,\sigma(t)\bigr),\sigma(t)\bigr)\,dt.$$

We can easily find that $x(\cdot;x_0,u(\cdot))$ and $J(x_0,u(\cdot))$ coincide with $x(\cdot;x_0,\delta_{u(\cdot)})$ and $J(x_0,\delta_{u(\cdot)})$ respectively. Thus, $\mathcal{U}_{\mathrm{ad}}[0,T]$ can be viewed as a subset of $M_{\mathrm{ad}}[0,T]$ in the sense of identifying $(x_0,u(\cdot)) \in \mathcal{U}_{\mathrm{ad}}[0,T]$ and $(x_0,\delta_{u(\cdot)}) \in M_{\mathrm{ad}}[0,T]$. Because the elements of $M_{\mathrm{ad}}[0,T]$ are very simple, we need neither pose additional assumptions like that the control domain is compact as Warga [11] did nor introduce the relaxed control defined by finite-additive probability measure as Fattorini [12] did. Now, we can see that $M_{\mathrm{ad}}[0,T]$ already has a linear structure at $(\bar{x}_0,\bar{u}(\cdot))$. First, we give some lemmas to show that $(\bar{x}_0,\delta_{\bar{u}(\cdot)})$ is a minimizer of $J(x_0,\sigma(\cdot))$ over $M_{\mathrm{ad}}[0,T]$.

Lemma 2.2 (Lou [7]) *Let (H_1)-(H_2) hold. Then there exists a positive constant $C > 0$, such that for any $(x_0,\sigma(\cdot)) \in M_{\mathrm{ad}}[0,T]$ and $t_1,t_2 \in [0,T]$,*

$$\begin{cases} \|x(\cdot;x_0,\sigma(\cdot))\|_{C[0,T]} \le C, \\ |x(t_1;x_0,\sigma(\cdot)) - x(t_2;x_0,\sigma(\cdot))| \le C|t_1 - t_2|. \end{cases} \tag{2.7}$$

Lemma 2.3 (Lou [7]) *Let (H_1)-(H_2) hold and $(\bar{x}_0,\bar{u}(\cdot))$ be a minimizer of $J(x_0,u(\cdot))$ over $\mathcal{U}_{\mathrm{ad}}[0,T]$. Then $(\bar{x}_0,\delta_{\bar{u}(\cdot)})$ is a minimizer of $J(x_0,\sigma(\cdot))$ over $M_{\mathrm{ad}}[0,T]$.*

Remark 2.4 We can see that Lemma 2.3 shows the equivalence of the above two problems. And this lemma is essential for our following maximum principle.

3 Pontryagin's maximum principle

In this section, we give our main result first and then prove it. Let f_x denote the derivative of f on x, and others can be defined in the same way. $\langle\cdot,\cdot\rangle$ denotes the inner product in R^n.

Theorem 3.1 (Pontryagin's maximum principle) *We assume* (H_1)-(H_2) *hold. Let* $(\bar{x}(\cdot),\bar{u}(\cdot))$ *be a solution of the optimal control problem* (2.1). *Then there exists a nontrivial pair* $(\psi^0,p(\cdot))\in R\times C^1([0,T];R^n)$, *i.e.*, $(\psi^0,p(\cdot))\neq 0$ *such that*

$$\psi^0\leq 0, \tag{3.1}$$

$$\begin{aligned} p(t)=p(T)+\int_t^T\bigl[&f_x\bigl(s,\bar{x}(s),\bar{u}(s)\bigr)p(s)\\ &+\psi^0 l_x\bigl(s,\bar{x}(s),\bar{u}(s)\bigr)\bigr]\,ds,\quad a.e.\ t\in[0,T],\end{aligned} \tag{3.2}$$

$$\bigl\langle p(0),x_1-\bar{x}_0\bigr\rangle-\bigl\langle p(T),x_2-\bar{x}(T)\bigr\rangle\leq 0,\quad \forall(x_1,x_2)\in S, \tag{3.3}$$

$$\begin{aligned} &H\bigl(t,\bar{x}(t),\bar{u}(t),p(t),\psi^0\bigr)\\ &\quad=\max_{(x_0,u(\cdot))\in\mathcal{U}_{\mathrm{ad}}[0,T]}H\bigl(t,\bar{x}(t),u(t),p(t),\psi^0\bigr),\quad a.e.\ t\in[0,T],\end{aligned} \tag{3.4}$$

where for any $(t,x,u,p,\psi^0)\in[0,T]\times R^n\times V\times R^n\times R$,

$$H\bigl(t,x,u,p,\psi^0\bigr)=\bigl\langle f(t,x,u),p\bigr\rangle+\psi^0 l(t,x,u), \tag{3.5}$$

and we also have the following Hamilton system:

$$\begin{cases} \dot{x}(t)=H_p(t,x(t),u(t),p(t),\psi^0),\quad a.e.\ t\in[0,T],\\ \dot{p}(t)=-H_x(t,x(t),u(t),p(t),\psi^0),\quad a.e.\ t\in[0,T],\\ x(0)=x_0,\qquad p(T)=-\bar{\psi},\\ H(t,x(t),u(t),p(t),\psi^0)\\ \quad=\max_{(x_0,u(\cdot))\in\mathcal{U}_{\mathrm{ad}}[0,T]}H(t,x(t),u(t),p(t),\psi^0),\quad a.e.\ t\in[0,T],\end{cases} \tag{3.6}$$

where $\bar{\psi}$ *will be defined in the following part.*

We recall that under the conditions (H_1)-(H_2), for any $(x_0,u(\cdot))\in R^n\times\mathcal{U}[0,T]$, the state equation (2.1) admits a unique solution $x(\cdot)$ with $x(0)=x_0$. So, the cost functional is uniquely determined by $(x_0,u(\cdot))$. In the sequel, we denote the unique solution of (2.1) with $x(0)=x_0$ by $x(\cdot;x_0,u(\cdot))$. Now, let $(\bar{x}_0,\bar{u}(\cdot))\in\mathcal{U}_{\mathrm{ad}}[0,T]$ be an optimal control. We denote $\bar{x}(0)=\bar{x}_0$. Without loss of generality, we may assume that $J(\bar{x}_0,\bar{u}(\cdot))=0$; otherwise, we may consider the optimal control problem with a cost functional of $J(x_0,u(\cdot))-J(\bar{x}_0,\bar{u}(\cdot))$.

Now, we give some definitions and lemmas for Theorem 3.1.

First, we define a penalty functional, via which, for convenience, we can transform the original problem to another one called the approximate problem, which has no endpoint constraint.

Let us introduce some notations. For all $(x_1,u_1(\cdot)),(x_2,u_2(\cdot))\in R^n\times\mathcal{U}[0,T]$, we define

$$\bar{d}\bigl(\bigl(x_1,u_1(\cdot)\bigr),\bigl(x_2,u_2(\cdot)\bigr)\bigr)=|x_1-x_2|+d\bigl(u_1(\cdot),u_2(\cdot)\bigr), \tag{3.7}$$

where $d(u_1(\cdot), u_2(\cdot))$ denote the measure of $\{t \in [0,T] | u_1(t) \neq u_2(t)\}$ (it obviously is a metric).

For $\forall \varepsilon > 0$ and $\forall (x_0, u(\cdot)) \in R^n \times U[0,T]$, the penalty functional is defined as follows:

$$J_\varepsilon\big(x_0, u(\cdot)\big) = \Big\{d_S^2\big(x_0, x\big(T; x_0, u(\cdot)\big)\big) + \Big[\big(J\big(x_0, u(\cdot)\big) + \varepsilon\big)^+\Big]^2\Big\}^{\frac{1}{2}}, \tag{3.8}$$

where $(a)^+ = \max\{0, a\}$, and for any $(x_0, x_1) \in R^n \times R^n$,

$$d_S(x_0, x_1) = d\big((x_0, x_1), S\big) \triangleq \inf_{(y_0, y_1) \in S} \big\{|y_0 - x_0|^2 + |y_1 - x_1|^2\big\}^{\frac{1}{2}}. \tag{3.9}$$

Obviously, d_S is a convex function and it is Lipschitz continuous with the Lipschitz constant being 1. We define the subdifferential of the function d_S as follows:

$$\begin{aligned} \partial d_S(x_0, x_1) = \big\{(a,b) \in R^n \times R^n | d_S\big(x_0', x_1'\big) - d_S(x_0, x_1) \\ \geq \big\langle a, x_0' - x_0\big\rangle + \big\langle b, x_1' - x_1\big\rangle, \forall \big(x_0', x_1'\big) \in R^n \times R^n\big\}, \end{aligned} \tag{3.10}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in R^n. For more properties of subdifferential ∂d_S, one can see p.146 in [6].

Lemma 3.2 *Let* (H_1)-(H_2) *hold. Then there exists a constant* $C > 0$ *such that for all* $(x_0, u(\cdot)), (\hat{x}_0, \hat{u}(\cdot)) \in R^n \times U[0,T]$,

$$\begin{cases} \sup_{t \in [0,T]} |x(t; x_0, u(\cdot)) - x(t; \hat{x}_0, \hat{u}(\cdot))| \\ \quad \leq C(1 + |x_0| \vee |\hat{x}_0|)\bar{d}((x_0, u(\cdot)), (\hat{x}_0, \hat{u}(\cdot))), \\ |J(x_0, u(\cdot)) - J(\hat{x}_0, \hat{u}(\cdot))| \\ \quad \leq C(1 + |x_0| \vee |\hat{x}_0|)\bar{d}((x_0, u(\cdot)), (\hat{x}_0, \hat{u}(\cdot))). \end{cases} \tag{3.11}$$

Proof Denote $x(\cdot) = x(t; x_0, u(\cdot))$ and $\hat{x}(\cdot) = x(t; \hat{x}_0, \hat{u}(\cdot))$. From the state equation (2.1) and condition (H_2), we have

$$\begin{aligned} \big|x(t)\big| &\leq |x_0| + \int_0^t \big|f\big(s, x(s), u(s)\big)\big|\,ds \\ &\leq |x_0| + \int_0^t \big(\big|f\big(s, 0, u(s)\big)\big| + \big|f\big(s, x(s), u(s)\big) - f\big(s, 0, u(s)\big)\big|\big)\,ds \\ &\leq |x_0| + \int_0^t \big(L + L\big|x(s)\big|\big)\,ds. \end{aligned}$$

By Gronwall's inequality, it follows that

$$\big|x(t)\big| \leq C\big(1 + |x_0|\big), \quad \forall t \in [0,T].$$

Similarly,

$$\big|\hat{x}(t)\big| \leq C\big(1 + |\hat{x}_0|\big), \quad \forall t \in [0,T],$$

where the constant C is independent of controls $u(\cdot)$ and $\hat{u}(\cdot)$, and may be different at different places throughout this paper. Further, noting the definition of $\bar{d}$, we have

$$\begin{aligned}
\left|x(t)-\hat{x}(t)\right| &\leq C|x_0-\hat{x}_0| + C\int_0^t \left|x(s)-\hat{x}(s)\right| ds \\
&\quad + \int_0^t \left|f\left(s,x(s),u(s)\right) - f\left(s,x(s),\hat{u}(s)\right)\right| ds \\
&\leq C|x_0-\hat{x}_0| + C\left(1+|x_0|\vee|\hat{x}_0|\right)d\left(u(\cdot),\hat{u}(\cdot)\right) + C\int_0^t \left|x(s)-\hat{x}(s)\right| ds \\
&\leq C\left(1+|x_0|\vee|\hat{x}_0|\right)\bar{d}\left(\left(x_0,u(\cdot)\right),\left(\hat{x}_0,\hat{u}(\cdot)\right)\right) + C\int_0^t \left|x(s)-\hat{x}(s)\right| ds.
\end{aligned}$$

Thus, by Gronwall's inequality, we get

$$\begin{aligned}
\left|x(t)-\hat{x}(t)\right| &\leq C\left(1+|x_0|\vee|\hat{x}_0|\right)\bar{d}\left(\left(x_0,u(\cdot)\right),\left(\hat{x}_0,\hat{u}(\cdot)\right)\right)e^{Ct} \\
&\leq C\left(1+|x_0|\vee|\hat{x}_0|\right)\bar{d}\left(\left(x_0,u(\cdot)\right),\left(\hat{x}_0,\hat{u}(\cdot)\right)\right)e^{CT}, \quad \forall t\in[0,T].
\end{aligned}$$

Taking supremum in the above inequality, the first inequality in (3.11) is obtained. The second inequality can be proved similarly. □

By the definition of $J_\varepsilon(x_0,u(\cdot))$ and Lemma 3.2, we can easily obtain the following result.

Corollary 3.3 *The functional $J_\varepsilon(x_0,u(\cdot))$ is continuous on the space $(R^n\times U[0,T],\bar{d})$.*

Remark 3.4 By the definition of $J_\varepsilon(x_0,u(\cdot))$ (3.8) and Corollary 3.3, we can see that

$$\begin{cases}
J_\varepsilon(x_0,u(\cdot))>0, \quad \forall(x_0,u(\cdot))\in R^n\times U[0,T], \\
J_\varepsilon(\bar{x}_0,\bar{u}(\cdot))=\varepsilon\leq \inf_{R^n\times U[0,T]} J_\varepsilon(x_0,u(\cdot))+\varepsilon.
\end{cases} \tag{3.12}$$

Thus, by the Ekeland variational principle (see p.135 in [6] for details), there exists a pair $(x_0^\varepsilon,u^\varepsilon(\cdot))\in R^n\times U[0,T]$, such that

$$\bar{d}\left(\left(x_0^\varepsilon,u^\varepsilon(\cdot)\right),\left(\bar{x}_0,\bar{u}(\cdot)\right)\right)\leq\sqrt{\varepsilon}, \tag{3.13}$$

$$\begin{aligned}
&J_\varepsilon\left(x_0,u(\cdot)\right)+\sqrt{\varepsilon}\bar{d}\left(\left(x_0,u(\cdot)\right),\left(x_0^\varepsilon,u^\varepsilon(\cdot)\right)\right) \\
&\quad \geq J_\varepsilon\left(x_0^\varepsilon,u^\varepsilon(\cdot)\right), \quad \forall\left(x_0,u(\cdot)\right)\in R^n\times U[0,T].
\end{aligned} \tag{3.14}$$

The above implies that if we let $x^\varepsilon(\cdot)=x(\cdot;x_0^\varepsilon,u^\varepsilon(\cdot))$, then $(x^\varepsilon(\cdot),u^\varepsilon(\cdot))$ is an optimal pair for the problem where the state equation is (2.1) and the cost functional is $J_\varepsilon(x_0,u(\cdot))$.

Now, we derive the necessary conditions for $(x^\varepsilon(\cdot),u^\varepsilon(\cdot))$.

Lemma 3.5 *Let $(x^\varepsilon(\cdot),u^\varepsilon(\cdot))$ be an optimal pair for the problem where the state equation is* (2.1) *and the cost functional is $J_\varepsilon(x_0,u(\cdot))$, then there exists a nontrivial triple $(\bar{\psi}_\varepsilon{}^0,\bar{\varphi}_\varepsilon,\bar{\psi}_\varepsilon)\in R\times R^n\times R^n$ such that*

$$|\bar{\varphi}_\varepsilon|^2+|\bar{\psi}_\varepsilon|^2+\left|\bar{\psi}_\varepsilon{}^0\right|^2=1,$$

and

$$-\sqrt{\varepsilon}\big(|\eta| + C\big) \leq \langle \bar{\varphi}_\varepsilon, \eta\rangle + \big\langle \bar{\psi}_\varepsilon, X_\varepsilon(T)\big\rangle + \bar{\psi}_\varepsilon{}^0 Y_\varepsilon,$$

where η, $X_\varepsilon(\cdot)$, Y_ε will be defined in the following proof.

Proof Similar to Section 2, we linearize $U[0,T]$ along $u^\varepsilon(\cdot)$ and denote $\sigma^{\alpha,\varepsilon}(\cdot) = (1-\alpha)\delta_{u^\varepsilon(\cdot)} + \alpha\delta_{u(\cdot)}$, $x^{\alpha,\varepsilon}(\cdot) = x(\cdot; x_0^\varepsilon + \alpha\eta, \sigma^{\alpha,\varepsilon}(\cdot))$, where $\eta \in B_1(0) \subset R^n$ ($B_1(0)$ denotes the ball whose center is 0 and radius is 1). Recall that $x^\varepsilon(\cdot) = x(\cdot; x_0^\varepsilon, u^\varepsilon(\cdot))$, then we derive the variational equation.

From (2.5), we have

$$\begin{aligned} X^{\alpha,\varepsilon}(t) &\triangleq \frac{x^{\alpha,\varepsilon}(t) - x^\varepsilon(t)}{\alpha} \\ &= \int_0^t \bigg[\frac{f(s, x^{\alpha,\varepsilon}(s), u^\varepsilon(s)) - f(s, x^\varepsilon(s), u^\varepsilon(s))}{\alpha} \\ &\quad + f\big(s, x^{\alpha,\varepsilon}(s), u(s)\big) - f\big(s, x^{\alpha,\varepsilon}(s), u^\varepsilon(s)\big)\bigg]\,ds \\ &= \int_0^t \bigg[\int_0^1 f_x\big(s, x^\varepsilon(s) + \tau\big(x^{\alpha,\varepsilon}(s) - x^\varepsilon(s)\big), u^\varepsilon(s)\big)^\top\,d\tau X^{\alpha,\varepsilon}(s) \\ &\quad + f\big(s, x^{\alpha,\varepsilon}(s), u(s)\big) - f\big(s, x^{\alpha,\varepsilon}(s), u^\varepsilon(s)\big)\bigg]\,ds, \end{aligned}$$

where $f_x(t,x,u)$ denotes the transpose of the Jacobi matrix of f on x. By virtue of (H_2), and using the convergence of $x^{\alpha,\varepsilon}(t) \to x^\varepsilon(t)$ (see the proof of Lemma 2.2 in [7]), we can easily obtain

$$X^{\alpha,\varepsilon}(t) \to X^\varepsilon(t), \quad \text{as } \alpha \to 0, \forall t \in [0,T],$$

where $X^\varepsilon(t)$ is the solution of the following variational equation:

$$\begin{cases} \dot{X}^\varepsilon(t) = f_x(t, x^\varepsilon(t), u^\varepsilon(t))^\top X^\varepsilon(t) + f(t, x^\varepsilon(t), u(t)) - f(t, x^\varepsilon(t), u^\varepsilon(t)), \\ X^\varepsilon(0) = \eta. \end{cases} \tag{3.15}$$

From the definition of $\sigma^{\alpha,\varepsilon}(\cdot)$, by Lemma 2.3, for any fixed α, we can define

$$\tilde{d}\big(\sigma^{\alpha,\varepsilon}(\cdot), \delta_{u^\varepsilon(\cdot)}\big) = \alpha d\big(u(\cdot), u^\varepsilon(\cdot)\big),$$

and

$$\tilde{d}\big(\big(x_1, \sigma^{\alpha,\varepsilon}(\cdot)\big), \big(x_2, \delta_{u^\varepsilon(\cdot)}\big)\big) = |x_1 - x_2| + \tilde{d}\big(\sigma^{\alpha,\varepsilon}(\cdot), \delta_{u^\varepsilon(\cdot)}\big).$$

So, by virtue of (3.14) and Lemma 2.3, we have

$$J_\varepsilon\big(x_0^\varepsilon + \alpha\eta, \sigma^{\alpha,\varepsilon}(\cdot)\big) + \sqrt{\varepsilon}\tilde{d}\big(\big(x_0^\varepsilon + \alpha\eta, \sigma^{\alpha,\varepsilon}(\cdot)\big), \big(x_0^\varepsilon, \delta_{u^\varepsilon(\cdot)}\big)\big) \geq J_\varepsilon\big(x_0^\varepsilon, \delta_{u^\varepsilon(\cdot)}\big). \tag{3.16}$$

It means that when $\alpha \to 0$,

$$\begin{aligned}
-\sqrt{\varepsilon}\big(|\eta| + d\big(u(\cdot), u^\varepsilon(\cdot)\big)\big) \le \lim_{\alpha\to 0} \frac{1}{\alpha}\big(J_\varepsilon\big(x_0^\varepsilon + \alpha\eta, \sigma^{\alpha,\varepsilon}(\cdot)\big) - J_\varepsilon\big(x_0^\varepsilon, \delta_{u^\varepsilon(\cdot)}\big)\big) \\
= \frac{1}{2J_\varepsilon(x_0^\varepsilon, \delta_{u^\varepsilon(\cdot)})} \lim_{\alpha\to 0} \frac{1}{\alpha}\big\{d_S^2\big(x_0^\varepsilon + \alpha\eta, x^{\alpha,\varepsilon}(T)\big) - d_S^2\big(x_0^\varepsilon, x^\varepsilon(T)\big) \\
+ \big[\big(J\big(x_0^\varepsilon + \alpha\eta, \sigma^{\alpha,\varepsilon}(\cdot)\big) + \varepsilon\big)^+\big]^2 - \big[\big(J\big(x_0^\varepsilon, \delta_{u^\varepsilon(\cdot)}\big) + \varepsilon\big)^+\big]^2\big\}.
\end{aligned}$$

Note that the map $(x_0, x_1) \to d_S^2(x_0, x_1)$ is continuously differentiable on $R^n \times R^n$, then we get

$$\begin{aligned}
&\lim_{\alpha\to 0} \frac{1}{\alpha}\big\{d_S^2\big(x_0^\varepsilon + \alpha\eta, x^{\alpha,\varepsilon}(T)\big) - d_S^2\big(x_0^\varepsilon, x^\varepsilon(T)\big)\big\} \\
&\quad = 2d_S\big(x_0^\varepsilon, x^\varepsilon(T)\big)\big(\langle a_\varepsilon, \eta\rangle + \big\langle b_\varepsilon, X^\varepsilon(T)\big\rangle\big),
\end{aligned} \tag{3.17}$$

with $(a_\varepsilon, b_\varepsilon) \in \partial d_S(x_0^\varepsilon, x^\varepsilon(T))$ and

$$|a_\varepsilon|^2 + |b_\varepsilon|^2 = 1. \tag{3.18}$$

Similarly, we can obtain

$$\begin{aligned}
&\lim_{\alpha\to 0} \frac{1}{\alpha}\big\{\big[\big(J\big(x_0^\varepsilon + \alpha\eta, \sigma^{\alpha,\varepsilon}(\cdot)\big) + \varepsilon\big)^+\big]^2 - \big[\big(J\big(x_0^\varepsilon, \delta_{u^\varepsilon(\cdot)}\big) + \varepsilon\big)^+\big]^2\big\} \\
&\quad = 2\big(J\big(x_0^\varepsilon, \delta_{u^\varepsilon(\cdot)}\big) + \varepsilon\big)^+ Y_\varepsilon,
\end{aligned} \tag{3.19}$$

where

$$\begin{aligned}
Y_\varepsilon &= \lim_{\alpha\to 0} \frac{J((1-\alpha)\delta_{u^\varepsilon(\cdot)} + \alpha\delta_{u(\cdot)}) - J(\delta_{u^\varepsilon(\cdot)})}{\alpha} \\
&= \lim_{\alpha\to 0} \int_0^T \bigg[\frac{l(t, x^{\alpha,\varepsilon}(t), u^\varepsilon(t)) - l(t, x^\varepsilon(t), u^\varepsilon(t))}{\alpha} \\
&\qquad + l\big(t, x^{\alpha,\varepsilon}(t), u(t)\big) - l\big(t, x^{\alpha,\varepsilon}(t), u^\varepsilon(t)\big)\bigg]\,dt \\
&= \lim_{\alpha\to 0} \int_0^T \bigg[\int_0^1 \big\langle l_x\big(t, x^\varepsilon(t) + s\big(x^{\alpha,\varepsilon}(t) - x^\varepsilon(t)\big), u^\varepsilon(t)\big), X^{\alpha,\varepsilon}(t)\big\rangle\,ds \\
&\qquad + l\big(t, x^{\alpha,\varepsilon}(t), u(t)\big) - l\big(t, x^{\alpha,\varepsilon}(t), u^\varepsilon(t)\big)\bigg]\,dt \\
&= \int_0^T \big[\big\langle l_x\big(t, x^\varepsilon(t), u^\varepsilon(t)\big), X^\varepsilon(t)\big\rangle + l\big(t, x^\varepsilon(t), u(t)\big) - l\big(t, x^\varepsilon(t), u^\varepsilon(t)\big)\big]\,dt.
\end{aligned}$$

Combining (3.17) and (3.19), by sending $\alpha \to 0$, we obtain

$$-\sqrt{\varepsilon}\big(|\eta| + C\big) \le \langle\bar{\varphi}_\varepsilon, \eta\rangle + \big\langle\bar{\psi}_\varepsilon, X_\varepsilon(T)\big\rangle + \bar{\psi}_\varepsilon{}^0 Y_\varepsilon, \tag{3.20}$$

where

$$\begin{cases}
(\bar{\varphi}_\varepsilon, \bar{\psi}_\varepsilon) = \frac{d_S(x_0^\varepsilon, x^\varepsilon(T))}{J_\varepsilon(x_0^\varepsilon, \delta_{u^\varepsilon(\cdot)})}(a_\varepsilon, b_\varepsilon), \\
\bar{\psi}_\varepsilon{}^0 = \frac{(J(x_0^\varepsilon, \delta_{u^\varepsilon(\cdot)}) + \varepsilon)^+}{J_\varepsilon(x_0^\varepsilon, \delta_{u^\varepsilon(\cdot)})}.
\end{cases} \tag{3.21}$$

From (3.18), it is obvious that

$$|\bar{\varphi}_\varepsilon|^2 + |\bar{\psi}_\varepsilon|^2 + \left|\bar{\psi}_\varepsilon{}^0\right|^2 = 1. \tag{3.22}$$

□

Remark 3.6 By the definition of subdifferential of the function $d_S(\cdot)$, for any $(x_1, x_2) \in S$, we have

$$\left\langle \bar{\varphi}_\varepsilon, x_1 - x_0^\varepsilon \right\rangle + \left\langle \bar{\psi}_\varepsilon, x_2 - x^\varepsilon(T) \right\rangle \le 0. \tag{3.23}$$

In order to pass to the limit as $\varepsilon \to 0$, we give the following lemma first, which is necessary for the derivation of our maximum principle.

Lemma 3.7 *It holds that*

$$\lim_{\varepsilon \to 0}\Big[\left|x_0^\varepsilon - \bar{x}_0\right| + \sup_{t\in[0,T]}\left|X^\varepsilon(t) - X(t)\right| + |Y_\varepsilon - Y|\Big] = 0,$$

where $X(t)$ and Y satisfy the following equations:

$$\begin{cases} \dot{X}(t) = f_x(t, \bar{x}(t), \bar{u}(t))^\top X(t) + f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t)), \\ X(0) = \eta, \end{cases} \tag{3.24}$$

$$Y = \int_0^T \left[l_x\big(t, \bar{x}(t), \bar{u}(t)\big)^\top X(t) + l\big(t, \bar{x}(t), u(t)\big) - l\big(t, \bar{x}(t), \bar{u}(t)\big)\right] dt. \tag{3.25}$$

Proof By (3.13) and the definition of $\bar{d}$, it is easy to see that

$$\lim_{\varepsilon \to 0}\left|x_0^\varepsilon - \bar{x}_0\right| = 0. \tag{3.26}$$

From (3.15) and (3.24), we have

$$\begin{aligned} \left|X^\varepsilon(t) - X(t)\right| \le{}& \int_0^t \Big| f_x\big(s, x^\varepsilon(s), u^\varepsilon(s)\big)^\top X^\varepsilon(s) + f\big(s, x^\varepsilon(s), u(s)\big) - f\big(s, x^\varepsilon(s), u^\varepsilon(s)\big) \\ &- \Big(f_x\big(s, \bar{x}(s), \bar{u}(s)\big)^\top X(s) + f\big(s, \bar{x}(s), u(s)\big) - f\big(s, \bar{x}(s), \bar{u}(s)\big)\Big)\Big| \, ds \\ \le{}& \int_0^t \Big[\left|f_x\big(s, x^\varepsilon(s), u^\varepsilon(s)\big)^\top\right|\left|X^\varepsilon(s) - X(s)\right| \\ &+ \left|f_x\big(s, x^\varepsilon(s), u^\varepsilon(s)\big)^\top - f_x\big(s, \bar{x}(s), \bar{u}(s)\big)^\top\right|\left|X(s)\right| \\ &+ \left|f\big(s, x^\varepsilon(s), u(s)\big) - f\big(s, \bar{x}(s), u(s)\big)\right| \\ &+ \left|f\big(s, x^\varepsilon(s), u^\varepsilon(s)\big) - f\big(s, \bar{x}(s), \bar{u}(s)\big)\right|\Big] \, ds. \end{aligned} \tag{3.27}$$

By virtue of Lemma 3.2, then

$$\begin{aligned} &\sup_{t\in[0,T]}\left|x^\varepsilon(t) - \bar{x}(t)\right| \\ &\quad \le C\big(1 + \left|x_0^\varepsilon\right| \vee |\bar{x}_0|\big)\bar{d}\Big(\big(x_0^\varepsilon, u^\varepsilon(\cdot)\big), \big(\bar{x}_0, \bar{u}(\cdot)\big)\Big) \to 0, \quad \text{as } \varepsilon \to 0. \end{aligned} \tag{3.28}$$

Now, combining the condition (H_2) and (3.27), we deduce that

$$\begin{aligned}&\left|X^\varepsilon(t)-X(t)\right|\\&\quad\le C\bigl(1+\left|x_0^\varepsilon\right|\vee|\bar{x}_0|\bigr)\bar{d}\bigl(\bigl(x_0^\varepsilon,u^\varepsilon(\cdot)\bigr),\bigl(\bar{x}_0,\bar{u}(\cdot)\bigr)\bigr)+C\int_0^t\left|X^\varepsilon(s)-X(s)\right|ds.\end{aligned}\tag{3.29}$$

Then by Gronwall's inequality, we obtain

$$\lim_{\varepsilon\to0}\sup_{t\in[0,T]}\left|X^\varepsilon(t)-X(t)\right|=0.\tag{3.30}$$

Similarly, we can prove that

$$\lim_{\varepsilon\to0}|Y_\varepsilon-Y|=0.\tag{3.31}$$

Thus, the proof of this lemma is completed. □

Remark 3.8 Now, we can let $\varepsilon\to0$. By (3.23), it is obvious that for any $(x_1,x_2)\in S$, we have

$$\langle\bar{\varphi}_\varepsilon,x_1-\bar{x}_0\rangle+\bigl\langle\bar{\psi}_\varepsilon,x_2-\bar{x}(T)\bigr\rangle\le\bigl(\left|x_0^\varepsilon-\bar{x}_0\right|^2+\left|x^\varepsilon(T)-\bar{x}(T)\right|^2\bigr)^{\frac12}\triangleq\delta_\varepsilon\to0.\tag{3.32}$$

Hence, combining (3.20) and (3.32), for any $(x_1,x_2)\in S$, we obtain

$$\begin{aligned}&\bigl\langle\bar{\varphi}_\varepsilon,\eta-(x_1-\bar{x}_0)\bigr\rangle+\bigl\langle\bar{\psi}_\varepsilon,X(T)-\bigl(x_2-\bar{x}(T)\bigr)\bigr\rangle+\bar{\psi}_\varepsilon{}^0Y\\&\quad\ge-\sqrt{\varepsilon}\bigl(|\eta|+C\bigr)-\delta_\varepsilon-\left|X_\varepsilon(T)-X(T)\right|-|Y_\varepsilon-Y|.\end{aligned}\tag{3.33}$$

From (3.22), we can find a subsequence (still denoted by itself) such that

$$\bigl(\bar{\varphi}_\varepsilon,\bar{\psi}_\varepsilon,\bar{\psi}_\varepsilon{}^0\bigr)\to\bigl(\bar{\varphi},\bar{\psi},\bar{\psi}^0\bigr)\ne0.\tag{3.34}$$

Now, by Lemma 3.7 and sending $\varepsilon\to0$ in (3.33), for any $(x_1,x_2)\in S$, $u(\cdot)\in U[0,T]$ and $\eta\in B_1(0)$, we have

$$\bigl\langle\bar{\varphi},\eta-(x_1-\bar{x}_0)\bigr\rangle+\bigl\langle\bar{\psi},X(T)-\bigl(x_2-\bar{x}(T)\bigr)\bigr\rangle+\bar{\psi}^0Y\ge0.\tag{3.35}$$

Based on the above preparation, now we start to prove Theorem 3.1 by the duality relations.

Proof of Theorem 3.1 Let $p(t)$ solve the adjoint equation

$$\begin{cases}\dot{p}(t)=-f_x(t,\bar{x}(t),\bar{u}(t))p(t)-\psi^0l_x(t,\bar{x}(t),\bar{u}(t)),\\p(T)=-\bar{\psi},\end{cases}\tag{3.36}$$

where $\psi^0=-\bar{\psi}^0$.

So, by virtue of (3.35), we get

$$\langle\bar{\varphi}, x_1 - \bar{x}_0 - \eta\rangle + \bigl\langle p(T), X(T) - \bigl(x_2 - \bar{x}(T)\bigr)\bigr\rangle + \psi^0 Y \le 0. \tag{3.37}$$

Simultaneously, we have the following duality equality:

$$\begin{aligned}
&\bigl\langle p(T), X(T)\bigr\rangle - \bigl\langle p(0), \eta\bigr\rangle + \psi^0 Y \\
&\quad= \int_0^T \bigl\langle -f_x\bigl(t, \bar{x}(t), \bar{u}(t)\bigr)p(t) - \psi^0 l_x\bigl(t, \bar{x}(t), \bar{u}(t)\bigr), X(t)\bigr\rangle \\
&\qquad + \bigl\langle p(t), f_x\bigl(t, \bar{x}(t), \bar{u}(t)\bigr)^\top X(t) + f\bigl(t, \bar{x}(t), u(t)\bigr) - f\bigl(t, \bar{x}(t), \bar{u}(t)\bigr)\bigr\rangle\,dt \\
&\qquad + \psi^0 \int_0^T \bigl[l_x\bigl(t, \bar{x}(t), \bar{u}(t)\bigr)^\top X(t) + l\bigl(t, \bar{x}(t), u(t)\bigr) - l\bigl(t, \bar{x}(t), \bar{u}(t)\bigr)\bigr]\,dt \\
&\quad= \int_0^T \bigl\langle p(t), f\bigl(t, \bar{x}(t), u(t)\bigr) - f\bigl(t, \bar{x}(t), \bar{u}(t)\bigr)\bigr\rangle + \psi^0\bigl(l\bigl(t, \bar{x}(t), u(t)\bigr) - l\bigl(t, \bar{x}(t), \bar{u}(t)\bigr)\bigr)\,dt \\
&\quad= \int_0^T \bigl[H\bigl(t, \bar{x}(t), u(t), p(t), \psi^0\bigr) - H\bigl(t, \bar{x}(t), \bar{u}(t), p(t), \psi^0\bigr)\bigr]\,dt,
\end{aligned} \tag{3.38}$$

where H is the Hamilton function defined in (3.5). Now, setting $\eta = 0$ and $(x_1, x_2) = (\bar{x}_0, \bar{x}(T))$ in (3.37), we obtain

$$\int_0^T \bigl[H\bigl(t, \bar{x}(t), u(t), p(t), \psi^0\bigr) - H\bigl(t, \bar{x}(t), \bar{u}(t), p(t), \psi^0\bigr)\bigr]\,dt \le 0.$$

As U is separable and $(x_0, u(\cdot)) \in \mathcal{U}_{\mathrm{ad}}[0, T]$ is arbitrary, the above inequality can be written as

$$H\bigl(t, \bar{x}(t), u(t), p(t), \psi^0\bigr) - H\bigl(t, \bar{x}(t), \bar{u}(t), p(t), \psi^0\bigr) \le 0, \quad \text{a.e. } t \in [0, T]. \tag{3.39}$$

Thus,

$$\begin{aligned}
&H\bigl(t, \bar{x}(t), \bar{u}(t), p(t), \psi^0\bigr) \\
&\quad= \max_{(x_0, u(\cdot)) \in \mathcal{U}_{\mathrm{ad}}[0, T]} H\bigl(t, \bar{x}(t), u(t), p(t), \psi^0\bigr), \quad \text{a.e. } t \in [0, T].
\end{aligned} \tag{3.40}$$

Next, by taking $u(t) = \bar{u}(t)$ and $(x_1, x_2) = (\bar{x}_0, \bar{x}(T))$ in (3.37), using the duality equality (3.38), we get

$$\langle\bar{\varphi}, \eta\rangle \ge \bigl\langle p(T), X(T)\bigr\rangle + \psi^0 Y = \bigl\langle p(0), \eta\bigr\rangle, \quad \forall \eta \in B_1(0). \tag{3.41}$$

Thus, $\bar{\varphi} = p(0)$. Then taking $\eta = 0$ and $u(t) = \bar{u}(t)$ in (3.37), combining with the duality equality (3.38), we obtain the transversality condition

$$\bigl\langle p(0), x_1 - \bar{x}_0\bigr\rangle - \bigl\langle p(T), x_2 - \bar{x}(T)\bigr\rangle \le 0, \quad \forall (x_1, x_2) \in S. \tag{3.42}$$

Finally, we claim that $(\psi^0, p(\cdot)) \neq 0$. Otherwise, in particular, we have that

$$\bar{\psi} = -p(T) = 0, \qquad \bar{\varphi} = p(0) = 0. \tag{3.43}$$

Note that $\psi^0 = 0$ in this case, so this gives a contradiction to (3.22). The Hamilton system (3.6) is obvious. Then the proof of the maximum principle is completed. □

Remark 3.9 We give some important special cases of our control problem.

(i) The control problem with fixed endpoints. In this case, the constraint set is $S = \{(x_0, x_T)\}$ and the endpoint constraint becomes of the following form:

$$x(0) = x_0, \qquad x(T) = x_T.$$

(ii) The control problem with a terminal state constraint, *i.e.*,

$$x(0) = x_0, \qquad x(T) \in S_2.$$

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors considered this problem and carried out the proof. We all also read and approved the final manuscript.

Acknowledgements
We would like to thank the referees for the useful suggestions which helped to improve the previous version of this paper. This work was partially supported by NNSF of China (Grant No. 11171122).

Received: 30 May 2012 Accepted: 28 September 2012 Published: 12 October 2012

References
1. Barbu, V: Optimal of Variational Inequalities. Pitman, London (1984)
2. Barbu, V, Da Prato, G: Hamilton Jacobi Equations in Hilbert Spaces. Pitman, London (1983)
3. Barbu, V, Precupanu, T: Convexity and Optimization in Banach Spaces. Reidel, Dordrecht (1986)
4. Capuzzo-Dolcetta, I, Evans, LC: Optimal switching for ordinary differential equations. SIAM J. Control Optim. **22**, 143-161 (1984)
5. Pontryagin, LS: The maximum principle in the theory of control processes. In: Proc. 1st. Congress IFAC, Moscow (1960)
6. Li, XJ, Yong, YM: Optimal Control Theory for Infinite Dimensional Systems. Birkhäuser, Boston (1995)
7. Lou, HW: Second-order necessary/sufficient conditions for optimal control problems in the absence of linear structure. Discrete Contin. Dyn. Syst., Ser. B, **14**, 1445-1464 (2010)
8. Casas, E: Control of an elliptic problem with pointwise state constraints. SIAM J. Control Optim. **24**, 1309-1318 (1986)
9. Casas, E, Fernández, LA: Optimal control of semilinear elliptic equations with pointwise constraints on the gradient of the state. Appl. Math. Optim. **27**, 35-56 (1993)
10. Yong, JM, Zhou, XY: Stochastic Controls: Hamiltonian System and HJB Equations. Springer, New York (1999)
11. Warga, J: Optimal Control of Differential and Functional Equations. Academic Press, New York (1972)
12. Fattorini, HO: Relaxed controls in infinite dimensional systems. Int. Ser. Numer. Math. **100**, 115-128 (1991)

doi:10.1186/1029-242X-2012-231
Cite this article as: Wang and Liu: A maximum principle for optimal control system with endpoint constraints. *Journal of Inequalities and Applications* 2012 **2012**:0.