Automatica 47 (2011) 630–633

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Technical communique

A simple proof of the Pontryagin maximum principle on manifolds☆

Dong Eui Chang *

Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada

ARTICLE INFO

Article history:
Available online 18 February 2011

Keywords:
Optimal control
Pontryagin maximum principle
Tubular neighborhood
Whitney embedding

ABSTRACT

Applying the tubular neighborhood theorem, we give a simple proof of the Pontryagin maximum principle on a smooth manifold. The idea is as follows. Given a control system on a manifold M, we embed it into some $\mathbb{R}^n$ and extend the control system to $\mathbb{R}^n$. Then, we apply the Pontryagin maximum principle on $\mathbb{R}^n$ to the extended system and project the consequence to M.

1. Introduction

The classic book by Pontryagin, Boltyanskii, Gamkrelidze, and Mishchenko (1962) gives a proof of the celebrated Pontryagin Maximum Principle (PMP) for control systems on $\mathbb{R}^n$. See also Boltyanskii (1971) and Lee and Markus (1967) for another proof of the PMP on $\mathbb{R}^n$. Since many control systems are defined on manifolds (Bloch, 2003), the PMP on manifolds is as important as that on $\mathbb{R}^n$. Despite its importance, proofs of the PMP on manifolds are generally long and are not easily accessible. In general, there can be three kinds of proof of the PMP on manifolds. The first is to *mutatis mutandis* translate the proof in Pontryagin et al. (1962) into the modern differential-geometric language (Barbero-Linan and Munoz-Lecanda (2009) and Agrachev and Sachkov (2004)). Although this approach gives a good geometric insight into the principle, it has the drawback that the proof becomes long due to the repetition of many arguments in the original proof in Pontryagin et al. (1962). The second kind of proof is to adapt the proof in Pontryagin et al. (1962) to manifolds by patching up a finite number of local charts covering an optimal trajectory; see the remark on p. 357 in Jurdjevic (1997). Similarly, a drawback of the second approach is that the proof becomes long, involving coordinate transformations and repeating the proof in Pontryagin et al. (1962). Furthermore, it requires some knowledge of the proof of the PMP on $\mathbb{R}^n$. The third kind of proof is the one that we present in this paper by combining the Whitney embedding theorem and the PMP on $\mathbb{R}^n$. This approach uses the PMP on $\mathbb{R}^n$ as a lemma to prove the PMP on manifolds, avoiding repeating the arguments in the original proof in Pontryagin et al. (1962). Our proof is short, not requiring knowledge of the proof of the PMP on $\mathbb{R}^n$.

The idea in our proof is simple. Given a control system on a manifold M, we embed M into some $\mathbb{R}^n$ and construct a control system on $\mathbb{R}^n$ whose restriction to M agrees with the original system such that M becomes an invariant manifold of the extended control system on $\mathbb{R}^n$. We then reformulate the original optimal control problem with a point-to-point transfer on M into an equivalent optimal control problem with a point-to-submanifold transfer in $\mathbb{R}^n$ where the submanifold is transversal to M. We apply the PMP on $\mathbb{R}^n$ to the equivalent problem, and then project (or, restrict) the result to M, to prove the PMP on M. Our proof is pedagogically meaningful, illustrating a nice application of the tubular neighborhood theorem to control theory. For the sake of simplicity, we consider only the case of free terminal time since the fixed terminal time case follows similarly.

2. Main results

Let us first introduce the notation that will be used in the paper; refer to Chapters 1 and 2 of Abraham and Marsden (1978) for more on notation in differential manifolds theory. For a manifold M, TM and T^*M denote the tangent bundle and the cotangent bundle of M, respectively. The canonical symplectic form Ω on a cotangent bundle T^*M is a two-form that is expressed as $\Omega = dx^i \wedge dp_i$ in any local bundle charts (x^i, p_j) on T^*M. The canonical symplectic form Ω induces a bundle map $\Omega^\sharp : T^*M \to TM$ as follows: $\Omega^\sharp df = \frac{\partial f}{\partial p_i}\frac{\partial}{\partial x^i} - \frac{\partial f}{\partial x^j}\frac{\partial}{\partial p_j}$ for each smooth function $f : T^*M \to \mathbb{R}$. Let $H : M \times L \to \mathbb{R}$ be a smooth function on a product manifold $M \times L$, and (x, y) coordinates for $M \times L$. When we write $H(x; y)$

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Ulf T. Jonsson under the direction of Editor André L. Tits.

* Tel.: +1 519 888 4567x37213; fax: +1 519 746 4319.
E-mail address: dechang@uwaterloo.ca.

0005-1098/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2011.01.037

instead of $H(x, y)$, we interpret the variable y as parameter. Hence, $dH(\cdot; y) = dH(x; y) = \frac{\partial H}{\partial x^i}(x; y)dx^i$ is understood as a one-form on X parameterized by y. Let $X : M \times L \to TM$ be a map such that $X(x; y) \in T_xM$ for all $(x, y) \in M \times L$. In this case we interpret $X(x; y)$ as a vector field on M parameterized by y. However, we do not strictly follow this rule when there is no possibility of confusion. For an alternative way of treating $dH(x; y)$ and $X(x; y)$, see Definition 2.1 in Barbero-Linan and Munoz-Lecanda (2009).

2.1. Review of the Pontryagin maximum principle on $\mathbb{R}^n$

Consider a control system on $\mathbb{R}^n$:

$$\dot{x} = f(x, u), \quad x \in \mathbb{R}^n, \ u \in W \tag{1}$$

where W is a subset of some Euclidean space. We want to find a control $u(t)$, taking values in W of course, for the system (1) such that

$$\int_{t_0}^{t_1} f^0(x, u)dt \quad \text{is minimized} \tag{2}$$

with respect to piecewise continuous[1] u and to t_1, and

$$x(t_0) = x_0, \qquad x(t_1) = x_1 \tag{3}$$

where the terminal time t_1 is free. For convenience, we assume that $f : \mathbb{R}^n \times W \to \mathbb{R}^n$ and $f^0 : \mathbb{R}^n \times W \to \mathbb{R}$ are smooth.

Theorem 1 (*Theorem 1 on p.19 Pontryagin et al. (1962)*). *Suppose that $u(t), t_0 \le t \le t_1$, is a piecewise continuous optimal control and $x(t)$ is the corresponding optimal trajectory for (1) – (3). Then, there exists a nowhere-vanishing continuous curve $(p_0(t), p(t)) \in \mathbb{R} \times \mathbb{R}^n = \mathbb{R} \times T^*_{x(t)}\mathbb{R}^n$ such that:*

1. *The trajectory $(x(t), p(t))$ satisfies*

$$\begin{cases} \dot{x}^i = \dfrac{\partial H}{\partial p_i}(x, p; p_0, u(t)), \\ \dot{p}_i = -\dfrac{\partial H}{\partial x^i}(x, p; p_0, u(t)), \end{cases} \quad i = 1, \dots, n, \tag{4}$$

i.e., it is the flow of the Hamiltonian vector field

$$X_H(x, p; p_0, u(t)) = \Omega^\sharp dH(x, p; p_0, u(t)) \tag{5}$$

where Ω is the canonical symplectic form on $T^*\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$ and the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times W \to \mathbb{R}$ is given by

$$H(x, p; p_0, u) = p_0 f^0(x, u) + \langle p, f(x, u)\rangle.$$

2. *$p_0 \le 0$ and is constant in t.*
3. *$u(t) = \arg\max_{v \in W} H(x(t), p(t); p_0, v)$ for every $t \in [t_0, t_1]$.*
4. *$H(x(t), p(t); p_0, u(t)) = 0$ for every $t \in [t_0, t_1]$.*

When the two endpoints $x(t_0)$ and $x(t_1)$ are variable on some manifolds in $\mathbb{R}^n$, we replace the fixed endpoint conditions in (3) with the following conditions:

$$x(t_0) \in S_0, \quad x(t_1) \in S_1 \tag{6}$$

where S_0 and S_1 are smooth submanifolds of $\mathbb{R}^n$. The terminal time t_1 is still assumed to be free. In this case, the PMP on $\mathbb{R}^n$ is modified as follows:

Theorem 2 (*Theorem 3 on p. 50 Pontryagin et al. (1962)*). *Let $u(t), t_0 \le t \le t_1$, be a piecewise continuous optimal[2] control and $x(t)$ the corresponding trajectory for (1), (2) and (6). Then, all of the conclusions in Theorem 1 hold, and additionally the transversality conditions*

$$\langle p(t_0), T_{x(t_0)}S_0\rangle = 0, \qquad \langle p(t_1), T_{x(t_1)}S_1\rangle = 0$$

are satisfied, i.e., $\langle p(t_0), v\rangle = 0$ for all $v \in T_{x(t_0)}S_0$ and $\langle p(t_1), v\rangle = 0$ for all $v \in T_{x(t_1)}S_1$.

2.2. The Pontryagin maximum principle on manifolds

We consider the optimal control problem of finding a control $u(t)$ for the control system on an n-dimensional manifold M

$$\dot{x} = f(x, u), \quad x \in M, \ u \in W \tag{7}$$

such that

$$\int_{t_0}^{t_1} f^0(x, u)dt \quad \text{is minimized} \tag{8}$$

with respect to piecewise continuous u and to t_1, and

$$x(t_0) = x_0, \qquad x(t_1) = x_1 \tag{9}$$

where the terminal time t_1 is free. For convenience, we assume that W is a subset of a Euclidean space and that f and f^0 are smooth.

Theorem 3. *Suppose that $u(t), t_0 \le t \le t_1$ is a piecewise continuous optimal control and $x(t)$ is the corresponding trajectory for (7)–(9). Then, there exists a nowhere-vanishing continuous curve $(\lambda_0(t), \lambda(t)) \in \mathbb{R} \times T^*_{x(t)}M$ such that:*

1. *The trajectory $(x(t), \lambda(t))$ is the flow of*

$$X_H(x, \lambda; \lambda_0, u(t)) = \Omega^\sharp dH(x, \lambda; \lambda_0, u(t)) \tag{10}$$

where Ω is the canonical symplectic form on T^*M and

$$H(x, \lambda; \lambda_0, u) = \lambda_0 f^0(x, u) + \langle \lambda, f(x, u)\rangle. \tag{11}$$

2. *$\lambda_0 \le 0$ and is constant in t.*
3. *$u(t) = \arg\max_{v \in W} H(x(t), \lambda(t); \lambda_0, v)$ for every $t \in [t_0, t_1]$.*
4. *$H(x(t), \lambda(t); \lambda_0, u(t)) = 0$ for every $t \in [t_0, t_1]$.*

Proof. By the Whitney Embedding Theorem (Lee, 2002), we may assume that M is an embedded submanifold and a closed subset of $\mathbb{R}^N$ for some $N \in \mathbb{N}$ where $N > n$. Then, by the tubular neighborhood theorem, there is an open neighborhood V of M in $\mathbb{R}^N$ and a smooth retraction π_V of V onto M such that for each $x \in M$, $\pi_V^{-1}(x)$ is a submanifold of dimension $(N - n)$ in V,

$$\pi_V^{-1}(x) \cap M = \{x\}, \quad \text{and} \quad T_x(\pi_V^{-1}(x)) \oplus T_xM = \mathbb{R}^N.$$

Refer to Lee (2002) for a construction of V and π_V. Since M is a closed subset of V, by Proposition 2.26 in Lee (2002) there is a smooth function $\rho : \mathbb{R}^N \to [0, 1]$ such that the closure of the support of the function ρ is contained in V and $\rho(z) = 1$ for every $z \in M$. Define a control vector field $F : \mathbb{R}^N \times W \to \mathbb{R}^N$ by

$$F(z, u) = \begin{cases} \rho(z) f(\pi_V(z), u) & \text{if } (z, u) \in V \times W, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Note that the restriction of F to M agrees with f, and M is invariant under the flow of F. Define a function $F^0 : \mathbb{R}^N \times W \to \mathbb{R}$ by

$$F^0(z, u) = \begin{cases} \rho(z) f^0(\pi_V(z), u) & \text{if } (z, u) \in V \times W, \\ 0 & \text{otherwise,} \end{cases}$$

so as to extend the integrand f^0 in (8) to $\mathbb{R}^N \times W$.

[1] In this paper we follow the definition of a piecewise continuous function given on p. 10 in Pontryagin et al. (1962).

[2] The optimization is again with respect to piecewise continuous u and to t_1.