Statement and discussion of Pontryagin maximum principle *(Lecture 9)*

Yuri Sachkov

Program Systems Institute
Russian Academy of Sciences
Pereslavl-Zalessky, Russia
yusachkov@gmail.com

«Elements of Control Theory»

Lecture course in Program Systems Institute, Pereslavl-Zalessky

13 June 2023

Reminder: Plan of previous lecture

1. Lie derivative of differential forms
2. Liouville form and symplectic form
3. Hamiltonian vector fields

Plan of this lecture

1. Linear on fibers Hamiltonians
2. Geometric statement of PMP and discussion

Linear on fibers Hamiltonians

- We introduce a construction that works only on T^*M. Given a vector field $X \in \operatorname{Vec} M$, we define a Hamiltonian function

$$X^* \in C^\infty(T^*M),$$

which is linear on fibers T_q^*M, as follows:

$$X^*(\lambda) = \langle \lambda, X(q) \rangle, \qquad \lambda \in T^*M, \quad q = \pi(\lambda).$$

- In canonical coordinates (ξ, x) on T^*M we have:

$$X = \sum_{i=1}^{n} a_i(x) \frac{\partial}{\partial x_i}, \qquad X^* = \sum_{i=1}^{n} \xi_i a_i(x). \tag{1}$$

- This coordinate representation implies that

$$\{X^*, Y^*\} = [X, Y]^*, \qquad X, Y \in \operatorname{Vec} M,$$

i.e., Poisson brackets of Hamiltonians linear on fibers in T^*M contain usual Lie brackets of vector fields on M.

- The Hamiltonian vector field $\overrightarrow{X^*} \in \mathrm{Vec}(T^*M)$ corresponding to the Hamiltonian function X^* is called the *Hamiltonian lift* of the vector field $X \in \mathrm{Vec}\, M$.
- It is easy to see from the coordinate representation (1) that

$$\pi_* \overrightarrow{X^*} = X.$$

- Now we pass to nonautonomous vector fields. Let X_t be a nonautonomous vector field and

$$P_{\tau,t} = \overrightarrow{\exp} \int_{\tau}^{t} X_\theta \, d\theta$$

the corresponding flow on M.
- The flow $P = P_{\tau,t}$ acts on M:

$$P \, : \, M \to M, \qquad P \, : \, q_0 \mapsto q_1 ,$$

its differential pushes tangent vectors forward:

$$P_* \, : \, T_{q_0} M \to T_{q_1} M,$$

and the dual mapping P^* pulls covectors back:

$$P^* \, : \, T^*_{q_1} M \to T^*_{q_0} M.$$

- Thus we have a flow on covectors (i.e., on points of the cotangent bundle):

$$P^*_{\tau,t} \, : \, T^* M \to T^* M.$$

- Let V_t be the nonautonomous vector field on T^*M that generates the flow $P^*_{\tau,t}$:

$$V_t = \left.\frac{d}{d\,\varepsilon}\right|_{\varepsilon=0} P^*_{t,t+\varepsilon}.$$

- Then

$$\frac{d}{d\,t} P^*_{\tau,t} = \left.\frac{d}{d\,\varepsilon}\right|_{\varepsilon=0} P^*_{\tau,t+\varepsilon} = \left.\frac{d}{d\,\varepsilon}\right|_{\varepsilon=0} P^*_{t,t+\varepsilon} \circ P^*_{\tau,t} = V_t \circ P^*_{\tau,t},$$

so the flow $P^*_{\tau,t}$ is a solution to the Cauchy problem

$$\frac{d}{d\,t} P^*_{\tau,t} = V_t \circ P^*_{\tau,t}, \qquad P^*_{\tau,\tau} = \mathrm{Id},$$

i.e., it is the left chronological exponential:

$$P^*_{\tau,t} = \overleftarrow{\exp} \int_\tau^t V_\theta \, d\theta.$$

- It turns out that the nonautonomous field V_t is simply related with the Hamiltonian vector field corresponding to the Hamiltonian X_t^*:

$$V_t = -\overrightarrow{X_t^*} \,. \qquad (2)$$

- Indeed, the flow $P^*_{\tau,t}$ preserves the tautological form s, thus

$$L_{V_t} s = 0.$$

- By Cartan's formula,

$$i_{V_t}\sigma = -d\langle s, V_t\rangle,$$

i.e., the field V_t is Hamiltonian:

$$V_t = \overrightarrow{\langle s, V_t\rangle} \,.$$

- But $\pi_* V_t = -X_t$, consequently,

$$\langle s, V_t\rangle = -X_t^*,$$

and equality (2) follows.

- Taking into account the relation between the left and right chronological exponentials, we obtain

$$P^*_{\tau,t} = \overleftarrow{\exp} \int_\tau^t -\overrightarrow{X^*_\theta}\, d\theta = \overrightarrow{\exp} \int_t^\tau \overrightarrow{X^*_\theta}\, d\theta.$$

- We proved the following statement.

Proposition 1

Let X_t be a complete nonautonomous vector field on M. Then

$$\left(\overrightarrow{\exp} \int_\tau^t X_\theta \, d\theta\right)^* = \overrightarrow{\exp} \int_t^\tau \overrightarrow{X^*_\theta}\, d\theta.$$

- In particular, for autonomous vector fields $X \in \operatorname{Vec} M$,

$$\left(e^{tX}\right)^* = e^{-t\overrightarrow{X^*}}.$$

Pontryagin Maximum Principle

Geometric statement of PMP and discussion

- Consider an optimal control problem for a control system

$$\dot{q} = f_u(q), \qquad q \in M, \quad u \in U \subset \mathbb{R}^m, \tag{3}$$

with the initial condition

$$q(0) = q_0. \tag{4}$$

- Define the following family of Hamiltonians:

$$h_u(\lambda) = \langle \lambda, f_u(q) \rangle, \qquad \lambda \in T_q^*M,\ q \in M,\ u \in U.$$

- In terms of the previous slides,

$$h_u(\lambda) = f_u^*(\lambda).$$

- Fix an arbitrary instant $t_1 > 0$.

- In Lecture 2 we reduced the optimal control problem to the study of boundary of attainable sets.
- Now we give a *necessary optimality condition* in this geometric setting.

Theorem 1 (PMP)

Let $\widetilde{u}(t)$, $t \in [0, t_1]$, *be an admissible control and* $\widetilde{q}(t) = q_{\widetilde{u}}(t)$ *the corresponding solution of Cauchy problem* (3), (4). *If* $\widetilde{q}(t_1) \in \partial \mathcal{A}_{q_0}(t_1)$, *then there exists a Lipschitzian curve in the cotangent bundle*

$$\lambda_t \in T^*_{\widetilde{q}(t)} M, \qquad 0 \leq t \leq t_1,$$

such that

$$\lambda_t \neq 0, \tag{5}$$

$$\dot{\lambda}_t = \vec{h}_{\widetilde{u}(t)}(\lambda_t), \tag{6}$$

$$h_{\widetilde{u}(t)}(\lambda_t) = \max_{u \in U} h_u(\lambda_t) \tag{7}$$

for almost all $t \in [0, t_1]$.

- If $u(t)$ is an admissible control and λ_t a Lipschitzian curve in T^*M such that conditions (5)–(7) hold, then the pair $(u(t), \lambda_t)$ is said to satisfy PMP
- In this case the curve λ_t is called an *extremal*, and its projection $\widetilde{q}(t) = \pi(\lambda_t)$ is called an *extremal trajectory*.

Remark 1

If a pair $(\widetilde{u}(t), \lambda_t)$ satisfies PMP, then

$$h_{\widetilde{u}(t)}(\lambda_t) = \text{const}, \qquad t \in [0, t_1]. \tag{8}$$

Indeed, since the admissible control $\widetilde{u}(t)$ is bounded, we can take maximum in (7) over the compact $\overline{\{\widetilde{u}(t) \mid t \in [0, t_1]\}} = \widetilde{U}$.
Further, the function $\varphi(\lambda) = \max_{u \in \widetilde{U}} h_u(\lambda)$ is Lipschitzian w.r.t. $\lambda \in T^*M$. We show that this function has zero derivative.

For optimal control $\widetilde{u}(t)$,

$$\varphi(\lambda_t) \geq h_{\widetilde{u}(\tau)}(\lambda_t), \qquad \varphi(\lambda_\tau) = h_{\widetilde{u}(\tau)}(\lambda_\tau),$$

thus

$$\frac{\varphi(\lambda_t) - \varphi(\lambda_\tau)}{t - \tau} \geq \frac{h_{\widetilde{u}(\tau)}(\lambda_t) - h_{\widetilde{u}(\tau)}(\lambda_\tau)}{t - \tau}, \qquad t > \tau.$$

Consequently,

$$\left.\frac{d}{d\,t}\right|_{t=\tau} \varphi(\lambda_t) \geq \{h_{\widetilde{u}(\tau)}, h_{\widetilde{u}(\tau)}\} = 0$$

if τ is a differentiability point of $\varphi(\lambda_t)$. Similarly,

$$\frac{\varphi(\lambda_t) - \varphi(\lambda_\tau)}{t - \tau} \leq \frac{h_{\widetilde{u}(\tau)}(\lambda_t) - h_{\widetilde{u}(\tau)}(\lambda_\tau)}{t - \tau}, \qquad t < \tau,$$

thus $\left.\frac{d}{d\,t}\right|_{t=\tau} \varphi(\lambda_t) \leq 0$. So

$$\frac{d}{d\,t}\varphi(\lambda_t) = 0,$$

and identity (8) follows.

- The Hamiltonian system of PMP

$$\dot{\lambda}_t = \vec{h}_{u(t)}(\lambda_t) \tag{9}$$

is an extension of the initial control system (3) to the cotangent bundle.

- Indeed, in canonical coordinates $\lambda = (\xi, x) \in T^*M$, the Hamiltonian system yields

$$\dot{x} = \frac{\partial h_{u(t)}}{\partial \xi} = f_{u(t)}(x).$$

- That is, solutions λ_t to (9) are Hamiltonian lifts of solutions $q(t)$ to (3):

$$\pi(\lambda_t) = q_u(t).$$

- Before proving Pontryagin Maximum Principle, we discuss its statement.

- First we give a heuristic explanation of the way the covector curve λ_t appears naturally in the study of trajectories coming to boundary of the attainable set.
- Let
$$q_1 = \widetilde{q}(t_1) \in \partial \mathcal{A}_{q_0}(t_1). \tag{10}$$
- The idea is to take a normal covector to the attainable set $\mathcal{A}_{q_0}(t_1)$ near q_1, more precisely — a normal covector to a kind of a convex tangent cone to $\mathcal{A}_{q_0}(t_1)$ at q_1.
- By virtue of inclusion (10), this convex cone is proper.
- Thus it has a hyperplane of support, i.e., a linear hyperplane in $T_{q_1}M$ bounding a half-space that contains the cone.

- Further, the hyperplane of support is a kernel of a normal covector $\lambda_{t_1} \in T^*_{q_1} M$, $\lambda_{t_1} \neq 0$, see fig. below:

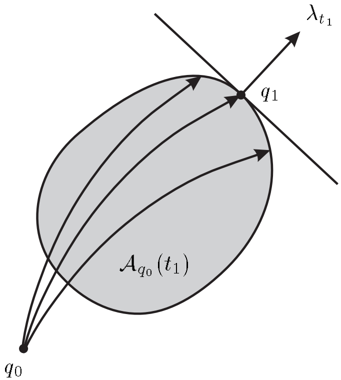

Figure: Hyperplane of support and normal covector to attainable set $\mathcal{A}_{q_0}(t_1)$ at the point q_1

- The covector λ_{t_1} is an analog of Lagrange multipliers.

- In order to construct the whole curve λ_t, $t \in [0, t_1]$, consider the flow generated by the control $\widetilde{u}(\cdot)$:

$$P_{t,t_1} = \overrightarrow{\exp} \int_t^{t_1} f_{\widetilde{u}(\tau)}\, d\tau, \qquad t \in [0, t_1].$$

- It is easy to see that

$$P_{t,t_1}(\mathcal{A}_{q_0}(t)) \subset \mathcal{A}_{q_0}(t_1), \qquad t \in [0, t_1].$$

- Indeed, if a point $q \in \mathcal{A}_{q_0}(t)$ is reachable from q_0 by a control $u(\tau)$, $\tau \in [0, t]$, then the point $P_{t,t_1}(q)$ is reachable from q_0 by the control

$$v(\tau) = \begin{cases} u(\tau), & \tau \in [0, t], \\ \widetilde{u}(\tau), & \tau \in [t, t_1]. \end{cases}$$

- Further, the diffeomorphism $P_{t,t_1} : M \to M$ satisfies the condition

$$P_{t,t_1}(\widetilde{q}(t)) = \widetilde{q}(t_1) = q_1, \qquad t \in [0, t_1].$$

- Thus if $\widetilde{q}(t) \in \operatorname{int} \mathcal{A}_{q_0}(t)$, then $q_1 \in \operatorname{int} \mathcal{A}_{q_0}(t_1)$.
- By contradiction, inclusion (10) implies that

$$\widetilde{q}(t) \in \partial \mathcal{A}_{q_0}(t), \qquad t \in [0, t_1].$$

- The tangent cone to $\mathcal{A}_{q_0}(t)$ at the point $\widetilde{q}(t) = P_{t_1,t}(q_1)$ has the normal covector $\lambda_t = P^*_{t,t_1}(\lambda_{t_1})$.
- By Proposition 1, the curve λ_t, $t \in [0, t_1]$, is a trajectory of the Hamiltonian vector field $\vec{h}_{\widetilde{u}(t)}$, i.e., of the Hamiltonian system of PMP.

- One can easily get the maximality condition of PMP as well.
- The tangent cone to $\mathcal{A}_{q_0}(t_1)$ at q_1 should contain the infinitesimal attainable set from the point q_1:

$$f_U(q_1) - f_{\widetilde{u}(t_1)}(q_1),$$

i.e., the set of vectors obtained by variations of the control $\widetilde{u}$ near t_1.
- Thus the covector λ_{t_1} should determine a hyperplane of support to this set:

$$\langle \lambda_{t_1}, f_u - f_{\widetilde{u}(t_1)} \rangle \leq 0, \qquad u \in U.$$

- In other words,

$$h_u(\lambda_{t_1}) = \langle \lambda_{t_1}, f_u \rangle \leq \langle \lambda_{t_1}, f_{\widetilde{u}(t_1)} \rangle = h_{\widetilde{u}(t_1)}(\lambda_{t_1}), \qquad u \in U.$$

- Translating the covector λ_{t_1} by the flow P^*_{t,t_1}, we arrive at the maximality condition of PMP:

$$h_u(\lambda_t) \leq h_{\widetilde{u}(t)}(\lambda_t), \qquad u \in U, \quad t \in [0, t_1].$$

- The following statement shows the power of PMP.

Proposition 2

Assume that the maximized Hamiltonian of PMP

$$H(\lambda) = \max_{u \in U} h_u(\lambda), \qquad \lambda \in T^*M,$$

is defined and C^2-smooth on $T^*M \setminus \{\lambda = 0\}$.
If a pair $(\widetilde{u}(t), \lambda_t)$, $t \in [0, t_1]$, satisfies PMP, then

$$\dot{\lambda}_t = \vec{H}(\lambda_t), \qquad t \in [0, t_1]. \tag{11}$$

Conversely, if a Lipschitzian curve $\lambda_t \neq 0$ is a solution to the Hamiltonian system (11), then one can choose an admissible control $\widetilde{u}(t)$, $t \in [0, t_1]$, such that the pair $(\widetilde{u}(t), \lambda_t)$ satisfies PMP.

- That is, in the favorable case when the maximized Hamiltonian H is C^2-smooth, PMP reduces the problem to the study of solutions to just one Hamiltonian system (11).

- From the point of view of dimension, this reduction is the best one we can expect.
- Indeed, for a full-dimensional attainable set ($\dim \mathcal{A}_{q_0}(t_1) = n$) we have $\dim \partial \mathcal{A}_{q_0}(t_1) = n - 1$, i.e., we need an $(n-1)$-parameter family of curves to describe the boundary $\partial \mathcal{A}_{q_0}(t_1)$.
- On the other hand, the family of solutions to Hamiltonian system (11) with the initial condition $\pi(\lambda_0) = q_0$ is n-dimensional.
- Taking into account that the Hamiltonian H is homogeneous:

$$H(c\lambda) = cH(\lambda), \qquad c > 0,$$

thus

$$e^{t\vec{H}}(c\lambda_0) = ce^{t\vec{H}}(\lambda_0), \qquad \pi \circ e^{t\vec{H}}(c\lambda_0) = \pi \circ e^{t\vec{H}}(\lambda_0),$$

we obtain the required $(n-1)$-dimensional family of curves.

- Now we prove Proposition 2.

Proof.

- We show that if an admissible control $\widetilde{u}(t)$ satisfies the maximality condition (7), then

$$\vec{h}_{\widetilde{u}(t)}(\lambda_t) = \vec{H}(\lambda_t), \qquad t \in [0, t_1]. \tag{12}$$

- By definition of the maximized Hamiltonian H,

$$H(\lambda) - h_{\widetilde{u}(t)}(\lambda) \geq 0 \qquad \lambda \in T^*M, \quad t \in [0, t_1].$$

- On the other hand, by the maximality condition of PMP (7), along the extremal λ_t this inequality turns into equality:

$$H(\lambda_t) - h_{\widetilde{u}(t)}(\lambda_t) = 0, \qquad t \in [0, t_1].$$

- That is why

$$d_{\lambda_t} H = d_{\lambda_t} h_{\widetilde{u}(t)}, \qquad t \in [0, t_1].$$

- But a Hamiltonian vector field is obtained from differential of the Hamiltonian by a standard linear transformation, thus equality (12) follows.

- Conversely, let $\lambda_t \neq 0$ be a trajectory of the Hamiltonian system $\dot{\lambda}_t = \vec{H}(\lambda_t)$.
- In the same way as in the proof of Filippov's theorem, one can choose an admissible control $\widetilde{u}(t)$ that realizes maximum along λ_t:

$$H(\lambda_t) = h_{\widetilde{u}(t)}(\lambda_t) = \max_{u \in U} h_u(\lambda_t).$$

- As we have shown above, then there holds equality (12). So the pair $(\widetilde{u}(t), \lambda_t)$ satisfies PMP.

$\square$