Dynamic Programming & Optimal Control

Lecture 10
Pontryagin's Minimum Principle

Fall 2023

Prof. Raffaello D'Andrea

ETH Zurich

Learning Objectives

Topic: Pontryagin's Minimum Principle

Objectives

- You know the *Pontryagin's Minimum Principle*.
- You know the *connection* between the HJB equation and the Minimum Principle.
- You know the *properties and limitations* of the Minimum Principle.

Outline

Preliminaries

Lemma 10.1:

Let $F(t, \mathrm{x}, \mathrm{u})$ be a continuously differentiable function of $t \in \mathbb{R}$, $\mathrm{x} \in \mathbb{R}^n$, $\mathrm{u} \in \mathbb{R}^m$ and let $\mathcal{U} \subseteq \mathbb{R}^m$ be the constraint set for u. Furthermore, assume that $\mu^*(t, \mathrm{x}) := \arg\min_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u})$ exists and is continuously differentiable. Then, for all t and x:

$$\frac{\partial \left(\min\limits_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u}) \right)}{\partial t} = \left. \frac{\partial F(t, \mathrm{x}, \mathrm{u})}{\partial t} \right|_{\mathrm{u} = \mu^*(t, \mathrm{x})},$$

$$\frac{\partial \left(\min\limits_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u}) \right)}{\partial \mathrm{x}} = \left. \frac{\partial F(t, \mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{\mathrm{u} = \mu^*(t, \mathrm{x})}.$$

Proof of Lemma 10.1

We prove this for $\mathcal{U} = \mathbb{R}^m$. Let $G(t, \mathrm{x}) := \min_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u}) = F(t, \mathrm{x}, \mu^*(t, \mathrm{x}))$, then:

$$\begin{aligned}
\frac{\partial G(t, \mathrm{x})}{\partial t} &= \frac{\partial F(t, \mathrm{x}, \mu^*(t, \mathrm{x}))}{\partial t} \\
&= \left.\frac{\partial F(t, \mathrm{x}, \mathrm{u})}{\partial t}\right|_{\mathrm{u}=\mu^*(t,\mathrm{x})} + \underbrace{\left.\frac{\partial F(t, \mathrm{x}, \mathrm{u})}{\partial \mathrm{u}}\right|_{\mathrm{u}=\mu^*(t,\mathrm{x})}}_{=0, \text{ since } \mu^*(t,\mathrm{x}) \text{ minimizes } F(t,\mathrm{x},\mathrm{u})} \frac{\partial \mu^*(t, \mathrm{x})}{\partial t}.
\end{aligned}$$

Similarly we obtain the result for the partial derivative with respect to x.

Example (1/2)

Let $F(t, \mathrm{x}, \mathrm{u}) := (1+t)\mathrm{u}^2 + \mathrm{ux} + 1, t \geq 0$. Then:

$$\min_{\mathrm{u}\in\mathbb{R}} F(t, \mathrm{x}, \mathrm{u})$$
$$\Rightarrow \begin{cases} \mu^*(t, \mathrm{x}) = -\frac{\mathrm{x}}{2(1+t)}, \\ \min\limits_{\mathrm{u}\in\mathbb{R}} F(t, \mathrm{x}, \mathrm{u}) = \frac{(1+t)\mathrm{x}^2}{4(1+t)^2} - \frac{\mathrm{x}^2}{2(1+t)} + 1 = -\frac{\mathrm{x}^2}{4(1+t)} + 1. \end{cases}$$

Note that

$$\frac{\partial \left(\min\limits_{\mathrm{u}\in\mathbb{R}} F(t, \mathrm{x}, \mathrm{u}) \right)}{\partial t} = \frac{\partial F(t, \mathrm{x}, \mu^*(t, \mathrm{x}))}{\partial t} = \frac{\mathrm{x}^2}{4(1+t)^2}$$

and, in agreement with Lemma 10.1,

$$\frac{\partial \left(\min\limits_{\mathrm{u}\in\mathbb{R}} F(t, \mathrm{x}, \mathrm{u}) \right)}{\partial t} = \left. \frac{\partial F(t, \mathrm{x}, \mathrm{u})}{\partial t} \right|_{\mathrm{u}=\mu^*(t,\mathrm{x})} = \left. \mathrm{u}^2 \right|_{\mathrm{u}=-\frac{\mathrm{x}}{2(1+t)}} = \frac{\mathrm{x}^2}{4(1+t)^2}.$$

Example (2/2)

Let $F(t, \mathrm{x}, \mathrm{u}) := (1+t)\mathrm{u}^2 + \mathrm{ux} + 1, t \geq 0$. Then:

$$\min_{\mathrm{u}\in\mathbb{R}} F(t,\mathrm{x},\mathrm{u})$$
$$\Rightarrow \begin{cases} \mu^*(t,\mathrm{x}) = -\frac{\mathrm{x}}{2(1+t)}, \\ \min\limits_{\mathrm{u}\in\mathbb{R}} F(t,\mathrm{x},\mathrm{u}) = \frac{(1+t)\mathrm{x}^2}{4(1+t)^2} - \frac{\mathrm{x}^2}{2(1+t)} + 1 = -\frac{\mathrm{x}^2}{4(1+t)} + 1. \end{cases}$$

Note that

$$\frac{\partial \left(\min\limits_{\mathrm{u}\in\mathbb{R}} F(t,\mathrm{x},\mathrm{u}) \right)}{\partial \mathrm{x}} = \frac{\partial F(t,\mathrm{x},\mu^*(t,\mathrm{x}))}{\partial \mathrm{x}} = -\frac{\mathrm{x}}{2(1+t)}$$

and, in agreement with Lemma 10.1,

$$\frac{\partial \left(\min\limits_{\mathrm{u}\in\mathbb{R}} F(t,\mathrm{x},\mathrm{u}) \right)}{\partial \mathrm{x}} = \left. \frac{\partial F(t,\mathrm{x},\mathrm{u})}{\partial \mathrm{x}} \right|_{\mathrm{u}=\mu^*(t,\mathrm{x})} = \mathrm{u}\big|_{\mathrm{u}=-\frac{\mathrm{x}}{2(1+t)}} = -\frac{\mathrm{x}}{2(1+t)}.$$

Outline

Problem Setup (1/2)

Dynamics

$$\dot{x}(t) = f(x(t), u(t)), \qquad 0 \leq t \leq T,$$

where:

- time $t \in \mathbb{R}_{\geq 0}$ and T is the terminal time;
- state $x(t) \in \mathcal{S} := \mathbb{R}^n, \ \forall t \in [0, T]$;
- control input $u(t) \in \mathcal{U} \subseteq \mathbb{R}^m, \ \forall t \in [0, T]$, $\mathcal{U}$ is the input space;
- $f(\cdot, \cdot)$ is the function capturing the system evolution.

Problem Setup (2/2)

Cost

We consider the scalar-valued cost function

$$h(x(T)) + \int_0^T g(x(\tau), u(\tau))\mathrm{d}\tau.$$

Objective

Given an initial condition $x(0) = \mathrm{x} \in \mathcal{S}$, construct an optimal control trajectory $u(t)$ such that the cost is minimized.

The Minimum Principle

One could solve the problem using the HJB equation, which gives an optimal policy $\mu^*(\cdot,\cdot)$. The optimal control trajectory $u(t)$ can then be inferred from the policy: $u(t) = \mu^*(t, x(t))$, where $x(t)$ is the solution to $\dot{x}(t) = f(x(t), \mu^*(t, x(t)))$, $x(0) = \mathrm{x}$.

However, solving the HJB equation is in general very difficult.

Instead, one can study the necessary conditions for optimality, introduced in the following theorem.

The Minimum Principle

Theorem 10.1: The Minimum Principle

For a given initial condition $x(0) = \mathrm{x} \in \mathcal{S}$, let $u(t)$ be an optimal control trajectory with associated state trajectory $x(t)$ for the system

$$\dot{x}(t) = f(x(t), u(t)), \quad 0 \leq t \leq T \qquad \text{s.t.} \quad x(0) = \mathrm{x}.$$

Then a trajectory $p(t)$ exists such that:

$$\dot{p}(t) = -\left.\frac{\partial H(\mathrm{x}, \mathrm{u}, \mathrm{p})}{\partial \mathrm{x}}\right|^{\top}_{\substack{x(t)\\ u(t)\\ p(t)}}, \quad p(T) = \left.\frac{\partial h(\mathrm{x})}{\partial \mathrm{x}}\right|^{\top}_{x(T)},$$

$$u(t) = \arg\min_{\mathrm{u} \in \mathcal{U}} H(x(t), \mathrm{u}, p(t)),$$

$$H(x(t), u(t), p(t)) = \text{ constant}, \quad \forall t \in [0, T],$$

where $H(\mathrm{x}, \mathrm{u}, \mathrm{p}) := g(\mathrm{x}, \mathrm{u}) + \mathrm{p}^{\top} f(\mathrm{x}, \mathrm{u})$ is called the *Hamiltonian function*.

Proof of the Minimum Principle (1/6)

We provide an informal proof which assumes that the optimal cost-to-go $J(t, \mathrm{x})$ and the optimal policy $\mu(\cdot, \cdot)$ are continuously differentiable.

These assumptions are not needed, but they simplify the proof.

With continuously differentiable cost-to-go, the HJB equation is also a necessary condition for optimality and we have:

$$0 = \min_{\mathrm{u} \in \mathcal{U}} \underbrace{\left(g(\mathrm{x}, \mathrm{u}) + \frac{\partial J(t, \mathrm{x})}{\partial t} + \frac{\partial J(t, \mathrm{x})}{\partial \mathrm{x}} f(\mathrm{x}, \mathrm{u}) \right)}_{=: F(t, \mathrm{x}, \mathrm{u})}, \quad \forall t \in [0, T], \ \forall \mathrm{x} \in \mathcal{S}, \tag{1}$$

$$= \min_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u})$$

$$J(T, \mathrm{x}) = h(\mathrm{x}), \quad \forall \mathrm{x} \in \mathcal{S}. \tag{2}$$

Proof of the Minimum Principle (2/6)

Let $\mu(t, \mathrm{x})$ be the minimizing input: $\min_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u}) = F(t, \mathrm{x}, \mu(t, \mathrm{x}))$.

Since $F(t, \mathrm{x}, \mu(t, \mathrm{x})) = 0$ for all $t \geq 0$, $\mathrm{x} \in \mathcal{S}$ and by Lemma 10.1:

$$\begin{aligned} 0 &= \frac{\partial \left(\min_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u}) \right)}{\partial t} = \left. \frac{\partial F(t, \mathrm{x}, \mathrm{u})}{\partial t} \right|_{\mu(t, \mathrm{x})} \\ &= \frac{\partial^2 J(t, \mathrm{x})}{\partial t^2} + f(\mathrm{x}, \mu(t, \mathrm{x}))^\top \frac{\partial^2 J(t, \mathrm{x})}{\partial t \partial \mathrm{x}}, \end{aligned} \tag{3}$$

Similarly,

$$\begin{aligned} 0 &= \frac{\partial \left(\min_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u}) \right)}{\partial \mathrm{x}} = \left. \frac{\partial F(t, \mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{\mu(t, \mathrm{x})} \\ &= \left. \frac{\partial g(\mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{\mu(t, \mathrm{x})} + \frac{\partial^2 J(t, \mathrm{x})}{\partial \mathrm{x} \partial t} + f(\mathrm{x}, \mu(t, \mathrm{x}))^\top \frac{\partial^2 J(t, \mathrm{x})}{\partial \mathrm{x}^2} \\ &\quad + \left. \frac{\partial J(t, \mathrm{x})}{\partial \mathrm{x}} \frac{\partial f(\mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{\mu(t, \mathrm{x})}. \end{aligned} \tag{4}$$

Proof of the Minimum Principle (3/6)

Now consider the optimal input trajectory $u(t) := \mu(t, x(t))$ and optimal state trajectory $x(t)$, where

$$\dot{x}(t) = f(x(t), u(t)) = \left.\frac{\partial H(\mathrm{x}, \mathrm{u}, \mathrm{p}))}{\partial \mathrm{p}}\right|_{\substack{x(t)\\u(t)}}^{\top},$$

$$x(0) = \mathrm{x}.$$

Along this optimal trajectory, (3) evaluates to

$$\begin{aligned} 0 &= \left.\frac{\partial^2 J(t, \mathrm{x})}{\partial t^2}\right|_{x(t)} + \dot{x}(t)^\top \left.\frac{\partial^2 J(t, \mathrm{x})}{\partial t \partial \mathrm{x}}\right|_{x(t)} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\underbrace{\left.\frac{\partial J(t, \mathrm{x})}{\partial t}\right|_{x(t)}}_{=:r(t)}\right). \end{aligned}$$

In particular, $r(t)$ is a constant.

Proof of the Minimum Principle (4/6)

Along the same optimal trajectory, (4) evaluates to

$$
\begin{aligned}
0 = &\left.\frac{\partial g(\mathrm{x},\mathrm{u})}{\partial \mathrm{x}}\right|_{\substack{x(t)\\u(t)}} + \left.\frac{\partial^2 J(t,\mathrm{x})}{\partial \mathrm{x}\partial t}\right|_{x(t)} + \dot{x}(t)^\top \left.\frac{\partial^2 J(t,\mathrm{x})}{\partial \mathrm{x}^2}\right|_{x(t)} \\
&+ \left.\frac{\partial J(t,\mathrm{x})}{\partial \mathrm{x}}\right|_{x(t)} \left.\frac{\partial f(\mathrm{x},\mathrm{u})}{\partial \mathrm{x}}\right|_{\substack{x(t)\\u(t)}} \\
= &\left.\frac{\partial g(\mathrm{x},\mathrm{u})}{\partial \mathrm{x}}\right|_{\substack{x(t)\\u(t)}} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\underbrace{\left.\frac{\partial J(t,\mathrm{x})}{\partial \mathrm{x}}\right|_{x(t)}}_{=:p(t)^\top}\right) + \left.\frac{\partial J(t,\mathrm{x})}{\partial \mathrm{x}}\right|_{x(t)} \left.\frac{\partial f(\mathrm{x},\mathrm{u})}{\partial \mathrm{x}}\right|_{\substack{x(t)\\u(t)}} .
\end{aligned}
$$

In particular,

$$
\dot{p}(t) = -\left.\frac{\partial g(\mathrm{x},\mathrm{u})}{\partial \mathrm{x}}\right|_{\substack{x(t)\\u(t)}}^\top - \left.\frac{\partial f(\mathrm{x},\mathrm{u})}{\partial \mathrm{x}}\right|_{\substack{x(t)\\u(t)}}^\top p(t) = -\left.\frac{\partial H(\mathrm{x},\mathrm{u},\mathrm{p})}{\partial \mathrm{x}}\right|_{\substack{x(t)\\u(t)\\p(t)}}^\top .
$$

Proof of the Minimum Principle (5/6)

The partial derivative of the boundary condition (2) with respect to x yields

$$\frac{\partial J(T,\mathrm{x})}{\partial \mathrm{x}} = \frac{\partial h(\mathrm{x})}{\partial \mathrm{x}}, \quad \forall \mathrm{x} \in \mathcal{S}$$

and thus

$$p(T) = \left. \frac{\partial J(t,\mathrm{x})}{\partial \mathrm{x}} \right|_{T,x(T)}^{\top} = \left. \frac{\partial h(\mathrm{x})}{\partial \mathrm{x}} \right|_{x(T)}^{\top}.$$

From (1), we have

$$\begin{aligned} -\frac{\partial J(t,\mathrm{x})}{\partial t} &= \min_{\mathrm{u} \in \mathcal{U}} \left(g(\mathrm{x},\mathrm{u}) + \frac{\partial J(t,\mathrm{x})}{\partial \mathrm{x}} f(\mathrm{x},\mathrm{u}) \right) \\ &= \min_{\mathrm{u} \in \mathcal{U}} H\left(\mathrm{x}, \mathrm{u}, \frac{\partial J(t,\mathrm{x})}{\partial \mathrm{x}}^{\top} \right), \end{aligned}$$

which along the optimal trajectory is

$$H(x(t), u(t), p(t)) = -r(t) = \text{constant}.$$

Proof of the Minimum Principle (6/6)

Furthermore:

$$
\begin{aligned}
u(t) &= \arg\min_{\mathrm{u}\in\mathcal{U}} F(t, x(t), \mathrm{u}) \\
&= \arg\min_{\mathrm{u}\in\mathcal{U}} \left(g(x(t), \mathrm{u}) + \left.\frac{\partial J(t, \mathrm{x})}{\partial \mathrm{x}}\right|_{x(t)} f(x(t), \mathrm{u}) \right) \\
&= \arg\min_{\mathrm{u}\in\mathcal{U}} \big(g(x(t), \mathrm{u}) + p(t)^\top f(x(t), \mathrm{u}) \big) \\
&= \arg\min_{\mathrm{u}\in\mathcal{U}} H(x(t), \mathrm{u}, p(t)).
\end{aligned}
$$

The Minimum Principle – Remarks

The Minimum Principle requires solving an ODE with split boundary conditions. It is not trivial to solve, but it is easier than solving the PDE in the HJB equation.

The Minimum Principle provides necessary conditions for optimality. If a control trajectory satisfies these conditions, it is not necessarily optimal.

Further analysis is needed to guarantee optimality. One method that often works is to prove that an optimal control trajectory exists and to verify that there is only one control trajectory satisfying the conditions of the Minimum Principle.

Example: Preparing a Martian Base (1/6)

Consider a problem where a fleet of general purpose robots is sent to Mars at time 0 to help build a Martian base.

They can be used in two ways:

1. they can replicate themselves;
2. they can make habitats for human-beings.

The number of robots at time t is denoted by $x(t)$, and the number of habitats at time t by $z(t)$.

We want to maximize the size of the Martian base at the terminal time T.

Example: Preparing a Martian Base (2/6)

The dynamics are given by:

$$\begin{aligned} \dot{x}(t) &= u(t)x(t), & x(0) > 0, \\ \dot{z}(t) &= (1-u(t))x(t), & z(0) = 0, \\ 0 &\leq u(t) \leq 1, \end{aligned}$$

where $u(t)$ denotes the fraction of $x(t)$ used to reproduce themselves.

Objective: find a control input $u(t)$ that maximizes $z(T)$.

Note that the utility function to maximize can be written as a function of $x(t)$ and $u(t)$,

$$z(T) = \int_0^T \dot{z}(t)\mathrm{d}t = \int_0^T (1-u(t))x(t)\mathrm{d}t,$$

and that $z(t)$ does not enter into the dynamics of $x(t)$. Therefore, we consider $x(t)$ as the only state and $u(t)$ as the control input.

The stage utility rate is $g(\mathrm{x},\mathrm{u}) = (1-\mathrm{u})\mathrm{x}$, the terminal utility is $h(\mathrm{x}) = 0$, and the dynamics are $f(\mathrm{x},\mathrm{u}) = \mathrm{ux}$.

Example: Preparing a Martian Base (3/6)

Thus the Hamiltonian is:

$$H(\mathrm{x}, \mathrm{u}, \mathrm{p}) = (1 - \mathrm{u})\mathrm{x} + \mathrm{pux}.$$

We apply Theorem 10.1:

$$\begin{aligned}
\dot{p}(t) &= - \left. \frac{\partial H(\mathrm{x}, \mathrm{u}, \mathrm{p})}{\partial \mathrm{x}} \right|^{\top}_{x(t), u(t), p(t)} = -1 + u(t) - p(t)u(t), \\
p(T) &= \left. \frac{\partial h(\mathrm{x})}{\partial \mathrm{x}} \right|^{\top}_{x(T)} = 0, \\
\dot{x}(t) &= x(t)u(t), \quad x(0) = \mathrm{x}, \\
u(t) &= \underset{0 \le \mathrm{u} \le 1}{\arg\max}\, H(x(t), \mathrm{u}, p(t)) = \underset{0 \le \mathrm{u} \le 1}{\arg\max}\, (x(t) + x(t)(p(t) - 1)\mathrm{u}).
\end{aligned}$$

Since $x(t) > 0$ for $t \in [0, T]$, we can find the following solution:

$$u(t) = \begin{cases} 0 & \text{if } p(t) < 1 \\ 1 & \text{if } p(t) > 1 \\ \text{undetermined} & \text{if } p(t) = 1 \end{cases}.$$

Example: Preparing a Martian Base (4/6)

To have $p(t) = 1$ for a non-trivial time interval, we need

$$0 = \dot{p}(t) = -1 + u - p(t)u(t) = -1 \Rightarrow \text{absurd.}$$

Thus, $p(t) = 1$ may happen only for an inconsequential time interval, and we can arbitrarily pick $u(t) = 1$ for $p(t) = 1$.

We will now work backwards from $t = T$.

Since $p(T) = 0$, for t close to T, we have $u(t) = 0$ and therefore $\dot{p}(t) = -1$.

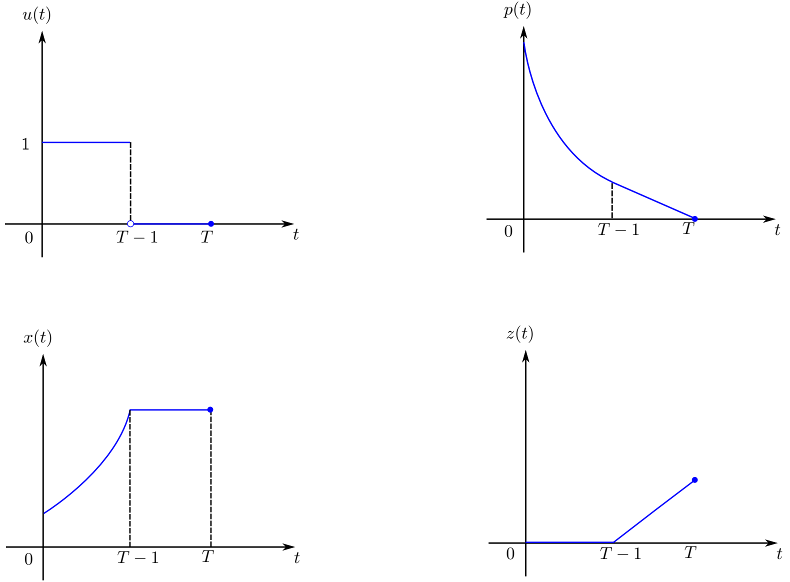
Therefore at time $t = T - 1$, $p(t) = 1$ and that is when the control input switches to $u(t) = 1$. Thus for $t \leq T - 1$:

$$\begin{aligned} \dot{p}(t) &= -p(t), \quad p(T-1) = 1 \\ \Rightarrow p(t) &= e^{(T-1)}e^{-t}. \end{aligned}$$

Thus $p(t)$ is bigger than 1 for $t < T - 1$ till time 0, hence we have:

$$u(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq T-1 \\ 0 & \text{if } T-1 < t \leq T \end{cases}.$$

Example: Preparing a Martian Base (5/6)



Example: Preparing a Martian Base (6/6)

The optimal strategy is to use all the robots to replicate themselves from time 0 until $t = T - 1$, and then use all the robots to build habitats.

If $T < 1$, then the robots should only build habitats.

The derived strategy is optimal because it is the only one that satisfies the necessary optimality conditions (up to the inconsequential time instant when $p(t) = 1$, we will see how to handle these "singular arcs" in the next lecture). In fact, our solution is constructive.

In general, if the Hamiltonian is linear in u, the maximum or minimum of the Hamiltonian can only be attained on the boundaries of $\mathcal{U}$.

The resulting control trajectory is known as bang-bang control.

The Minimum Principle – Summary (1/2)

The Minimum Principle is a necessary condition for optimality:

- All the optimal trajectories must satisfy the conditions outlined in Theorem 10.1.
- It is possible that non-optimal trajectories satisfy the conditions outlined in Theorem 10.1.

As we saw in the previous lecture, the HJB equation is a sufficient condition for optimality:

- If a solution satisfies the HJB equation, then we are guaranteed that it is indeed optimal.
- If a candidate solution is differentiable, it must satisfy the HJB equation (in this case, the HJB equation is also necessary for optimality).

The Minimum Principle – Summary (2/2)

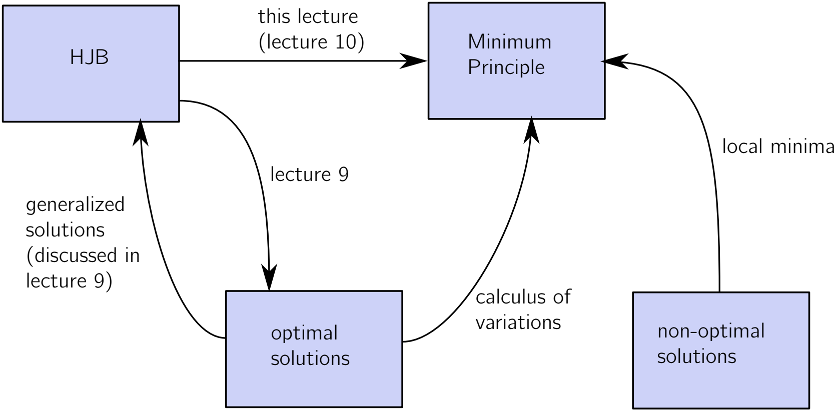


Figure: Optimal solutions and their relation to the HJB equation and the Minimum Principle.

Outline

Additional reading material

Just as the Minimum Principle seeks a control function that minimizes (or maximizes) the Hamiltonian, the *calculus of variations* seeks a function that minimizes (or maximizes) a functional. The necessary optimality conditions in calculus of variations are expressed via the *Euler-Lagrange equation*.

Calculus of variations is a fascinating field. If you are interested in a historical perspective on the development of the Minimum Principle and its relation with calculus of variations, you can read:
`https://ethz.ch/content/dam/ethz/special-interest/mavt/dynamic-systems-n-control/idsc-dam/Lectures/Optimal-Control/Additional%20Material/additional-material.pdf`