

# ECE276B: Planning & Learning in Robotics

## Lecture 16: Pontryagin's Minimum Principle

Instructor:

Nikolay Atanasov: [natanasov@ucsd.edu](mailto:natanasov@ucsd.edu)

Teaching Assistants:

Thai Duong: [tduong@eng.ucsd.edu](mailto:tduong@eng.ucsd.edu)

**UC San Diego**

**JACOBS SCHOOL OF ENGINEERING**  
Electrical and Computer Engineering

# Deterministic Continuous-time Optimal Control

► **Problem statement:**

$$\begin{aligned} \min_{\pi} \quad & V^{\pi}(0, \mathbf{x}_0) := \int_0^T \ell(\mathbf{x}(t), \pi(t, \mathbf{x}(t))) dt + q(\mathbf{x}(T)) \\ \text{s.t.} \quad & \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0 \\ & \mathbf{x}(t) \in \mathcal{X}, \quad \pi(t, \mathbf{x}(t)) \in PC^0([0, T], \mathcal{U}) \end{aligned}$$

- **Admissible policies:**  $PC^0([0, T], \mathcal{U})$  is the set of piecewise continuous functions from  $[0, T]$  to  $\mathcal{U}$
- **Optimal value function:**  $V^*(t, \mathbf{x}) = \min_{\pi} V^{\pi}(t, \mathbf{x})$

## Relationship to Mechanics

- ▶ **Costate:**  $\mathbf{p}(t)$  is the gradient/sensitivity of the optimal value function  $V^*(t, \mathbf{x})$  with respect to the state  $\mathbf{x}$ .
- ▶ **Hamiltonian:** captures the total energy of the system:

$$H(\mathbf{x}, \mathbf{u}, \mathbf{p}) := \ell(\mathbf{x}, \mathbf{u}) + \mathbf{p}^\top \mathbf{f}(\mathbf{x}, \mathbf{u})$$

- ▶ **Hamilton's principle of least action:** trajectories of mechanical systems are extremals of the action integral  $\int_0^T \ell(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt$ , where the Lagrangian  $\ell(\mathbf{x}, \dot{\mathbf{x}}) := K(\dot{\mathbf{x}}) - U(\mathbf{x})$  is the difference between kinetic and potential energy.
- ▶ If the stage cost is the Lagrangian of a mechanical system, the Hamiltonian is the (negative) total energy (kinetic plus potential)

# Lagrangian Mechanics

- ▶ Consider a point mass  $m$  with position  $\mathbf{x}$  and velocity  $\dot{\mathbf{x}}$
- ▶ Kinetic energy  $K(\dot{\mathbf{x}}) := \frac{1}{2} m \|\dot{\mathbf{x}}\|_2^2$  and momentum  $\mathbf{p} := m\dot{\mathbf{x}}$
- ▶ Potential energy  $U(\mathbf{x})$  and conservative force  $\mathbf{F} = -\frac{\partial U(\mathbf{x})}{\partial \mathbf{x}}$
- ▶ Newtonian equations of motion:  $\mathbf{F} = m\ddot{\mathbf{x}}$
- ▶ Note that  $-\frac{\partial U(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{F} = m\ddot{\mathbf{x}} = \frac{d}{dt}\mathbf{p} = \frac{d}{dt}\left(\frac{\partial K(\dot{\mathbf{x}})}{\partial \dot{\mathbf{x}}}\right)$
- ▶ Note that  $\frac{\partial U(\mathbf{x})}{\partial \dot{\mathbf{x}}} = 0$  and  $\frac{\partial K(\dot{\mathbf{x}})}{\partial \mathbf{x}} = 0$
- ▶ Lagrangian:  $\ell(\mathbf{x}, \dot{\mathbf{x}}) := K(\dot{\mathbf{x}}) - U(\mathbf{x})$
- ▶ Euler-Lagrange equation:  $\frac{d}{dt}\left(\frac{\partial \ell(\mathbf{x}, \dot{\mathbf{x}})}{\partial \dot{\mathbf{x}}}\right) - \frac{\partial \ell(\mathbf{x}, \dot{\mathbf{x}})}{\partial \mathbf{x}} = 0$

## Conservation of Energy

▶ Total energy  $E(\mathbf{x}, \dot{\mathbf{x}}) = K(\dot{\mathbf{x}}) + U(\mathbf{x}) = 2K(\dot{\mathbf{x}}) - \ell(\mathbf{x}, \dot{\mathbf{x}}) = \mathbf{p}^\top \dot{\mathbf{x}} - \ell(\mathbf{x}, \dot{\mathbf{x}})$

▶ Note that:

$$\frac{d}{dt} \left( \mathbf{p}^\top \dot{\mathbf{x}} \right) = \frac{d}{dt} \left( \frac{\partial \ell(\mathbf{x}, \dot{\mathbf{x}})^\top}{\partial \dot{\mathbf{x}}} \dot{\mathbf{x}} \right) = \left( \frac{d}{dt} \frac{\partial \ell(\mathbf{x}, \dot{\mathbf{x}})}{\partial \dot{\mathbf{x}}} \right)^\top \dot{\mathbf{x}} + \frac{\partial \ell(\mathbf{x}, \dot{\mathbf{x}})^\top}{\partial \dot{\mathbf{x}}} \ddot{\mathbf{x}}$$

$$\frac{d}{dt} \ell(\mathbf{x}, \dot{\mathbf{x}}) = \frac{\partial \ell(\mathbf{x}, \dot{\mathbf{x}})^\top}{\partial \mathbf{x}} \dot{\mathbf{x}} + \frac{\partial \ell(\mathbf{x}, \dot{\mathbf{x}})^\top}{\partial \dot{\mathbf{x}}} \ddot{\mathbf{x}} + \frac{\partial}{\partial t} \ell(\mathbf{x}, \dot{\mathbf{x}})$$

▶ Conservation of energy using the Euler-Lagrange equation:

$$\frac{d}{dt} E(\mathbf{x}, \dot{\mathbf{x}}) = \frac{d}{dt} \left( \frac{\partial \ell(\mathbf{x}, \dot{\mathbf{x}})^\top}{\partial \dot{\mathbf{x}}} \dot{\mathbf{x}} \right) - \frac{d}{dt} \ell(\mathbf{x}, \dot{\mathbf{x}}) = -\frac{\partial}{\partial t} \ell(\mathbf{x}, \dot{\mathbf{x}}) = 0$$

▶ In our formulation, the costate is the negative momentum and the Hamiltonian is the negative total energy

- ▶ **Extremal open-loop trajectories** (i.e., local minima) can be computed by solving a boundary-value ODE with initial **state**  $\mathbf{x}(0)$  and terminal **costate**  $\mathbf{p}(T) = \nabla_{\mathbf{x}}q(\mathbf{x})$

## Theorem: Pontryagin's Minimum Principle (PMP)

- ▶ Let  $\mathbf{u}^*(t) : [0, T] \rightarrow \mathcal{U}$  be an optimal control trajectory
- ▶ Let  $\mathbf{x}^*(t) : [0, T] \rightarrow \mathcal{X}$  be the associated state trajectory from  $\mathbf{x}_0$
- ▶ Then, there exists a **costate trajectory**  $\mathbf{p}^*(t) : [0, T] \rightarrow \mathcal{X}$  satisfying:
  1. **Canonical equations with boundary conditions:**

$$\begin{aligned}\dot{\mathbf{x}}^*(t) &= \nabla_{\mathbf{p}}H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t)), & \mathbf{x}^*(0) &= \mathbf{x}_0 \\ \dot{\mathbf{p}}^*(t) &= -\nabla_{\mathbf{x}}H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t)), & \mathbf{p}^*(T) &= \nabla_{\mathbf{x}}q(\mathbf{x}^*(T))\end{aligned}$$

2. **Minimum principle with constant (holonomic) constraint:**

$$\begin{aligned}\mathbf{u}^*(t) &= \arg \min_{\mathbf{u} \in \mathcal{U}(\mathbf{x}^*(t))} H(\mathbf{x}^*(t), \mathbf{u}, \mathbf{p}^*(t)), & \forall t \in [0, T] \\ H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t)) &= \text{constant}, & \forall t \in [0, T]\end{aligned}$$

- ▶ **Proof:** Liberzon, Calculus of Variations & Optimal Control, Ch. 4.2

## Proof of PMP (Step 0: Preliminaries)

### Lemma: $\nabla$ -min Exchange

Let  $F(t, \mathbf{x}, \mathbf{u})$  be continuously differentiable in  $t \in \mathbb{R}$ ,  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{u} \in \mathbb{R}^m$  and let  $\mathcal{U} \subseteq \mathbb{R}^m$  be a convex set. Assume  $\pi^*(t, \mathbf{x}) = \arg \min_{\mathbf{u} \in \mathcal{U}} F(t, \mathbf{x}, \mathbf{u})$  exists and is continuously differentiable. Then, for all  $t$  and  $\mathbf{x}$ :

$$\frac{\partial}{\partial t} \left( \min_{\mathbf{u} \in \mathcal{U}} F(t, \mathbf{x}, \mathbf{u}) \right) = \frac{\partial}{\partial t} F(t, \mathbf{x}, \mathbf{u}) \Big|_{\mathbf{u}=\pi^*(t, \mathbf{x})} \quad \nabla_{\mathbf{x}} \left( \min_{\mathbf{u} \in \mathcal{U}} F(t, \mathbf{x}, \mathbf{u}) \right) = \nabla_{\mathbf{x}} F(t, \mathbf{x}, \mathbf{u}) \Big|_{\mathbf{u}=\pi^*(t, \mathbf{x})}$$

► **Proof:** Let  $G(t, \mathbf{x}) := \min_{\mathbf{u} \in \mathcal{U}} F(t, \mathbf{x}, \mathbf{u}) = F(t, \mathbf{x}, \pi^*(t, \mathbf{x}))$ . Then:

$$\frac{\partial}{\partial t} G(t, \mathbf{x}) = \frac{\partial}{\partial t} F(t, \mathbf{x}, \mathbf{u}) \Big|_{\mathbf{u}=\pi^*(t, \mathbf{x})} + \underbrace{\frac{\partial}{\partial \mathbf{u}} F(t, \mathbf{x}, \mathbf{u}) \Big|_{\mathbf{u}=\pi^*(t, \mathbf{x})}}_{=0 \text{ by 1st order optimality condition}} \frac{\partial \pi^*(t, \mathbf{x})}{\partial t}$$

A similar derivation can be used for the partial derivative wrt  $\mathbf{x}$ .

## Proof of PMP (Step 1: HJB PDE gives $V^*(t, \mathbf{x})$ )

- ▶ **Extra Assumptions:**  $V^*(t, \mathbf{x})$  and  $\pi^*(t, \mathbf{x})$  are continuously differentiable in  $t$  and  $\mathbf{x}$  and  $\mathcal{U}$  is convex. These assumptions can be avoided in a more general proof.
- ▶ With a continuously differentiable value function, the HJB PDE is also a necessary condition for optimality:

$$V^*(T, \mathbf{x}) = q(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{X}$$

$$0 = \min_{\mathbf{u} \in \mathcal{U}} \underbrace{\left( \ell(\mathbf{x}, \mathbf{u}) + \frac{\partial}{\partial t} V^*(t, \mathbf{x}) + \nabla_{\mathbf{x}} V^*(t, \mathbf{x})^\top \mathbf{f}(\mathbf{x}, \mathbf{u}) \right)}_{:=F(t, \mathbf{x}, \mathbf{u})}, \quad \forall t \in [0, T], \mathbf{x} \in \mathcal{X}$$

with a corresponding optimal policy  $\pi^*(t, \mathbf{x})$ .



## Proof of PMP (Step 2: $\nabla$ -min Exchange Lemma)

- ▶ Apply the  $\nabla$ -min Exchange Lemma to the HJB PDE:

$$0 = \frac{\partial}{\partial t} \left( \min_{\mathbf{u} \in \mathcal{U}} F(t, \mathbf{x}, \mathbf{u}) \right) = \frac{\partial^2}{\partial t^2} V^*(t, \mathbf{x}) + \left[ \frac{\partial}{\partial t} \nabla_{\mathbf{x}} V^*(t, \mathbf{x}) \right]^\top \mathbf{f}(\mathbf{x}, \pi^*(t, \mathbf{x}))$$

$$\begin{aligned} 0 &= \nabla_{\mathbf{x}} \left( \min_{\mathbf{u} \in \mathcal{U}} F(t, \mathbf{x}, \mathbf{u}) \right) \\ &= \nabla_{\mathbf{x}} \ell(\mathbf{x}, \mathbf{u}^*) + \nabla_{\mathbf{x}} \frac{\partial}{\partial t} V^*(t, \mathbf{x}) + [\nabla_{\mathbf{x}}^2 V^*(t, \mathbf{x})] \mathbf{f}(\mathbf{x}, \mathbf{u}^*) + [\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{u}^*)]^\top \nabla_{\mathbf{x}} V^*(t, \mathbf{x}) \end{aligned}$$

where  $\mathbf{u}^* := \pi^*(t, \mathbf{x})$

- ▶ Evaluate these along the trajectory  $\mathbf{x}^*(t)$  resulting from  $\pi^*(t, \mathbf{x}^*(t))$ :

$$\dot{\mathbf{x}}^*(t) = \mathbf{f}(\mathbf{x}^*(t), \mathbf{u}^*(t)) = \nabla_{\mathbf{p}} H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}), \quad \mathbf{x}^*(0) = \mathbf{x}_0$$

## Proof of PMP (Step 3: Evaluate along $\mathbf{x}^*(t), \mathbf{u}^*(t)$ )

- Evaluate the results of Step 2 along  $\mathbf{x}^*(t)$ :

$$\begin{aligned} 0 &= \frac{\partial^2 V^*(t, \mathbf{x})}{\partial t^2} \Big|_{\mathbf{x}=\mathbf{x}^*(t)} + \left[ \frac{\partial}{\partial t} \nabla_{\mathbf{x}} V^*(t, \mathbf{x}) \Big|_{\mathbf{x}=\mathbf{x}^*(t)} \right]^\top \dot{\mathbf{x}}^*(t) \\ &= \frac{d}{dt} \left( \underbrace{\frac{\partial}{\partial t} V^*(t, \mathbf{x}) \Big|_{\mathbf{x}=\mathbf{x}^*(t)}}_{:=r(t)} \right) = \frac{d}{dt} r(t) \Rightarrow r(t) = \text{const. } \forall t \end{aligned}$$

$$\begin{aligned} 0 &= \nabla_{\mathbf{x}} \ell(\mathbf{x}, \mathbf{u}^*) \Big|_{\mathbf{x}=\mathbf{x}^*(t)} + \frac{d}{dt} \left( \underbrace{\nabla_{\mathbf{x}} V^*(t, \mathbf{x}) \Big|_{\mathbf{x}=\mathbf{x}^*(t)}}_{:=\mathbf{p}^*(t)} \right) \\ &\quad + [\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{u}^*) \Big|_{\mathbf{x}=\mathbf{x}^*(t)}]^\top [\nabla_{\mathbf{x}} V^*(t, \mathbf{x}) \Big|_{\mathbf{x}=\mathbf{x}^*(t)}] \\ &= \nabla_{\mathbf{x}} \ell(\mathbf{x}, \mathbf{u}^*) \Big|_{\mathbf{x}=\mathbf{x}^*(t)} + \dot{\mathbf{p}}^*(t) + [\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{u}^*) \Big|_{\mathbf{x}=\mathbf{x}^*(t)}]^\top \mathbf{p}^*(t) \\ &= \dot{\mathbf{p}}^*(t) + \nabla_{\mathbf{x}} H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t)) \end{aligned}$$

## Proof of PMP (Step 4: Done)

- ▶ The boundary condition  $V^*(T, \mathbf{x}) = q(\mathbf{x})$  implies that  $\nabla_{\mathbf{x}} V^*(T, \mathbf{x}) = \nabla_{\mathbf{x}} q(\mathbf{x})$  for all  $\mathbf{x} \in \mathcal{X}$  and thus  $\mathbf{p}^*(T) = \nabla_{\mathbf{x}} q(\mathbf{x}^*(T))$
- ▶ From the HJB PDE we have:

$$-\frac{\partial}{\partial t} V^*(t, \mathbf{x}) = \min_{\mathbf{u} \in \mathcal{U}} H(\mathbf{x}, \mathbf{u}, \nabla_{\mathbf{x}} V^*(t, \cdot))$$

which along the optimal trajectory  $\mathbf{x}^*(t)$ ,  $\mathbf{u}^*(t)$  becomes:

$$-r(t) = H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t)) = \text{const}$$

- ▶ Finally, note that

$$\begin{aligned} \mathbf{u}^*(t) &= \arg \min_{\mathbf{u} \in \mathcal{U}} F(t, \mathbf{x}^*(t), \mathbf{u}) \\ &= \arg \min_{\mathbf{u} \in \mathcal{U}} \left\{ \ell(\mathbf{x}^*(t), \mathbf{u}) + [\nabla_{\mathbf{x}} V^*(t, \mathbf{x})|_{\mathbf{x}=\mathbf{x}^*(t)}]^\top \mathbf{f}(\mathbf{x}^*(t), \mathbf{u}) \right\} \\ &= \arg \min_{\mathbf{u} \in \mathcal{U}} \left\{ \ell(\mathbf{x}^*(t), \mathbf{u}) + \mathbf{p}^*(t)^\top \mathbf{f}(\mathbf{x}^*(t), \mathbf{u}) \right\} \\ &= \arg \min_{\mathbf{u} \in \mathcal{U}} H(\mathbf{x}^*(t), \mathbf{u}, \mathbf{p}^*(t)) \end{aligned}$$

## HJB PDE vs PMP

- ▶ The HJB PDE provides a lot of information – the optimal value function and an optimal policy for all time and all states!
- ▶ Often, we only care about the optimal trajectory for a specific initial condition  $\mathbf{x}_0$ . Exploiting that we need less information, we can arrive at simpler conditions for optimality – Pontryagin's Minimum Principle
- ▶ The PMP does **not apply to infinite horizon problems**, so one has to use the HJB PDE in that case
- ▶ The HJB PDE is a **sufficient condition** for optimality: it is possible that the optimal solution does not satisfy it but a solution that satisfies it is guaranteed to be optimal
- ▶ The PMP is a **necessary condition** for optimality: it is possible that non-optimal trajectories satisfy it so further analysis is necessary to determine if a candidate PMP policy is optimal
- ▶ The PMP requires solving an ODE with split boundary conditions (not easy but much easier than the nonlinear HJB PDE!)

## Example: Resource Allocation for a Martian Base

- ▶ A fleet of reconfigurable, general purpose robots is sent to Mars at  $t = 0$
- ▶ The robots can 1) replicate or 2) make human habitats
- ▶ The number of robots at time  $t$  is  $x(t)$ , while the number of habitats is  $z(t)$  and they evolve according to:

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

where  $u(t)$  denotes the percentage of the  $x(t)$  robots used for replication

- ▶ Goal: Maximize the size of the Martian base by a terminal time  $T$ , i.e.:

$$\max z(T) = \int_0^T (1 - u(t))x(t)dt$$

with  $f(x, u) = ux$ ,  $\ell(x, u) = -(1 - u)x$  and  $q(x) = 0$

## Example: Resource Allocation for a Martian Base

▶ Hamiltonian:  $H(x, u, p) = -(1 - u)x + pux$

▶ Apply the PMP:

$$\dot{x}^*(t) = \nabla_p H(x^*, u^*, p^*) = x^*(t)u^*(t), \quad x^*(0) = x$$

$$\dot{p}^*(t) = -\nabla_x H(x^*, u^*, p^*) = (1 - u^*(t)) - p^*(t)u^*(t), \quad p^*(T) = 0$$

$$u^*(t) = \arg \min_{0 \leq u \leq 1} H(x^*(t), u, p^*(t)) = \arg \min_{0 \leq u \leq 1} (x^*(t)(p^*(t) + 1)u)$$

▶ Since  $x^*(t) > 0$  for  $t \in [0, T]$ :

$$u^*(t) = \begin{cases} 0 & \text{if } p^*(t) > -1 \\ 1 & \text{if } p^*(t) \leq -1 \end{cases}$$

## Example: Resource Allocation for a Martian Base

- ▶ Work backwards from  $t = T$  to determine  $p^*(t)$ :
  - ▶ Since  $p^*(T) = 0$  for  $t$  close to  $T$ , we have  $u^*(t) = 0$  and the costate dynamics become  $\dot{p}^*(t) = 1$
  - ▶ At time  $t = T - 1$ ,  $p^*(t) = -1$  and the control input switches to  $u^*(t) = 1$
  - ▶ For  $t \leq T - 1$ :

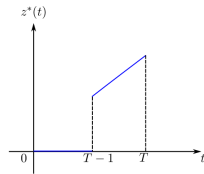
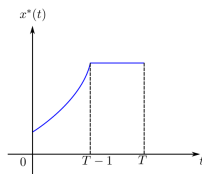
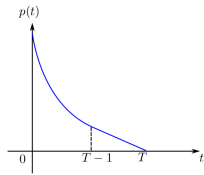
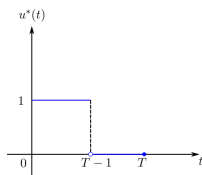
$$\begin{aligned} \dot{p}^*(t) &= -p^*(t), \quad p(T - 1) = -1 \\ \Rightarrow p^*(t) &= e^{-[(T-1)-t]} p(T - 1) \leq -1 \quad \text{for } t < T - 1 \end{aligned}$$

- ▶ Optimal control:

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

## Example: Resource Allocation for a Martian Base

- ▶ Optimal trajectories for the Martian resource allocation problem:



- ▶ **Conclusions:**

- ▶ All robots replicate themselves from  $t = 0$  to  $t = T - 1$  and then all robots build habitats
- ▶ If  $T < 1$ , then the robots should only build habitats
- ▶ If the Hamiltonian is linear in  $u$ , its min can only be attained on the boundary of  $\mathcal{U}$ , known as **bang-bang control**



## PMP with Fixed Terminal State

- ▶ Suppose that in addition to  $\mathbf{x}(0) = \mathbf{x}_s$ , a final state  $\mathbf{x}(T) = \mathbf{x}_\tau$  is given.
- ▶ The terminal cost  $q(\mathbf{x}(T))$  is not useful since  $V^*(T, \mathbf{x}) = \infty$  if  $\mathbf{x}(T) \neq \mathbf{x}_\tau$ . The terminal boundary condition for the costate  $\mathbf{p}(T) = \nabla_{\mathbf{x}} q(\mathbf{x}(T))$  does not hold but as compensation we have a different boundary condition  $\mathbf{x}(T) = \mathbf{x}_\tau$ .
- ▶ We still have  $2n$  ODEs with  $2n$  boundary conditions:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), & \mathbf{x}(0) &= \mathbf{x}_s, \mathbf{x}(T) = \mathbf{x}_\tau \\ \dot{\mathbf{p}}(t) &= -\nabla_{\mathbf{x}} H(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}(t))\end{aligned}$$

- ▶ If only some terminal state are fixed  $\mathbf{x}_j(T) = \mathbf{x}_{\tau,j}$  for  $j \in I$ , then:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), & \mathbf{x}(0) &= \mathbf{x}_s, \mathbf{x}_j(T) = \mathbf{x}_{\tau,j}, \quad \forall j \in I \\ \dot{\mathbf{p}}(t) &= -\nabla_{\mathbf{x}} H(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}(t)), & \mathbf{p}_j(T) &= \frac{\partial}{\partial x_j} q(\mathbf{x}(T)), \quad \forall j \notin I\end{aligned}$$

## PMP with Fixed Terminal Set

- ▶ **Terminal set:** a  $k$  dim surface in  $\mathbb{R}^n$  requiring:

$$\mathbf{x}(T) \in \mathcal{X}_\tau = \{\mathbf{x} \in \mathbb{R}^n \mid h_j(\mathbf{x}) = 0, j = 1, \dots, n - k\}$$

- ▶ The costate boundary condition requires that  $\mathbf{p}(T)$  is orthogonal to the tangent space  $D = \{\mathbf{d} \in \mathbb{R}^n \mid \nabla_x h_j(\mathbf{x}(T))^\top \mathbf{d} = 0, j = 1, \dots, n - k\}$ :

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(0) = \mathbf{x}_s, \quad h_j(\mathbf{x}(T)) = 0, j = 1, \dots, n - k$$

$$\dot{\mathbf{p}}(t) = -\nabla_x H(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}(t)), \quad \mathbf{p}(T) \in \mathbf{span}\{\nabla_x h_j(\mathbf{x}(T)), \forall j\}$$

$$\text{OR } \mathbf{d}^\top \mathbf{p}(T) = 0, \forall \mathbf{d} \in D$$

## PMP with Free Initial State

- ▶ Suppose that  $\mathbf{x}_0$  is free and subject to optimization with additional cost  $\ell_0(\mathbf{x}_0)$  term
- ▶ The total cost becomes  $\ell_0(\mathbf{x}_0) + V(0, \mathbf{x}_0)$  and the necessary condition for an optimal initial state  $\mathbf{x}_0$  is:

$$\nabla_{\mathbf{x}} \ell_0(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_0} + \underbrace{\nabla_{\mathbf{x}} V(0, \mathbf{x})|_{\mathbf{x}=\mathbf{x}_0}}_{=\mathbf{p}(0)} = 0 \quad \Rightarrow \quad \mathbf{p}(0) = -\nabla_{\mathbf{x}} \ell_0(\mathbf{x}_0)$$

- ▶ We lose the initial state boundary condition but gain an adjoint state boundary condition:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t))$$

$$\dot{\mathbf{p}}(t) = -\nabla_{\mathbf{x}} H(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}(t)), \quad \mathbf{p}(0) = -\nabla_{\mathbf{x}} \ell_0(\mathbf{x}_0), \quad \mathbf{p}(T) = \nabla_{\mathbf{x}} q(\mathbf{x}(T))$$

- ▶ Similarly, we can deal with some parts of the initial state being free and some not

## PMP with Free Terminal Time

- ▶ Suppose that the initial and/or terminal state are given but the terminal time  $T$  is free and subject to optimization
- ▶ We can compute the total cost of optimal trajectories for various terminal times  $T$  and look for the best choice, i.e.:

$$\left. \frac{\partial}{\partial t} V^*(t, \mathbf{x}) \right|_{t=T, \mathbf{x}=\mathbf{x}(T)} = 0$$

- ▶ Recall that on the optimal trajectory:

$$H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t)) = - \left. \frac{\partial}{\partial t} V^*(t, \mathbf{x}) \right|_{\mathbf{x}=\mathbf{x}^*(t)} = \text{const.} \quad \forall t$$

- ▶ Hence, in the free terminal time case, we gain an extra degree of freedom with free  $T$  but lose one degree of freedom by the constraint:

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

## PMP with Time-varying System and Cost

- ▶ Suppose that the system and stage cost vary with time:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \quad \ell(\mathbf{x}(t), \mathbf{u}(t), t)$$

- ▶ A usual trick is to convert the problem to a time-invariant one by making  $t$  part of the state. Let  $y(t) = t$  with dynamics:

$$\dot{y}(t) = 1, \quad y(0) = 0$$

- ▶ Augmented state  $\mathbf{z}(t) := (\mathbf{x}(t), y(t))$  and system:

$$\dot{\mathbf{z}}(t) = \bar{\mathbf{f}}(\mathbf{z}(t), \mathbf{u}(t)) := \begin{bmatrix} \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), y(t)) \\ 1 \end{bmatrix}$$

$$\bar{\ell}(\mathbf{z}, \mathbf{u}) := \ell(\mathbf{x}, \mathbf{u}, y) \quad \bar{q}(\mathbf{z}) := q(\mathbf{x})$$

- ▶ The Hamiltonian need not to be constant along the optimal trajectory:

$$H(\mathbf{x}, \mathbf{u}, \mathbf{p}, t) = \ell(\mathbf{x}, \mathbf{u}, t) + \mathbf{p}^\top \mathbf{f}(\mathbf{x}, \mathbf{u}, t)$$

$$\dot{\mathbf{x}}^*(t) = \mathbf{f}(\mathbf{x}^*(t), \mathbf{u}^*(t), t), \quad \mathbf{x}^*(0) = \mathbf{x}_0$$

$$\dot{\mathbf{p}}^*(t) = -\nabla_{\mathbf{x}} H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t), \quad \mathbf{p}^*(T) = \nabla_{\mathbf{x}} q(\mathbf{x}^*(T))$$

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

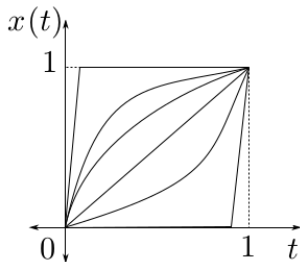
$$H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \neq \text{const}$$

# Singular Problems

- ▶ The minimum condition  $\mathbf{u}(t) = \arg \min_{\mathbf{u} \in \mathcal{U}} H(\mathbf{x}^*(t), \mathbf{u}, \mathbf{p}^*(t), t)$  may be insufficient to determine  $\mathbf{u}^*(t)$  for all  $t$  in some cases because the values of  $\mathbf{x}^*(t)$  and  $\mathbf{p}^*(t)$  are such that  $H(\mathbf{x}^*(t), \mathbf{u}, \mathbf{p}^*(t), t)$  is independent of  $\mathbf{u}$  over a nontrivial interval of time
- ▶ The optimal trajectories consist of portions where  $\mathbf{u}^*(t)$  can be determined from the minimum condition (**regular arcs**) and where  $\mathbf{u}^*(t)$  cannot be determined from the minimum condition since the Hamiltonian is independent of  $\mathbf{u}$  (**singular arcs**)

## Example: Fixed Terminal State

- ▶ System:  $\dot{x}(t) = u(t)$ ,  $x(0) = 0$ ,  $x(1) = 1$ ,  $u(t) \in \mathbb{R}$
- ▶ Cost:  $\min \frac{1}{2} \int_0^1 (x(t)^2 + u(t)^2) dt$
- ▶ Want  $x(t)$  and  $u(t)$  to be small but need to meet  $x(1) = 1$



- ▶ Approach: use PMP to find a locally optimal open-loop policy

## Example: Fixed Terminal State

### ► Pontryagin's Minimum Principle

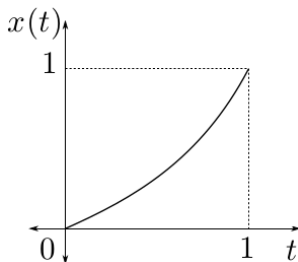
- Hamiltonian:  $H(x, u, p) = \frac{1}{2}(x^2 + u^2) + pu$
- Minimum principle:  $u(t) = \arg \min_{u \in \mathbb{R}} \left\{ \frac{1}{2}(x(t)^2 + u^2) + p(t)u \right\} = -p(t)$
- Canonical equations with boundary conditions:

$$\begin{aligned}\dot{x}(t) &= \nabla_p H(x(t), u(t), p(t)) = u(t) = -p(t), & x(0) &= 0, & x(1) &= 1 \\ \dot{p}(t) &= -\nabla_x H(x(t), u(t), p(t)) = -x(t)\end{aligned}$$

### ► Candidate trajectory: $\ddot{x}(t) = x(t) \Rightarrow x(t) = ae^t + be^{-t} = \frac{e^t - e^{-t}}{e - e^{-1}}$

- $x(0) = 0 \Rightarrow a + b = 0$
- $x(1) = 1 \Rightarrow ae + be^{-1} = 1$

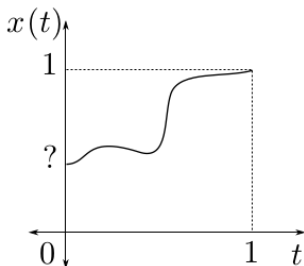
### ► Open-loop control: $u(t) = \dot{x}(t) = \frac{e^t + e^{-t}}{e - e^{-1}}$





## Example: Free Initial State

- ▶ System:  $\dot{x}(t) = u(t)$ ,  $x(0) = \text{free}$ ,  $x(1) = 1$ ,  $u(t) \in \mathbb{R}$
- ▶ Cost:  $\min \frac{1}{2} \int_0^1 (x(t)^2 + u(t)^2) dt$
- ▶ Picking  $x(0) = 1$  will allow  $u(t) = 0$  but we will accumulate cost due to  $x(t)$ . On the other hand, picking  $x(0) = 0$  will accumulate cost due to  $u(t)$  having to drive the state to  $x(1) = 1$ .



- ▶ Approach: use PMP to find a locally optimal open-loop policy

## Example: Free Initial State

### ► Pontryagin's Minimum Principle

- Hamiltonian:  $H(x, u, p) = \frac{1}{2}(x^2 + u^2) + pu$
- Minimum principle:  $u(t) = \arg \min_{u \in \mathbb{R}} \left\{ \frac{1}{2}(x(t)^2 + u^2) + p(t)u \right\} = -p(t)$
- Canonical equations with boundary conditions:

$$\dot{x}(t) = \nabla_p H(x(t), u(t), p(t)) = u(t) = -p(t), \quad x(1) = 1$$

$$\dot{p}(t) = -\nabla_x H(x(t), u(t), p(t)) = -x(t), \quad p(0) = 0$$

### ► Candidate trajectory:

$$\ddot{x}(t) = x(t) \Rightarrow x(t) = ae^t + be^{-t} = \frac{e^t + e^{-t}}{e + e^{-1}}$$

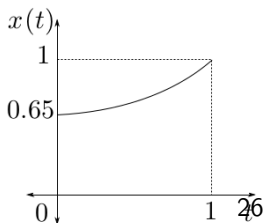
$$p(t) = -\dot{x}(t) = -ae^t + be^{-t} = \frac{-e^t + e^{-t}}{e + e^{-1}}$$

- $x(1) = 1 \Rightarrow ae + be^{-1} = 1$

- $p(0) = 0 \Rightarrow -a + b = 0$

- $x(0) \approx 0.65$

- Open-loop control:  $u(t) = \dot{x}(t) = \frac{e^t - e^{-t}}{e + e^{-1}}$



## Example: Free Terminal Time

- ▶ System:  $\dot{x}(t) = u(t)$ ,  $x(0) = 0$ ,  $x(T) = 1$ ,  $u(t) \in \mathbb{R}$
- ▶ Cost:  $\min \int_0^T 1 + \frac{1}{2}(x(t)^2 + u(t)^2)dt$
- ▶ Free terminal time:  $T = \text{free}$
- ▶ Note: if we do not include 1 in the stage-cost (i.e., use the same cost as before), we would get  $T^* = \infty$  (see next slide for details)
- ▶ Approach: use PMP to find a locally optimal open-loop policy

## Example: Free Terminal Time

### ▶ Pontryagin's Minimum Principle

- ▶ Hamiltonian:  $H(x(t), u(t), p(t)) = 1 + \frac{1}{2}(x(t)^2 + u(t)^2) + p(t)u(t)$
- ▶ Minimum principle:  $u(t) = \arg \min_{u \in \mathbb{R}} \left\{ \frac{1}{2}(x(t)^2 + u^2) + p(t)u \right\} = -p(t)$
- ▶ Canonical equations with boundary conditions:

$$\begin{aligned}\dot{x}(t) &= \nabla_p H(x(t), u(t), p(t)) = u(t) = -p(t), & x(0) &= 0, & x(T) &= 1 \\ \dot{p}(t) &= -\nabla_x H(x(t), u(t), p(t)) = -x(t)\end{aligned}$$

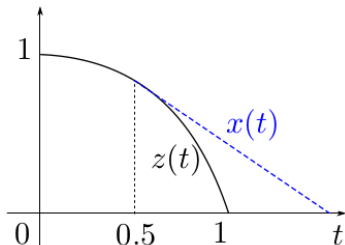
- ▶ Candidate trajectory:  $\ddot{x}(t) = x(t) \Rightarrow x(t) = ae^t + be^{-t} = \frac{e^t - e^{-t}}{e^T - e^{-T}}$ 
  - ▶  $x(0) = 0 \Rightarrow a + b = 0$
  - ▶  $x(T) = 1 \Rightarrow ae^T + be^{-T} = 1$

### ▶ Free terminal time:

$$\begin{aligned}0 &= H(x(t), u(t), p(t)) = 1 + \frac{1}{2}(x(t)^2 - p(t)^2) \\ &= 1 + \frac{1}{2} \left( \frac{(e^t - e^{-t})^2 - (e^t + e^{-t})^2}{(e^T - e^{-T})^2} \right) = 1 - \frac{2}{(e^T - e^{-T})^2} \\ &\Rightarrow T \approx 0.66\end{aligned}$$

## Example: Time-varying Singular Problem

- ▶ System:  $\dot{x}(t) = u(t)$ ,  $x(0) = \text{free}$ ,  $x(1) = \text{free}$ ,  $u(t) \in [-1, 1]$
- ▶ Time-varying cost:  $\min \frac{1}{2} \int_0^1 (x(t) - z(t))^2 dt$  for  $z(t) = 1 - t^2$
- ▶ Example feasible state trajectory that tracks the desired  $z(t)$  until the slope of  $z(t)$  becomes less than  $-1$  and the input  $u(t)$  saturates:



- ▶ Approach: use PMP to find a locally optimal open-loop policy

## Example: Time-varying Singular Problem

- ▶ Pontryagin's Minimum Principle

- ▶ Hamiltonian:  $H(x, u, p, t) = \frac{1}{2}(x - z(t))^2 + pu$

- ▶ Minimum principle:

$$u(t) = \arg \min_{|u| \leq 1} H(x(t), u, p(t), t) = \begin{cases} -1 & \text{if } p(t) > 0 \\ \text{undetermined} & \text{if } p(t) = 0 \\ 1 & \text{if } p(t) < 0 \end{cases}$$

- ▶ Canonical equations with boundary conditions:

$$\dot{x}(t) = \nabla_p H(x(t), u(t), p(t)) = u(t),$$

$$\dot{p}(t) = -\nabla_x H(x(t), u(t), p(t)) = -(x(t) - z(t)), \quad p(0) = 0, \quad p(1) = 0$$

- ▶ **Singular arc:** when  $p(t) = 0$  for a non-trivial time interval, the control cannot be determined from PMP

- ▶ In this example, the singular arc can be determined from the costate ODE. For  $p(t) = 0$ :

$$0 \equiv \dot{p}(t) = -x(t) + z(t) \quad \Rightarrow \quad u(t) = \dot{x}(t) = \dot{z}(t) = -2t$$

## Example: Time-varying Singular Problem

► Since  $p(0) = 0$ , the state trajectory follows a singular arc until  $t_s \leq \frac{1}{2}$  (since  $u(t) = -2t \in [-1, 1]$ ) when it switches to a regular arc with  $u(t) = -1$  (since  $z(t)$  is decreasing and we are trying to track it).

► For  $0 \leq t \leq t_s \leq \frac{1}{2}$ :  $x(t) = z(t)$   $p(t) = 0$

► For  $t_s < t \leq 1$ :

$$\dot{x}(t) = -1 \Rightarrow x(t) = z(t_s) - \int_{t_s}^t ds = 1 - t_s^2 - t + t_s$$

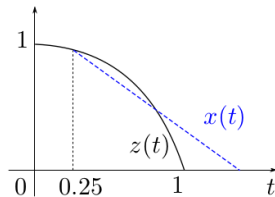
$$\dot{p}(t) = -(x(t) - z(t)) = t_s^2 - t_s - t^2 + t, \quad p(t_s) = p(1) = 0$$

$$\Rightarrow p(s) = p(t_s) + \int_{t_s}^s (t_s^2 - t_s - t^2 + t) dt, \quad s \in [t_s, 1]$$

$$\Rightarrow 0 = p(1) = t_s^2 - t_s - \frac{1}{3} + \frac{1}{2} - t_s^3 + t_s^2 + \frac{t_s^3}{3} - \frac{t_s^2}{2}$$

$$\Rightarrow 0 = (t_s - 1)^2(1 - 4t_s)$$

$$\Rightarrow \boxed{t_s = \frac{1}{4}}$$



## Discrete-time PMP

- ▶ Consider a discrete-time problem with dynamics  $\mathbf{x}_{t+1} = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t)$
- ▶ Introduce Lagrange multipliers  $\mathbf{p}_{0:T}$  to relax the constraints:

$$\begin{aligned} L(\mathbf{x}_{0:T}, \mathbf{u}_{0:T-1}, \mathbf{p}_{0:T}) &= q(\mathbf{x}_T) + \mathbf{x}_0^\top \mathbf{p}_0 + \sum_{t=0}^{T-1} \ell(\mathbf{x}_t, \mathbf{u}_t) + (\mathbf{f}(\mathbf{x}_t, \mathbf{u}_t) - \mathbf{x}_{t+1})^\top \mathbf{p}_{t+1} \\ &= q(\mathbf{x}_T) + \mathbf{x}_0^\top \mathbf{p}_0 - \mathbf{x}_T^\top \mathbf{p}_T + \sum_{t=0}^{T-1} H(\mathbf{x}_t, \mathbf{u}_t, \mathbf{p}_{t+1}) - \mathbf{x}_t^\top \mathbf{p}_t \end{aligned}$$

- ▶ Setting  $\nabla_{\mathbf{x}} L = \nabla_{\mathbf{p}} L = 0$  and explicitly minimizing wrt  $\mathbf{u}_{0:T-1}$  yields:

### Theorem: Discrete-time PMP

If  $\mathbf{x}_{0:T}^*$ ,  $\mathbf{u}_{0:T-1}^*$  is an optimal state-control trajectory starting at  $\mathbf{x}_0$ , then there exists a **costate trajectory**  $\mathbf{p}_{0:T}^*$  such that:

$$\begin{aligned} \mathbf{x}_{t+1}^* &= \nabla_{\mathbf{p}} H(\mathbf{x}_t^*, \mathbf{u}_t^*, \mathbf{p}_{t+1}^*) = \mathbf{f}(\mathbf{x}_t^*, \mathbf{u}_t^*), & \mathbf{x}_0^* &= \mathbf{x}_0 \\ \mathbf{p}_t^* &= \nabla_{\mathbf{x}} H(\mathbf{x}_t^*, \mathbf{u}_t^*, \mathbf{p}_{t+1}^*) = \nabla_{\mathbf{x}} \ell(\mathbf{x}_t^*, \mathbf{u}_t^*) + \nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}_t^*, \mathbf{u}_t^*)^\top \mathbf{p}_{t+1}^*, & \mathbf{p}_T^* &= \nabla_{\mathbf{x}} q(\mathbf{x}_T^*) \\ \mathbf{u}_t^* &= \arg \min_{\mathbf{u}} H(\mathbf{x}_t^*, \mathbf{u}, \mathbf{p}_{t+1}^*) \end{aligned}$$



## Gradient of the Value Function via the PMP

- ▶ The discrete-time PMP provides an efficient way to evaluate the gradient of the value function with respect to  $\mathbf{u}$  and thus optimize control trajectories locally and numerically

### Theorem: Value Function Gradient

Given an initial state  $\mathbf{x}_0$  and trajectory  $\mathbf{u}_{0:T-1}$ , let  $\mathbf{x}_{1:T}$ ,  $\mathbf{p}_{0:T}$  be such that:

$$\begin{aligned}\mathbf{x}_{t+1} &= \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t), & \mathbf{x}_0 \text{ given} \\ \mathbf{p}_t &= \nabla_{\mathbf{x}} \ell(\mathbf{x}_t, \mathbf{u}_t) + [\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t)]^\top \mathbf{p}_{t+1}, & \mathbf{p}_T = \nabla_{\mathbf{x}} q(\mathbf{x}_T)\end{aligned}$$

Then:

$$\nabla_{\mathbf{u}_t} V(\mathbf{x}_{0:T}, \mathbf{u}_{0:T-1}) = \nabla_{\mathbf{u}} H(\mathbf{x}_t, \mathbf{u}_t, \mathbf{p}_{t+1}) = \nabla_{\mathbf{u}} \ell(\mathbf{x}_t, \mathbf{u}_t) + \nabla_{\mathbf{u}} \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t)^\top \mathbf{p}_{t+1}$$

- ▶ Note that  $\mathbf{x}_t$  can be found in a forward pass (since it does not depend on  $\mathbf{p}$ ) and then  $\mathbf{p}_t$  can be found in a backward pass

## Proof by Induction

- ▶ The accumulated cost can be written recursively:

$$V_t(\mathbf{x}_{t:T}, \mathbf{u}_{t:T-1}) = \ell(\mathbf{x}_t, \mathbf{u}_t) + V_{t+1}(\mathbf{x}_{t+1:T}, \mathbf{u}_{t+1:T-1})$$

- ▶ Note that  $\mathbf{u}_t$  affects the future costs only through  $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t)$ :

$$\nabla_{\mathbf{u}_t} V_t(\mathbf{x}_{t:T}, \mathbf{u}_{t:T-1}) = \nabla_{\mathbf{u}} \ell(\mathbf{x}_t, \mathbf{u}_t) + [\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t)]^\top \nabla_{\mathbf{x}_{t+1}} V_{t+1}(\mathbf{x}_{t+1:T}, \mathbf{u}_{t+1:T-1})$$

- ▶ **Claim:**  $\mathbf{p}_t = \nabla_{\mathbf{x}_t} V_t(\mathbf{x}_{t:T}, \mathbf{u}_{t:T-1})$ :

- ▶ Base case:  $\mathbf{p}_T = \nabla_{\mathbf{x}_T} q(\mathbf{x}_T)$
- ▶ Induction: for  $t \in [0, T)$ :

$$\underbrace{\nabla_{\mathbf{x}_t} V_t(\mathbf{x}_{t:T}, \mathbf{u}_{t:T-1})}_{=\mathbf{p}_t} = \nabla_{\mathbf{x}} \ell(\mathbf{x}_t, \mathbf{u}_t) + [\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t)]^\top \underbrace{\nabla_{\mathbf{x}_{t+1}} V_{t+1}(\mathbf{x}_{t+1:T}, \mathbf{u}_{t+1:T-1})}_{=\mathbf{p}_{t+1}}$$

which is identical with the costate difference equation.