10. Pontryagin's Minimum Principle

The HJB equation provides a lot of information: the optimal cost-to-go and the optimal policy for all time and for all possible states. However, in many cases, we only care about the optimal control trajectory for a specific initial condition. We will see how to exploit the fact that we are asking for much less in order to arrive at simpler conditions for optimality (the Minimum Principle).

10.1 Notation

Let $F(t, \mathrm{x})$ be a continuous, differentiable function. Then,

- The partial derivative of F with respect to its first argument, t, is $\dfrac{\partial F(t, \mathrm{x})}{\partial t}$.
- The partial derivative of F with respect to t when subject to $\mathrm{x} = x(t)$ is

$$\frac{\partial F(t, x(t))}{\partial t} = \left.\frac{\partial F(t, \mathrm{x})}{\partial t}\right|_{\mathrm{x}=x(t)} + \left.\frac{\partial F(t, \mathrm{x})}{\partial \mathrm{x}}\right|_{\mathrm{x}=x(t)} \frac{\partial x(t)}{\partial t}.$$

- The total derivative of F with respect to t when subject to $\mathrm{x} = x(t)$ is

$$\frac{\mathrm{d}F(t, x(t))}{\mathrm{d}t} = \left.\frac{\partial F(t, \mathrm{x})}{\partial t}\right|_{\mathrm{x}=x(t)} + \left.\frac{\partial F(t, \mathrm{x})}{\partial \mathrm{x}}\right|_{\mathrm{x}=x(t)} \frac{\mathrm{d}x(t)}{\mathrm{d}t}$$

  Since $x(t)$ is only a function of t, then

$$\frac{\mathrm{d}F(t, x(t))}{\mathrm{d}t} = \frac{\partial F(t, x(t))}{\partial t}.$$

- We sometimes use the short-hand notation $\left.\dfrac{\partial F(t, \mathrm{x})}{\partial t}\right|_{x(t)} := \left.\dfrac{\partial F(t, \mathrm{x})}{\partial t}\right|_{\mathrm{x}=x(t)}$

Example 1:
Consider $F(t, \mathrm{x}) = t\mathrm{x}$. Then,

$$\begin{aligned}\frac{\partial F(t, \mathrm{x})}{\partial t} &= \mathrm{x} \\ \frac{\mathrm{d}F(t, x(t))}{\mathrm{d}t} &= x(t) + t\dot{x}(t).\end{aligned}$$

$\triangle$

Date compiled: December 5, 2018

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 a convex set. Furthermore, assume $\mu^*(t, \mathrm{x}) := \arg\min_{\mathrm{u}\in\mathcal{U}} F(t, \mathrm{x}, \mathrm{u})$ exists and is continuously differentiable. Then for all t and x,*

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

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

Proof. We prove this for when $\mathcal{U} = \mathbb{R}^m$. 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, this can be shown for the partial derivative with respect to x. □

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

$$\min_{\mathrm{u}\in\mathbb{R}} F(t,\mathrm{x},\mathrm{u}) : 2(1+t)\mathrm{u} + \mathrm{x} = 0, \quad \mathrm{u} = -\frac{\mathrm{x}}{2(1+t)}$$

$$\therefore \mu^*(t,\mathrm{x}) = -\frac{\mathrm{x}}{2(1+t)}$$

$$\therefore \min_{\mathrm{u}\in\mathbb{R}} F(t,\mathrm{x},\mathrm{u}) = \frac{(1+t)\mathrm{x}^2}{4(1+t)^2} - \frac{\mathrm{x}^2}{2(1+t)} + 1 = -\frac{\mathrm{x}^2}{4(1+t)} + 1$$

1.

$$\begin{aligned}\frac{\partial\left(\min_{\mathrm{u}\in\mathbb{R}} F(t,\mathrm{x},\mathrm{u})\right)}{\partial t} &= \frac{\partial F(t,\mathrm{x},\mu^*(t,\mathrm{x}))}{\partial t} \\ &= \frac{\mathrm{x}^2}{4(1+t)^2}\end{aligned}$$

And by Lemma 10.1,

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

2.

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

And by Lemma 10.1,

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

△

10.2 The Minimum Principle

Consider the following continuous-time set-up.

Dynamics

$$\dot{x}(t) = f(x(t), u(t)), \qquad 0 \le t \le T \tag{10.1}$$

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 $u(t) \in \mathcal{U} \subset \mathbb{R}^m,\ \forall t \in [0, T]$. $\mathcal{U}$ is the control constraint set;
- $f(\cdot, \cdot)$: function capturing system evolution that is continuously differentiable in x.

Cost

We consider the following scalar-valued cost function:

$$h(x(T)) + \int_0^T g(x(\tau), u(\tau))\mathrm{d}\tau \tag{10.2}$$

where g and h are both continuously differentiable in x.

Objective

Given an initial condition $x(0) = \mathrm{x} \in \mathcal{S}$, construct an optimal control trajectory $u(t)$ such that (10.2) subject to (10.1) is minimized.

One could potentially solve the above problem using the HJB which gives an optimal policy $\mu^*(t, \mathrm{x})$. The optimal input trajectory can then be inferred from the policy for a given initial condition and the solution to $\dot{x}(t) = f(x(t), \mu^*(t, x(t)))$. However, as we have seen, solving the HJB is very difficult in general. The following theorem can be used instead which gives necessary conditions on the optimal trajectory (henceforth the star superscript $\cdot^*$ will be dropped).

Theorem 10.1. *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* (10.1)*. Then there exists a trajectory* $p(t)$ *such that:*

$$
\begin{aligned}
\dot{p}(t) &= -\left.\frac{\partial H(\mathrm{x}, \mathrm{u}, \mathrm{p})}{\partial \mathrm{x}}\right|_{x(t), u(t), p(t)}^{\top}, \quad p(T) = \left.\frac{\partial h(\mathrm{x})}{\partial \mathrm{x}}\right|_{x(T)}^{\top} \\
u(t) &= \arg\min_{\mathrm{u} \in \mathcal{U}} H(x(t), \mathrm{u}, p(t)) \\
H(x(t), u(t), p(t)) &= \text{constant} \quad \forall t \in [0, T]
\end{aligned}
$$

where $H(\mathrm{x}, \mathrm{u}, \mathrm{p}) := g(\mathrm{x}, \mathrm{u}) + \mathrm{p}^\top f(\mathrm{x}, \mathrm{u})$.

The function $H(\cdot, \cdot, \cdot)$ in the above is referred to the Hamiltonian function, which comes from Hamilton's Principle of Least Action from mechanics.

Proof. We provide an informal proof which assumes that the cost-to-go $J(t, \mathrm{x})$ is continuously differentiable, the optimal policy $\mu(\cdot, \cdot)$ is continuously differentiable, and $\mathcal{U}$ is convex in order to make use of Lemma 10.1. However, these assumptions are actually not needed in a more formal proof.
With continuously differentiable cost-to-go, the HJB is also a necessary condition for optimality:

$$
\begin{aligned}
0 &= \min_{\mathrm{u} \in \mathcal{U}} \underbrace{\left(g(\mathrm{x}, \mathrm{u}) + \frac{\partial J(t, \mathrm{x})}{\partial t} + \frac{\partial J(t, \mathrm{x})}{\partial \mathrm{x}} f(\mathrm{x}, \mathrm{u}) \right)}_{=: F(t, \mathrm{x}, \mathrm{u})}, \quad \forall t \in [0, T], \ \forall \mathrm{x} \in \mathcal{S} && (10.3) \\
&= \min_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u}) && (10.4) \\
J(T, \mathrm{x}) &= h(\mathrm{x}) \quad \forall \mathrm{x} \in \mathcal{S} && (10.5)
\end{aligned}
$$

with $\mathrm{u} = \mu(t, \mathrm{x})$ the corresponding optimal control strategy.

Now take the partial derivatives of (10.4) with respect to t and x; by Lemma 10.1

$$
\begin{aligned}
0 &= \frac{\partial \left(\min_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u}) \right)}{\partial t} = \left.\frac{\partial F(t, \mathrm{x}, \mathrm{u})}{\partial t}\right|_{\mu(t, \mathrm{x})} \\
&= \frac{\partial^2 J(t, \mathrm{x})}{\partial t^2} + \frac{\partial^2 J(t, \mathrm{x})}{\partial t \partial \mathrm{x}} f(\mathrm{x}, \mu(t, \mathrm{x})), && (10.6)
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
0 &= \frac{\partial \left(\min\limits_{\mathrm{u} \in \mathcal{U}} F(t, \mathrm{x}, \mathrm{u}) \right)}{\partial \mathrm{x}} = \left. \frac{\partial F(t, \mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{\mu(t,\mathrm{x})} \\
&= \left. \frac{\partial g(\mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{\mu(t,\mathrm{x})} + \frac{\partial^2 J(t, \mathrm{x})}{\partial \mathrm{x} \partial t} + f(\mathrm{x}, \mu(t, \mathrm{x}))^\top \frac{\partial^2 J(t, \mathrm{x})}{\partial \mathrm{x}^2} + \left. \frac{\partial J(t, \mathrm{x})}{\partial \mathrm{x}} \frac{\partial f(\mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{\mu(t,\mathrm{x})} . \qquad (10.7)
\end{aligned}
$$

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

$$
\begin{aligned}
\dot{x}(t) &= f(x(t), u(t)) = \left. \frac{\partial H(\mathrm{x}, \mathrm{u}, \mathrm{p}))}{\partial \mathrm{p}} \right|_{x(t),u(t)}^\top \\
x(0) &= \mathrm{x}.
\end{aligned}
$$

Along this optimal trajectory, (10.6) becomes

$$
\begin{aligned}
0 &= \left. \frac{\partial^2 J(t, \mathrm{x})}{\partial t^2} \right|_{x(t)} + \left. \frac{\partial^2 J(t, \mathrm{x})}{\partial t \partial \mathrm{x}} \right|_{x(t)} \dot{x}(t) \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \left(\underbrace{\left. \frac{\partial J(t, \mathrm{x})}{\partial t} \right|_{x(t)}}_{=:r(t)} \right), \qquad (10.8)
\end{aligned}
$$

and (10.7) becomes

$$
\begin{aligned}
0 &= \left. \frac{\partial g(\mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{x(t),u(t)} + \left. \frac{\partial^2 J(t, \mathrm{x})}{\partial \mathrm{x} \partial t} \right|_{x(t)} + \dot{x}(t)^\top \left. \frac{\partial^2 J(t, \mathrm{x})}{\partial \mathrm{x}^2} \right|_{x(t)} + \left. \frac{\partial J(t, \mathrm{x})}{\partial \mathrm{x}} \right|_{x(t)} \left. \frac{\partial f(\mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{x(t),u(t)} \\
&= \left. \frac{\partial g(\mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{x(t),u(t)} + \frac{\mathrm{d}}{\mathrm{d}t} \left(\underbrace{\left. \frac{\partial J(t, \mathrm{x})}{\partial \mathrm{x}} \right|_{x(t)}}_{=:p(t)^\top} \right) + \left. \frac{\partial J(t, \mathrm{x})}{\partial \mathrm{x}} \right|_{x(t)} \left. \frac{\partial f(\mathrm{x}, \mathrm{u})}{\partial \mathrm{x}} \right|_{x(t),u(t)} . \qquad (10.9)
\end{aligned}
$$

With $r(t) := \left. \frac{\partial J(t, \mathrm{x})}{\partial t} \right|_{x(t)}$, (10.8) becomes

$$
\dot{r}(t) = 0 \qquad \Rightarrow r(t) = \text{ constant } \forall t
$$

and with $p(t) := \left. \frac{\partial J(t, \mathrm{x})}{\partial \mathrm{x}} \right|_{x(t)}^\top$, (10.9) becomes

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

Taking the partial derivative of the boundary condition (10.5) 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 (10.3), 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

$$-r(t) = H(x(t), u(t), p(t))$$

which is constant. Furthermore, note that

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

□

Remarks:

- The Minimum Principle requires solving an ODE with split boundary conditions. It is not trivial to solve, but easier than solving a PDE in the HJB.
- 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 3: Resource allocation problem: preparing a Martian base
Consider a problem where a fleet of reconfigurable, general purpose robots is sent to Mars at time 0 to help build a Martian base. They can be used for two things: 1) They can replicate themselves; 2) They can make habitats for human-beings. The number of the robots at time t is denoted by $x(t)$, and the number of the habitats by $z(t)$. We want to maximize the size of the Martian base at the terminal time T.

The dynamics are given by

$$\dot{x}(t) = u(t)x(t), \quad x(0) = \mathrm{x} > 0,$$
$$\dot{z}(t) = (1 - u(t))x(t), \quad z(0) = 0,$$
$$0 \leq u(t) \leq 1.$$

where $u(t)$ denotes the fraction of $x(t)$ used to reproduce themselves.
Objective: find a control trajectory $u(t)$ that maximizes $z(T)$.

Solution

Note that the cost function can be written as a function of $x(t)$ and $u(t)$:

$$z(T) = \int_0^T (1 - u(t))x(t)\mathrm{d}t \tag{10.10}$$

and $z(t)$ does not enter into the dynamics of $x(t)$, we can therefore consider $x(t)$ as the only state and $u(t)$ as the control input. The stage cost is then $g(\mathrm{x}, \mathrm{u}) = (1 - \mathrm{u})\mathrm{x}$, the terminal cost is $h(\mathrm{x}) = 0$, and the dynamics function is $f(\mathrm{x}, \mathrm{u}) = \mathrm{ux}$. Thus the Hamiltonian is

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

Applying Theorem 10.1,

$$\dot{p}(t) = -\left.\frac{\partial H(\mathrm{x}, \mathrm{u}, \mathrm{p})}{\partial \mathrm{x}}\right|_{x(t),u(t),p(t)}^{\top} = -1 + u(t) - p(t)u(t) \tag{10.11}$$

$$p(T) = \left.\frac{\partial h(\mathrm{x})}{\partial \mathrm{x}}\right|_{x(T)}^{\top} = 0$$

$$\dot{x}(t) = x(t)u(t), \quad x(0) = \mathrm{x}$$

$$u(t) = \underset{0 \leq \mathrm{u} \leq 1}{\arg\max}\, H(x(t), \mathrm{u}, p(t)) = \underset{0 \leq \mathrm{u} \leq 1}{\arg\max}\, (x(t) + x(t)(p(t) - 1)\mathrm{u}) \tag{10.12}$$

Since $x(t) > 0$ for $t \in [0, T]$, from (10.12) we can find the following solution[1]

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

We will now work backwards from $t = T$. Since $p(T) = 0$, for t close to T, we have $u(t) = 0$ and therefore (10.11) becomes $\dot{p}(t) = -1$. Therefore at time $t = T - 1$, $p(t) = 1$ and that is when the control input switches to $u(t) = 1$. Thus for $t \leq T - 1$:

$$\dot{p}(t) = -p(t), \quad p(T - 1) = 1$$

$$\Rightarrow p(t) = e^{(T-1)}e^{-t} \tag{10.13}$$

Note that by (10.13) $p(t)$ is bigger than 1 for $t < T - 1$ till time 0, hence we have

$$u(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq T - 1 \\ 0 & \text{if } T - 1 < t \leq T \end{cases}. \tag{10.14}$$

[1] Note that when $p(t) = 1$, $u(t)$ can be anything between 0 and 1. It can be shown that this choice does not make a difference in the incurred cost.

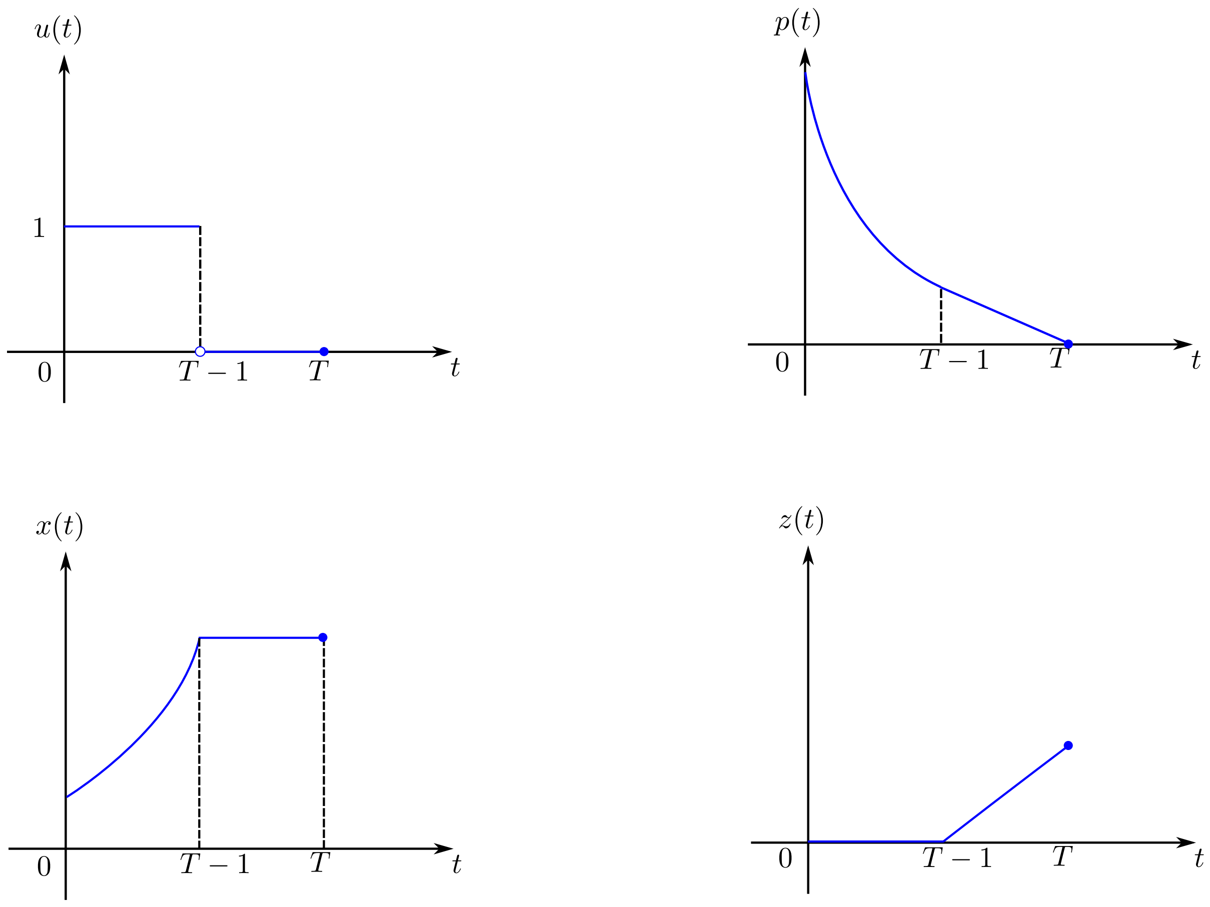

Figure 10.1: Optimal trajectories for the resource allocation example.

In conclusion, an optimal control trajectory 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. 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. $\triangle$

10.3 Summary

The Minimum Principle is a necessary condition for the optimal trajectory; in particular, it is possible that non-optimal trajectories satisfy the conditions outlined in Theorem 10.1. As we saw last lecture, the HJB is a sufficient condition for optimality; in particular, if a solution satisfies the HJB, then we are guaranteed that it is indeed optimal. This is summarized in Fig. 10.2.

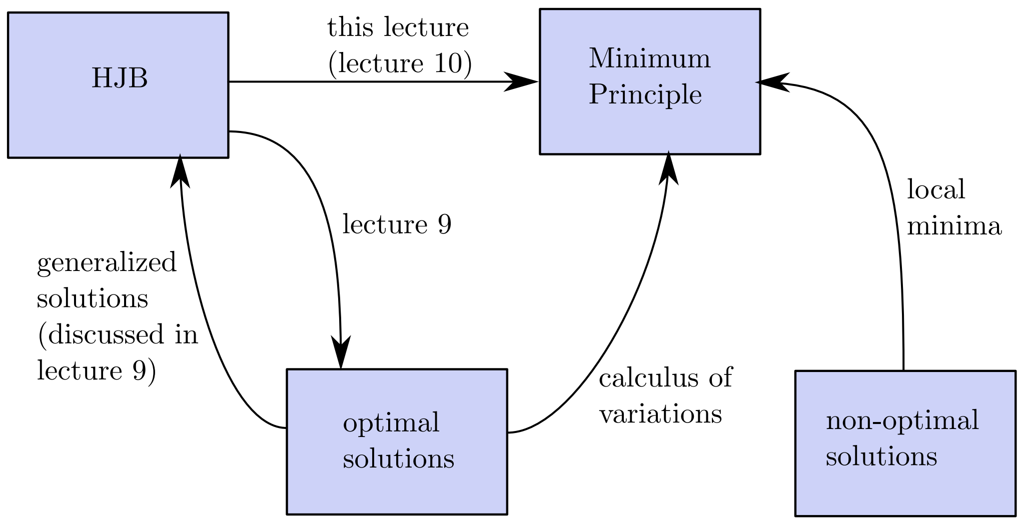

Figure 10.2: Optimal solutions and their relation to the HJB and the Minimum Principle.