1.4 Day 4: 弱水三千取一瓢

大自然说它想最小化 $S=-\frac{1}{4}\int F_{\mu\nu}F^{\mu\nu}\mathrm{d}x$，于是就有了光。众所周知，从学习理论力学这门学科开始，我们“给出一个系统动力学方程”的方法就发生了变化：相比于直接写出 ODEs，我们更倾向于找到在系统对称变换下保持不变的，系统运动轨迹 $x(t),t\in[t_0,t_f]$ 的标量泛函，“最小化该泛函”这一变分问题的解就是系统的动力学方程。Day 2 和 Day 3 中已经讨论了分析 LTI 系统性态的各种方法，现在我们将讨论泛函求极值的相关内容。本节中我们将抛弃通过权衡极点位置、超调量、稳定时间等指标来粗略设计控制器的方法，而是直接计算使得某性能指标最小化的开环/闭环控制律。

【阅读提示】为了偷懒，本节涉及矩阵与向量、向量与向量的乘法时，大量使用了 $\cdot$ 符号。阅读时请根据维数和上下文自行脑补哪个向量/矩阵需要转置 ^{T}。

1.4.1 控制律无约束的最优控制：乘子法和标准变分法

首先考虑一类微分约束的泛函极值问题：设有连续可微的状态变量 $x(t):[t_0,t_f]\to\mathbb{R}^N$，满足 $x(t_0)=x_0,x(t_f)=x_f$，满足约束 $F(x(t),\dot{x}(t),t)=0$，希望最小化性能指标 $J[x]=\int_{t_0}^{t_f}g(x,\dot{x},t)\mathrm{d}t$。对于这样的问题，我们可以直接使用拉格朗日乘子法处理微分约束，即构造增广的性能指标：

$$\bar{J}[x,p]=\int_{t_0}^{t_f}[g(x,\dot{x},t)+p(t)\cdot F(x,\dot{x},t)]\mathrm{d}t$$

为什么可以直接这样做?是因为这个泛函极值问题可以视作无穷维优化问题,不妨将时间等分作 $t_0,t_1,\cdots,t_{N-1},t_N=t_f$，则该问题的决策变量变成在各个分点上的 $x(t)$ 值 $x_0=x(t_0),x_1=x(t_1),\cdots,x_N=x(t_N)$，$\dot{x}_i=\dot{x}(t_i)$ 可以使用差分逼近，即取 $\dot{x}_i=\dfrac{x_{i+1}-x_i}{\Delta t},i=0,1,\cdots N-1,\Delta t=t_{i+1}-t_i$，这样 $\dot{x}_N$ 是未定义的，但是你完全可以直接取 $\dot{x}_N=0$，随着 $\Delta t\to 0$，在有限个点上对 $\dot{x}_i$ 的不正确逼近对性能指标/目标函数带来的影响将消失。离散化后，我们得到优化问题：

$$\min J(\{x_i\})=\sum_{i=0}^{N-1}g(x_i,\dot{x}_i,t_i)\Delta t,\quad \text{s.t.}F(x_i,\dot{x}_i,t_i)=0\ \ i=0,1,\cdots,N$$

我使用每个区间的左端点乘以区间长度来逼近积分结果，此时 $x_N,\dot{x}_N$ 根本不出现在性能指标中，真正在起效的决策变量只有 $x_0,x_1,\cdots,x_{N-1}$，真正在起效的约束也只有 $N+1$ 个中的前 N 个。我们已知这个有限维的优化问题可以使用拉格朗日乘子法解决：引入乘子 $p_0,p_1,\cdots,p_{N-1}$，令增广的拉格朗日函数为：

$$\bar{J}(\{x_i\},\{p_i\})=\sum_{i=0}^{N-1}[g(x_i,\dot{x}_i,t_i)+p_i\cdot F(x_i,\dot{x}_i,t_i)]\Delta t$$

令 $N\to\infty$ 或者说 $\Delta t=0$，这个增广拉格朗日函数就变成上面的增广拉格朗日泛函 $\bar{J}[x,p]$。要使得增广泛函取极小值，应满足关于 $x(t)$ 和 $p(t)$ 的 E-L 方程，令 $\bar{g}(x,\dot{x},p,t)=g(x(t),\dot{x}(t),t)+p(t)\cdot F(x(t),\dot{x}(t),t)$，E-L 方程给出：

$$0=\frac{\partial\bar{g}(x,\dot{x},p,t)}{\partial x}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\bar{g}(x,\dot{x},p,t)}{\partial\dot{x}}\right),\quad 0=F(x,\dot{x},t)\tag{1.47}$$

从离散化后的有限维优化问题中也可读出这个结果。令 $\bar{g}(x_i,\dot{x}_i,p_i,t_i)=g(x_i,\dot{x}_i,t_i)+p_i\cdot F(x_i,\dot{x}_i,t_i)$，则：

$$\begin{aligned}\frac{\partial J(\{x_i\},\{p_i\})}{\partial x_i}&=\frac{\partial\bar{g}(x_i,\dot{x}_i,p_i,t_i)}{\partial x_i}+\frac{\partial\bar{g}(x_i,\dot{x}_i,p_i,t_i)}{\partial x_{i-1}}\\&=\frac{\partial\bar{g}(x_i,\dot{x}_i,p_i,t_i)}{\partial x_i}+\frac{\partial\bar{g}(x_i,\dot{x}_i,p_i,t_i)}{\partial\dot{x}_i}\frac{\partial\dot{x}_i}{\partial x_i}+\frac{\partial\bar{g}(x_{i-1},\dot{x}_{i-1},p_{i-1},t_{i-1})}{\partial\dot{x}_{i-1}}\frac{\partial\dot{x}_{i-1}}{\partial x_{i-1}}\\&=\frac{\partial\bar{g}(x_i,\dot{x}_i,p_i,t_i)}{\partial x_i}-\frac{1}{\Delta t}\left(\frac{\partial\bar{g}(x_i,\dot{x}_i,p_i,t_i)}{\partial\dot{x}_i}-\frac{\partial\bar{g}(x_{i-1},\dot{x}_{i-1},p_{i-1},t_{i-1})}{\partial\dot{x}_{i-1}}\right)=0\\\frac{\partial J(\{x_i\},\{p_i\})}{\partial p_i}&=F(x_i,\dot{x}_i,t_i)=0\end{aligned}$$

这样，我们为直接使用拉格朗日乘子法求解最优控制问题提供了合法性。接下来我们看最简单的最优控制问题：考虑状态变量 $x(t):[t_0,t_f]\to\mathbb{R}^N$，控制变量 $u(t):[t_0,t_f]\to\mathbb{R}^M$ 均无约束。状态变量满足边界条件 $x(t_0)=$

$x_0, x(t_f) = x_f$，系统的动力学方程 $\dot{x}(t) = f(x(t), u(t), t)$，性能指标 $J[u] = \int_{t_0}^{t_f} g(x(t), u(t), t)\mathrm{d}t$，方便起见引入 $\mathcal{H}(x(t), u(t), p(t), t) = g(x(t), u(t), t) + p(t) \cdot f(x(t), u(t), t)$（我们暂且将其称为哈密顿量，下文中将反复使用），可得以下必要条件：

$$0 = \frac{\partial \mathcal{H}(x,u,p,t)}{\partial u}, \quad \dot{x}(t) = +\frac{\partial \mathcal{H}(x,u,p,t)}{\partial p}, \quad \dot{p}(t) = -\frac{\partial \mathcal{H}(x,u,p,t)}{\partial x} \tag{1.48}$$

很多问题可以转换成有微分约束的泛函极值问题处理，例如一类有积分约束的极值问题（e.g. 等周问题）：有状态变量 $x(t) : [t_0, t_d] \to \mathbb{R}^n$，边界条件 $x(t_0) = x_0, x(t_f) = x_f$，有约束 $\int_{t_0}^{t_f} b(x, \dot{x}, t)\mathrm{d}t = B$，仍然最小化 $J[x] = \int_{t_0}^{t_f} g(x, \dot{x}, t)\mathrm{d}t$，显然我们需要将积分约束换成微分约束，引入新的变量：

$$z(t) = \int_{t_0}^{t} b(x(\tau), \dot{x}(\tau), \tau)\mathrm{d}\tau \Rightarrow \dot{z}(t) = b(x(t), \dot{x}(t), t)$$

从而该问题立刻转换为有微分约束的问题，可使用上述方法处理。注意到此时 $\mathcal{H}$ 中不显含 $\dot{z}$ 的，因此与 z 对应的拉格朗日乘子 $p(t) = \text{Const} = p_0$。

下面考虑终端状态不固定的情形，也就是说我们不再把终端状态限制在一点 $x(t_f) = x_f$，而是允许系统落在某个目标集 $\mathcal{S} = \{x(t_f) : m(x(t_f), t_f) = 0\}$ 中。系统的性能指标为 $J(x) = h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x, \dot{x}, t)\mathrm{d}t$，我们可以直接将对终端状态和时间的约束使用拉格朗日乘子引入，即令 $\bar{J} = J + \lambda \cdot m(x(t_f), t_f)$，对 $\bar{J}$ 进行变分，注意此时对终端时间和终端状态也要变分：

$$\begin{aligned}
\delta \bar{J} &= \frac{\partial h}{\partial x} \cdot \delta x_f + \frac{\partial h}{\partial t}\delta t_f + m \cdot \delta\lambda + \lambda \cdot \left[\frac{\partial m}{\partial x} \cdot \delta x_f + \frac{\partial m}{\partial t}\delta t_f\right] + \int_{t_0}^{t_f} \left[\frac{\partial g}{\partial x} \cdot \delta x(t) + \frac{\partial g}{\partial \dot{x}}\delta\dot{x}\right]\mathrm{d}t + g\delta t_f \\
&= \left[g + \frac{\partial h}{\partial t} + \frac{\partial m}{\partial t}\right]\delta t_f + \left[\frac{\partial h}{\partial x} + \lambda \cdot \frac{\partial m}{\partial x}\right] \cdot \delta x_f + m \cdot \delta\lambda + \int_{t_0}^{t_f} \left[\frac{\partial g}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial g}{\partial \dot{x}}\right)\right] \cdot \delta x(t)\mathrm{d}t + \frac{\partial g}{\partial \dot{x}} \cdot \delta x(t_f) \\
&= \left[\frac{\partial h}{\partial t} + \lambda \cdot \frac{\partial m}{\partial t} + g - \frac{\partial g}{\partial \dot{x}} \cdot \dot{x}\right]\delta t_f + \left[\frac{\partial h}{\partial x} + \lambda \cdot \frac{\partial m}{\partial x} + \frac{\partial g}{\partial \dot{x}}\right] \cdot \delta x_f + \int_{t_0}^{t_f} \left[\frac{\partial g}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial g}{\partial \dot{x}}\right)\right] \cdot \delta x(t)\mathrm{d}t
\end{aligned}$$

在上面的推导中，我们已经利用了关系 $\delta x_f = \delta x(t_f) + \dot{x}(t_f)\delta t_f$，可以立刻读出三个极值条件：除了 E-L 方程外，剩下的两个是：

$$\frac{\partial h}{\partial t} + \lambda \cdot \frac{\partial m}{\partial t} + g - \frac{\partial g}{\partial \dot{x}} \cdot \dot{x} = 0 \quad \frac{\partial h}{\partial x} + \lambda \cdot \frac{\partial m}{\partial x} + \frac{\partial g}{\partial \dot{x}} = 0$$

通过这种对终端时刻和终端状态变分的方式，我们可求解有一般目标集的最简最优控制问题：仍考虑状态变量 $x(t) : [t_0, t_f] \to \mathbb{R}^N$、控制变量 $u(t) : [t_0, t_f] \to \mathbb{R}^M$，状态变量有初值条件 $x(t_0) = x_0$，和终端时刻的目标集 $\mathcal{S} = \{x(t_f) : m(x(t_f), t_f) = 0\}$，动力学方程 $\dot{x}(t) = f(x(t), u(t), t)$，性能指标 $J[u] = h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x, u, t)\mathrm{d}t$，使用拉格朗日乘子引入约束（包括动力学方程和终端时刻目标集），对轨迹 $x(t)$，终端时间 t_f，终端状态 x_f 变分，可获得最优控制的必要条件，除了方程1.48，还将得到三个边界条件：

$$0 = \left[\frac{\partial \bar{h}(x(t_f), t_f, \lambda)}{\partial x} - p(t_f)\right] \quad 0 = \left[\frac{\partial \bar{h}(x(t_f), t_f, \lambda)}{\partial t} + \mathcal{H}(x(t_f), u(t_f), t_f)\right] \quad 0 = m(x(t_f), t_f) \tag{1.49}$$

其中 $\bar{h}(x(t_f), t_f, \lambda) = h(x(t_f), t_f) + \lambda \cdot m(x(t_f), t_f)$

下面考虑有内点约束的情形。例如，我们希望控制一小车在起点和终点间进行折返跑，此时我们必须引入“在 $[t_0, t_f]$ 中某一时刻，小车到达终点”这个约束。我们考虑形如 $\psi(x(t_1), t_1) = 0$ 的约束，同样将约束使用拉格朗日乘子法添加到性能指标中：

$$\begin{aligned}
\bar{J} &= h(x(t_f), t_f) + \lambda \cdot \psi(x(t_1), t_1) + \int_{t_0}^{t_f} \left[g(x, u, t) + p(t) \cdot (f(x, u, t) - \dot{x})\right]\mathrm{d}t \\
&:= h(x(t_f), t_f) + \lambda \cdot \psi(x(t_1), t_1) + \int_{t_0}^{t_f} \left[\mathcal{H}(x, u, p, t) - p \cdot \dot{x}\right]\mathrm{d}t
\end{aligned}$$

在时间点 t_1 前后，可能有不同的控制策略 $u(t)$，因此 $\mathcal{H}$ 在 t_1 两侧可能是不连续的（假设 $x(t)$ 在 t_1 处是连续的）。

将 $t_1+\delta t_1$ 处的轨迹变化记为 $\delta\xi$，对上面的性能指标进行变分：

$$\begin{aligned}\delta\bar{J} &= \delta\bar{J}_1(\delta x_f)+\delta\bar{J}_2(\delta t_f)+\delta\lambda\cdot\psi(x(t_1),t_1)+\lambda\cdot\frac{\partial\psi}{\partial x}\cdot\delta\xi+\lambda\cdot\frac{\partial\psi}{\partial t_1}\delta t_1\\ &+\int_{t_0}^{t_f}\left[\frac{\partial\mathcal{H}}{\partial x}\cdot\delta x+\frac{\partial\mathcal{H}}{\partial u}\cdot\delta u+\frac{\partial\mathcal{H}}{\partial p}\cdot\delta p-\delta p\cdot\dot{x}-p\cdot\delta\dot{x}\right]\\ &=\delta\bar{J}_1(\delta x_f)+\delta\bar{J}_2(\delta t_f)+\delta\lambda\cdot\psi(x(t_1),t_1)+\lambda\cdot\frac{\partial\psi}{\partial x}\cdot\delta\xi+\lambda\cdot\frac{\partial\psi}{\partial t_1}\delta t_1\\ &+\int_{t_0}^{t_1}\left[\left(\frac{\partial\mathcal{H}}{\partial x}+\dot{p}\right)\cdot\delta x+\frac{\partial\mathcal{H}}{\partial u}\cdot\delta u+\left(\frac{\partial\mathcal{H}}{\partial p}-\dot{x}\right)\cdot\delta p\right]-[p(t_1)\cdot\delta x(t_1)]|_-\\ &+\int_{t_1}^{t_f}\left[\left(\frac{\partial\mathcal{H}}{\partial x}+\dot{p}\right)\cdot\delta x+\frac{\partial\mathcal{H}}{\partial u}\cdot\delta u+\left(\frac{\partial\mathcal{H}}{\partial p}-\dot{x}\right)\cdot\delta p\right]+[p(t_1)\cdot\delta x(t_1)]|_+\\ &+[H(x,u,p,t)-p\cdot\dot{x}]\,|_-\delta t_1-[H(x,u,p,t)-p\cdot\dot{x}]\,|_+\delta t_1\end{aligned}$$

这里，正比于 $\delta x_f,\delta t_f$ 的项被省略，分别记为 $\delta J_1(\delta x_f)$ 和 $\delta J_2(\delta t_f)$，所有标记为 $-$ 的项是在 t_1 左侧的项；所有标记为 $+$ 的项是在 t_1 右侧的项。注意到 $\delta x(t_1)|_+$、$\delta x(t_1)|_-$ 与 $\delta\xi$ 并非是独立的变分，因此需利用关系：

$$\delta\xi=\delta x(t_1)|_-+\dot{x}(t_1)|_-\delta t=\delta x(t_1)|_++\dot{x}(t_1)|_+\delta t$$

（这个关系的正确性在图1.4中简要说明）继续化简变分，将它们都变成 $\delta\xi$：

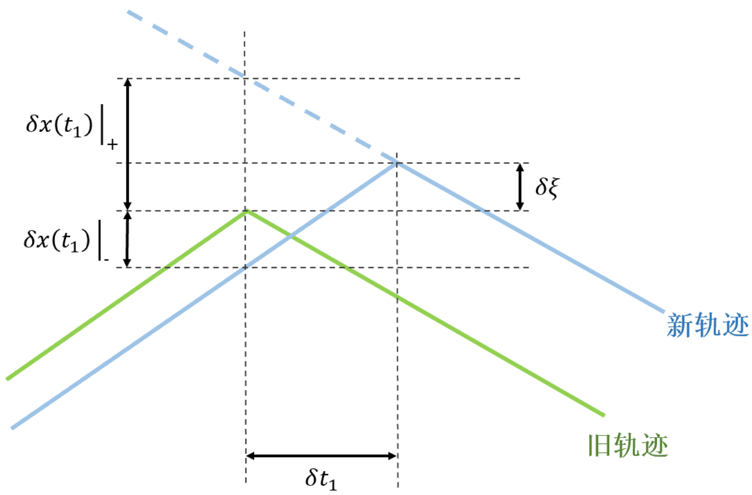

图 1.4: $\delta\xi$ 与 $\delta x(t_1)|^+,\delta x(t_1)|^-$ 的关系

$$\delta\bar{J}=\delta\bar{J}_1+\delta\bar{J}_2+\int_{t_0}^{t_1}\cdots\mathrm{d}t+\int_{t_1}^{t_f}\cdots\mathrm{d}t+\delta\lambda\cdot\psi+\left(\lambda\cdot\frac{\partial\psi}{\partial t_1}+\mathcal{H}|_--\mathcal{H}|_+\right)\delta t+\left(\lambda\cdot\frac{\partial\psi}{\partial x}-p|_-+p|_+\right)\cdot\delta\xi$$

从中读出三个由内点约束带来的额外的条件：

$$\lambda\cdot\frac{\partial\psi}{\partial t_1}+\mathcal{H}|_--\mathcal{H}|_+=0\quad\lambda\cdot\frac{\partial\psi}{\partial x}-p|_-+p|_+=0\quad\psi(x(t_1),t_1)=0$$

1.4.2 控制律有约束的最优控制：庞特里亚金极值原理

然而，我们知道，人或者发动机等装置能向系统内输入的力是有限制的，因此上述控制量 $u(t)$ 无约束的最优控制问题的解往往难以在现实中应用。例如，考虑“在最短时间内，将一个小物块从 $x(0)=1,\dot{x}(0)=0$ 状态转移到 $x(t_f)=0,\dot{x}(t_f)=0$ 状态”这一问题，只要我能输入尽可能大的力，这个过程就能在尽可能短的时间内被完成，这样的解显然是现实中行不通的。因此，我们有必要将 $u(t)$ 也限制在一个可行集合中。那么，$\delta u(t)$ 将不能随意取值，这导致了上文中标准变分法的失效。

我们将从最简单的问题开始研究：考虑系统状态 $x(t):[t_0,t_f]\to\mathbb{R}^N$，控制变量 $u(t):[t_0,t_f]\to\mathbb{R}^M$，且要求 $u(t)\in U,\forall t$，最小化一个只有终端代价的性能指标 $J[u]=h(x(t_f))$，终端时刻和终端状态均自由。现在声明：

最优控制 $u(t)$ 应满足如下条件:

- 对任意 $u'(t) \in U$，在几乎任意时刻 $t \in [t_0, t_f]$，$\mathcal{H}(x(t), u(t), p(t)) \leq \mathcal{H}(x(t), u'(t), p(t))$
- 状态变量 $x(t)$ 和协态变量(一组与状态变量“对偶”的变量,类似于上面的拉格朗日乘子)$p(t) : [t_0, t_f] \to \mathbb{R}^N$ 应满足正则方程组 $\dot{x}(t) = \dfrac{\partial \mathcal{H}}{\partial p}, \dot{p}(t) = -\dfrac{\partial \mathcal{H}}{\partial x}$
- 满足边界条件 $\dfrac{\partial h(x(t_f))}{\partial x} - p(t_f) = 0, \mathcal{H}(x(t_f), u(t_f), p(t_f)) = 0$
- $\mathcal{H}(x, u, p) = 0, \forall t \in [t_0, t_f]$

类比前文中控制量无约束的问题，此处只有终端代价，故哈密顿量 $\mathcal{H} = p(t) \cdot f(x(t), u(t))$。下面给出证明：设最优控制 $u(t)$，对应的状态 $x(t)$, 终端时刻 t_f。现在对最优控制施加扰动，变成 $u'(t) = u(t) + \delta u(t)$，对终端时刻施加扰动，变成 $t'_f = t_f + \delta t_f$，计算性能指标的增量:

$$
\begin{aligned}
\Delta J &= J[u'] - J[u] = h(x'(t'_f)) - h(x(t_f)) \\
&= \left[h(x'(t'_f)) - h(x(t'_f))\right] - \left[h(x(t'_f)) - h(x(t_f))\right] \\
&\approx \frac{\partial h}{\partial x}(x(t'_f)) \cdot \left[x'(t'_f) - x(t'_f)\right] + \frac{\partial h}{\partial x}(x(t_f)) \cdot \dot{x}(t_f)[t'_f - t_f] \\
&= \frac{\partial h}{\partial x}(x(t'_f)) \cdot \delta x(t'_f) + \frac{\partial h}{\partial x}(x(t_f)) \cdot f(x(t_f), u(t_f))\delta t_f \\
&= \left[\frac{\partial h}{\partial x}(x(t'_f)) - \frac{\partial h}{\partial x}(x(t_f))\right] \cdot \delta x(t_f) + \frac{\partial h}{\partial x}(x(t_f)) \cdot \delta x(t_f) + \frac{\partial h}{\partial x}(x(t_f)) \cdot f(x(t_f), u(t_f))\delta t_f \\
&\approx \frac{\mathrm{d}}{\mathrm{dt}}\left[\frac{\partial h}{\partial x}(x(t_f))\right] \cdot \delta x(t_f)\delta t_f + \frac{\partial h}{\partial x}(x(t_f)) \cdot \delta x(t_f) + \frac{\partial h}{\partial x}(x(t_f)) \cdot f(x(t_f), u(t_f))\delta t_f \\
&\approx \frac{\partial h}{\partial x}(x(t_f)) \cdot \delta x(t_f) + \frac{\partial h}{\partial x}(x(t_f)) \cdot f(x(t_f), u(t_f))\delta t_f
\end{aligned}
$$

也就是说我们将代价的变化拆成了两部分：t_f 时刻状态变化带来的代价变化，和终端时间变化带来的代价变化。首先考虑 t_f 时刻状态不变，单纯让终端时间变化带来的影响，令 $\delta \bar{J} = 0$ 得到:

$$
\frac{\partial h}{\partial x}(x(t_f)) \cdot f(x(t_f), u(t_f))\delta t_f \geq 0 \Rightarrow \frac{\partial h}{\partial x}(x(t_f)) \cdot f(x(t_f), u(t_f)) = 0 \tag{1.50}
$$

接着考虑终端时间不变，只有 t_f 时刻状态变化的情形，此时显然需要有:

$$
\frac{\partial h}{\partial x}(x(t_f)) \cdot \delta x(t_f) \geq 0
$$

我们需要分别求出 $\dfrac{\partial h(x(t_f))}{\partial x(t_f)}$ 和 $\delta x(t_f)$ 两部分，首先看 $\delta x(t)$ 随着时间的演化:

$$
\begin{aligned}
\delta \dot{x}(t) &\approx \frac{\partial f}{\partial x}(x(t), u(t) + \delta u(t)) \cdot \delta x(t) + [f(x(t), u(t) + \delta u(t)) - f(x(t), u(t))] \\
&= \frac{\partial f}{\partial x}(x(t), u(t)) \cdot \delta x(t) + \left[\frac{\partial f}{\partial x}(x(t), u(t) + \delta u(t)) - \frac{\partial f}{\partial x}(x(t), u(t))\right] \cdot \delta x(t) + [f(x(t), u(t) + \delta u(t)) - f(x(t), u(t))]
\end{aligned}
$$

为了求出 $\dfrac{\partial h(x(t_f))}{\partial x(t_f)}$，我们决定追踪 $p(t) = \dfrac{\partial h(x(t_f))}{\partial x(t)}$ 的演化，我们需要导出其演化方程，考虑:

$$
\frac{\partial h(x(t_f))}{\partial x(t)} = \frac{\partial h(x(t_f))}{\partial x(t + \Delta t)} \cdot \frac{\partial x(t + \Delta t)}{\partial x(t)}
$$

那么:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{dt}}\left[\frac{\partial h(x(t_f))}{\partial x(t)}\right] &\approx \frac{1}{\Delta t}\left[\frac{\partial h(x(t_f))}{\partial x(t + \Delta t)} - \frac{\partial h(x(t_f))}{\partial x(t + \Delta t)} \cdot \frac{\partial x(t + \Delta t)}{\partial x(t)}\right] = \frac{\partial h(x(t_f))}{\partial x(t + \Delta t)} \frac{1}{\Delta t}\left[1 - \frac{\partial x(t + \Delta t)}{\partial x(t)}\right] \\
&\approx \frac{\partial h(x(t_f))}{\partial x(t)} \frac{1}{\Delta t}\left[1 - \frac{\partial x(t + \Delta t)}{\partial x(t)}\right]
\end{aligned}
$$

而

$$
x(t + \Delta t) \approx x(t) + f(x(t), u(t), t)\Delta t \Rightarrow \frac{\partial x(t + \Delta t)}{\partial x(t)} \approx 1 + \frac{\partial f(x(t), u(t), t)}{\partial x(t)}\Delta t
$$

从而立刻得到 $p(t)=\dfrac{\partial h(x(t_f))}{\partial x(t)}$ 的演化方程：

$$\frac{\mathrm{d}}{\mathrm{d}t}p(t)=-\frac{\partial f(x(t),u(t),t)}{\partial x}\cdot p(t)=-\frac{\partial \mathcal{H}(x(t),u(t),p(t))}{\partial x} \tag{1.51}$$

配合上面求出的 $\delta\dot{x}(t)$，得到：

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[p(t)\cdot\delta x(t)\right]=p(t)\cdot\left[\frac{\partial f}{\partial x}(x(t),u(t)+\delta u(t))-\frac{\partial f}{\partial x}(x(t),u(t))\right]\cdot\delta x(t)+p(t)\cdot\left[f(x(t),u(t)+\delta u(t))-f(x(t),u(t))\right]$$

两侧积分得到：

$$p(t_f)\cdot\delta x(t_f)=\int_{t_0}^{t_f}\left\{p(t)\cdot\left[\frac{\partial f}{\partial x}(x(t),u(t)+\delta u(t))-\frac{\partial f}{\partial x}(x(t),u(t))\right]\cdot\delta x(t)+\mathcal{H}(x,u+\delta u,p)-\mathcal{H}(x,u,p)\right\}\mathrm{d}t$$

由于 $u(t)$ 要落在可行集合内，因此我们不能任意加上 $\delta u(t)$，引入所谓“针状变分”：

$$\delta u(t)=\begin{cases}\omega-u(t) & t\in[t_2,t_3],\omega\in U\\ 0 & \text{otherwise}\end{cases}$$

这意味着我们只在 $[t_2,t_3]$ 这一小段将控制律修改为 ω。那么，上面对 $[t_0,t_f]$ 的积分立刻转化为对 $[t_2,t_3=t_2+\Delta t]$ 的积分。设 $f(x,u)$ 对两个槽位都是 Lipschitz 连续的，系数分别为 $\gamma_f,\rho_f>0$，则可证明由控制律扰动带来的状态扰动：

$$\|\delta x(t)\|\leq\exp(\gamma_f t)\int_{t_0}^{t}\exp(-\gamma_f\tau)\rho_f\|\delta u(\tau)\|\mathrm{d}\tau$$

因此，只要把针状变分的区间长度 Δt 设置得足够小，上面 $p(t_f)\cdot\delta x(t_f)$ 中，积分号内的 $p(t)\cdot\left[\dfrac{\partial f}{\partial x}(x,u+\delta u)-\dfrac{\partial f}{\partial x}(x,u)\right]\cdot\delta x$ 就是比后面的部分更高一阶的小量，可以忽略。利用最优控制的条件 $p(t_f)\delta x(t_f)\geq 0$ 得到：

$$\int_{t_2}^{t_2+\Delta t}\left[\mathcal{H}(x(t),u(t)+\delta u(t),p(t))-\mathcal{H}(x(t),u(t),p(t))\right]\mathrm{d}t\geq 0$$

这意味着，对任意 $u_0\in U$，有：

$$\int_{t_2}^{t_2+\Delta t}\left[\mathcal{H}(x(t),u_0,p(t))-\mathcal{H}(x(t),u(t),p(t))\right]\mathrm{d}t\geq 0$$

由积分中值定理，存在 $\tau\in[t_2,t_2+\Delta t]$，使得：

$$\left[\mathcal{H}(x(\tau),u'(\tau),p(\tau))-\mathcal{H}(x(\tau),u(\tau),p(\tau))\right]\geq 0,\forall u'(\tau)\in U \tag{1.52}$$

由 $\mathcal{H}$ 的定义立刻知道 $\mathcal{H}(x(t_f),u(t_f),p(t_f))=0$，下面证明它在系统运行过程中恒为 0。首先证明它连续。考虑 $u(t)$ 的不连续点左侧的任一连续点 $t_1-\epsilon$，取任一容许控制 $u'(t_1-\epsilon)$，最优控制仍记为 $u(t_1-\epsilon)$，那么：

$$\mathcal{H}(x(t_1-\epsilon),u(t_1-\epsilon),p(t_1-\epsilon))\leq\mathcal{H}(x(t_1-\epsilon),u'(t_1-\epsilon),p(t_1-\epsilon))$$

取 $u'(t_1-\epsilon)=u(t_1+\epsilon)$，并令 $\epsilon\to 0$，立刻得到：

$$\mathcal{H}(x(t_1),u(t_1),p(t_1)|_-\leq\mathcal{H}(x(t_1),u(t_1),p(t_1)|_+$$

同样的操作将得到：

$$\mathcal{H}(x(t_1),u(t_1),p(t_1)|_+\leq\mathcal{H}(x(t_1),u(t_1),p(t_1)|_-$$

立刻知道 $\mathcal{H}$ 是对 t 连续的。考虑 $\mathcal{G}(x(t),p(t))$ 是最优控制轨迹上的 $\mathcal{H}$，即 $\mathcal{G}(x(t),p(t))=\min_{u'(t)}\mathcal{H}(x(t),u'(t),p(t))$。考虑任意区间 $[t_2,t_3]$，由极值条件：

$$\mathcal{H}(x(t_2),u(t_3),p(t_2))\geq\mathcal{H}(x(t_2),u(t_2),p(t_2))=\mathcal{G}(x(t_2),p(t_2))$$

以及

$$\mathcal{H}(x(t_3),u(t_2),p(t_3))\geq\mathcal{H}(x(t_3),u(t_3),p(t_3))=\mathcal{G}(x(t_3),p(t_3))$$

从而有不等式：

$$\mathcal{H}(x(t_3),u(t_3),p(t_3))-\mathcal{H}(x(t_2),u(t_3),p(t_2))\leq\mathcal{G}(x(t_3),p(t_3))-\mathcal{G}(x(t_2),p(t_2))\leq\mathcal{H}(x(t_3),u(t_2),p(t_3))-\mathcal{H}(x(t_2),u(t_2),p(t_2))$$

看这个不等式的两边，以左侧为例：

$$\begin{aligned}&\lim_{\Delta t\to 0}\frac{1}{\Delta t}\{\mathcal{H}(x(t_3),u(t_3),p(t_3))-\mathcal{H}(x(t_2),u(t_3),p(t_2))\}\\&\approx\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{-[x(t_2)-x(t_3)]\cdot\frac{\partial\mathcal{H}}{\partial x}(x(t_3),u(t_3),p(t_3))-[p(t_2)-p(t_3)]\cdot\frac{\partial\mathcal{H}}{\partial p}(x(t_3),u(t_3),p(t_3))\right\}\\&\approx\dot{x}\cdot\frac{\partial\mathcal{H}}{\partial x}(x(t_3),u(t_3),p(t_3))+\dot{p}\cdot\frac{\partial\mathcal{H}}{\partial p}(x(t_3),u(t_3),p(t_3))=0\end{aligned}$$

等式右侧可做类似处理。由不等式夹逼得到 $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{G}(x(t),u(t))=0$，从而在最优控制下，$\mathcal{H}(x(t),u(t),p(t))=0$ 恒成立。至此，我们证明了最优控制时，$x(t),u(t),p(t)$ 应满足的所有条件，通过利用这些条件给出最优控制的方式称为“庞特里亚金极值原理”。在这个证明中我们主要做了两件比较不平凡的事情：1）将标准变分换成针状变分，相较于标准变分“在整个路径上都有变化，但是变化无限小”的操作，针状变分的做法是“仅在一段时间上有变化，但变化的时间无穷短”；2）我们追踪了 $p(t)=\frac{\partial h(x(t_f))}{\partial x(t)}$ 的变化，也就是说，协态变量的意义是“影子价格”，它反映了 t 时刻轨迹稍稍变化，对性能指标的影响。

下面我们将给出其他更复杂情况下的极值原理，我们将尝试把更复杂的问题归约到之前求解过的简单问题。首先修改性能指标的形式，在性能指标中加入运行代价：$J[u]=h(x(t_f))+\int_{t_0}^{t_f}g(x(t),u(t))\mathrm{d}t$，为此，引入新的状态变量 $\dot{x}_{N+1}(t)=g(x(t),u(t))$，性能指标就写成 $J=h(x(t_f))+x_{N+1}(t_f)$。利用前面已经得到的极值条件求解这个新问题。记 $\bar{x}(t)=[x^T(t),x_{N+1}(t)]^T,\bar{p}(t)=[p^T(t),p_{N+1}(t)]^T$，此时的哈密顿量应当是：

$$\mathcal{H}(\bar{x}(t),u(t),\bar{p}(t))=p(t)\cdot f(x(t),u(t))+p_{N+1}(t)g(x(t),u(t))$$

由协态变量满足的方程：

$$\dot{p}(t)=-\left[\frac{\partial f}{\partial x}(x(t),u(t))\right]\cdot p(t)-p_{N+1}\frac{\partial g}{\partial x}(x(t),u(t)),\quad \dot{p}_{N+1}(t)=0$$

由 t_f 时刻边界条件：

$$\frac{\partial h}{\partial x}(x(t_f))-p(t_f)=0,\quad 1-p_{N+1}(t_f)=0$$

可以看出，新加入的 $p_{N+1}(t)=1$ 恒成立，因此我只需要将 $\mathcal{H}$ 设置为

$$\mathcal{H}(x(t),p(t),u(t))=p(t)\cdot f(x(t),u(t))+g(x(t),u(t))\tag{1.53}$$

就可以直接使用上文中仅有终端代价问题的四个极值条件了。下面我们再考虑动力学方程和性能指标都可以显含时间的情况，即 $\dot{x}(t)=f(x(t),u(t),t)$，$J[u]=h(x(t_f),t_f)+\int_{t_0}^{t_f}g(x(t),u(t),t)\mathrm{d}t$，这可以通过引入新的状态变量 $\dot{x}_{N+1}(t)=1,x_{N+1}(t_0)=t_0$ 来将此问题归约到刚才求解过的问题（性能指标中含有运行代价，但性能指标和动力学方程均不显含时间），此时的哈密顿量应当是：

$$\bar{\mathcal{H}}(\bar{x}(t),u(t),\bar{p}(t))=g(x(t),u(t),t)+p(t)\cdot f(x(t),u(t),t)+p_{N+1}(1)\cdot 1$$

为了和上面性能指标、动力学方程不显含 t 的情形下 $\mathcal{H}$ 的形式看起来一致，我们不希望最终的结论中出现 $\bar{\mathcal{H}}$，而要出现 $\mathcal{H}(x(t),u(t),p(t),t)$，它定义为：

$$\bar{\mathcal{H}}(\bar{x}(t),u(t),\bar{p}(t):=\mathcal{H}(x(t),u(t),p(t),t)+p_{N+1}(t)\Rightarrow\mathcal{H}(x(t),u(t),p(t),t)=g(x(t),u(t),t)+p(t)\cdot f(x(t),u(t),t)$$

由 $\bar{\mathcal{H}}$ 的极值条件 $\bar{\mathcal{H}}(\bar{x}(t),u(t),\bar{p}(t))\le\bar{\mathcal{H}}(\bar{x}(t),u'(t),\bar{p}(t)),\forall u'(t)\in U$，可知 $\mathcal{H}(x(t),u(t),p(t),t)\le\mathcal{H}(x(t),u'(t),p(t),t)$，$\forall u'(t)\in U$，这是关于 $\mathcal{H}$ 的极值条件。考虑原有的状态变量、协态变量的演化：

$$\dot{x}(t)=\frac{\partial\bar{\mathcal{H}}(\bar{x},u,p)}{\partial p}=\frac{\partial\mathcal{H}(x,u,p,t)}{\partial p}\quad \dot{p}(t)=-\frac{\partial\bar{\mathcal{H}}(\bar{x},u,p)}{\partial x}=\frac{\partial\mathcal{H}(x,u,p,t)}{\partial x}\tag{1.54}$$

利用边界条件：

$$0=\frac{\partial h}{\partial x}(x(t_f),t_f)-p(t_f)\quad 0=\frac{\partial h}{\partial t}(x(t_f),t_f)-p_{N+1}(t_f)\tag{1.55}$$

以及

$$0=\bar{\mathcal{H}}(\bar{x}(t_f),u(t_f),\bar{p}(t_f))=\mathcal{H}(x(t_f),u(t_f),p(t_f),t_f)+p_{N+1}(t_f)$$

将方程1.55中的第二式代入上式，得到：

$$0=\mathcal{H}(x(t_f),u(t_f),p(t_f),t_f)+\frac{\partial h}{\partial t}(x(t_f),t_f) \tag{1.56}$$

我们看到这样的非自治系统的最优控制律和状态演化仍然由 $\mathcal{H}$ 对 $u(t)$ 的极值条件，状态变量、协态变量满足的正则方程组（方程1.54）以及边界条件（方程1.55，1.56）给出。相比于之前不含时的版本，我们失去了 $\mathcal{H}(x(t),u(t),p(t),t)=0$ 恒成立这个性质。最后我们考虑系统的终端时刻、终端状态受约束的情形。设系统必须落在目标集 $S=\{x(t_f):m(x(t_f),t_f)=0\}$ 中，仍然使用乘子法将约束放到性能指标中：

$$\bar{J}[u,\lambda]=h(x(t_f),t_f)+\lambda\cdot m(x(t_f),t_f)+\int_{t_0}^{t_f}g(x(t),u(t),t)\mathrm{d}t:=\bar{h}(x(t_f),t_f)+\int_{t_0}^{t_f}g(x(t),u(t),t)\mathrm{d}t$$

可以看出，除了需要满足哈密顿量 $\mathcal{H}(x(t),u(t),p(t),\lambda,t)$ 对 $u(t)$ 的极值条件，状态变量、协态变量的演化方程之外，需要满足边界条件：

$$\frac{\partial\bar{h}}{\partial x}(x(t_f),t_f,\lambda)-p(t_f)=0,\ \ m(x(t_f),t_f)=0,\ \ \mathcal{H}(x(t_f),u(t_f),p(t_f),\lambda t_f)+\frac{\partial\bar{h}}{\partial t}(x(t_f),t_f,\lambda)=0 \tag{1.57}$$

1.4.3 连续时间动态规划：哈密顿-雅可比-贝尔曼方程

上面介绍的庞特里亚金极值原理是一种“面向策略”的方法，我们还有一种“面向值函数”的方法。定义值函数 $V(x,t)$ 是在初始时刻 t，从初始状态 x 出发，运行到终端时刻所需的最小代价，换言之：

$$V(x,t)=\min_{u(t)}\left[h(x(t_f),t_f)+\int_t^{t_f}g(x(\tau),u(\tau),\tau)\mathrm{d}\tau\right]\quad s.t.x(t)=x$$

简单起见，考虑一类终端时间固定、终端状态自由的最优控制问题：有状态变量 $x(t)$，控制变量 $u(t)$，系统动力学方程 $\dot{x}(t)=f(x(t),u(t),t)$，性能指标 $J[u]=h(x(t_f),t_f)+\int_t^{t_f}g(x(\tau),u(\tau),\tau)\mathrm{d}\tau$。若值函数 $V(x,t)$ 二阶连续可微，则如下 PDE（被称为 Hamilton-Jacobi-Bellman 方程）是最优控制的充要条件：

$$-\frac{\partial V}{\partial t}(x(t),t)=\min_{u(t)\in\mathbb{R}^M}\mathcal{H}\left(x(t),u(t),\frac{\partial V(x(t),t)}{\partial x},t\right) \tag{1.58}$$

边界条件是：

$$V(x(t_f),t_f)=h(x(t_f),t_f) \tag{1.59}$$

其中 $\mathcal{H}=g+p\cdot f$。我们将分别对 HJB 方程的必要性和充分性给予证明。

首先证明必要性：若 $u(t)$ 是最优控制，则 HJB 方程必成立。这需要将值函数中的积分拆成两段：

$$\begin{aligned}V(x(t),t)&=\min_{u(\tau)}\left[h(x(t_f),t_f)+\int_t^{t_f}g(x(\tau),u(\tau),\tau)\mathrm{d}\tau\right]\\&=\min_{u(\tau)}\left[h(x(t_f),t_f)+\int_{t+\Delta t}^{t_f}g(x(\tau),u(\tau),\tau)\mathrm{d}\tau+\int_t^{t+\Delta t}g(x(\tau),u(\tau),\tau)\mathrm{d}\tau\right]\\&=\min_{u(\tau)}\left[V(x(t+\Delta t),t+\Delta t)+\int_t^{t+\Delta t}g(x(\tau),u(\tau),\tau)\mathrm{d}\tau\right]\end{aligned}$$

这里我们已经使用了序贯决策问题的最优子结构性质：若 $u(\tau),\tau\in[t,t_f]$ 是 $\tau\in[t,t_f]$ 上的最优控制，则 $u(\tau),\tau\in[t+\Delta t,t_f]$ 必然是 $\tau\in[t+\Delta t,t_f]$ 上的最优控制，如果不是，我们可以把 $\tau\in[t+\Delta t,t_f]$ 这一段的控制律换掉，这与 $u(\tau),\tau\in[t,t_f]$ 是最优控制矛盾。继续推导：

$$V(x(t),t)\approx\min_{u(\tau)}\left[g(x(t),u(t),t)\Delta t+V(x(t),t)+\frac{\partial V}{\partial t}(x(t),t)\Delta t+\frac{\partial V(x)}{\partial x}(x(t),t)\cdot f(x(t),u(t),t)\Delta t\right]$$

两侧同时减去 $V(x(t),t)$，并把 $\frac{\partial V}{\partial t}(x(t),t)\Delta t$ 移到左侧（我们能做这两个操作是因为我们操作的项与 $u(t)$ 无关，不受“取最小”这个操作的影响），然后两边除以 Δt，直接得到 HJB 方程。

接下来证明充分性。我们需要证明 HJB 方程满足时：1）$V(x(t),t)$ 确实是值函数，2）且 $u(t)$ 确实是最优控制。分两步完成这个证明，第一步证明 $V(x_0,t_0)$ 是在 t_0 时刻从 x_0 状态出发，施加某个控制律 $u(t)$ 的性能指标，

这需要我们找到 $V(x(t),t)$ 的全导数。记：

$$u(t) = \arg\min_{\xi} \mathcal{H}\left(x(t), \xi, \frac{\partial V(x(t),t)}{\partial x}, t\right)$$

那么：

$$\begin{aligned}
-\frac{\partial V}{\partial t}(x(t),t) &= \min_{\xi} \mathcal{H}\left(x(t), \xi, \frac{\partial V(x(t),t)}{\partial x}, t\right) \\
&= \mathcal{H}\left(x(t), u(t), \frac{\partial V(x(t),t)}{\partial x}, t\right) \\
&= g(x(t),u(t),t) + \frac{\partial V(x(t),t)}{\partial x} \cdot f(x(t),u(t),t)
\end{aligned}$$

移项：

$$\begin{aligned}
0 &= g(x(t),u(t),t) + \frac{\partial V(x(t),t)}{\partial x} \cdot f(x(t),u(t),t) + \frac{\partial V}{\partial t}(x(t),t) \\
&= g(x(t),u(t),t) + \frac{\partial V(x(t),t)}{\partial x} \cdot \dot{x} + \frac{\partial V}{\partial t}(x(t),t) \\
&= g(x(t),u(t),t) + \frac{\mathrm{d}}{\mathrm{d}t}\left[V(x(t),t)\right]
\end{aligned}$$

对 $[t_0, t_f]$ 区间积分，得到：

$$V(x(t_0),t_0) = V(x(t_f),t_f) + \int_{t_0}^{t_f} g(x(t),u(t),t)\mathrm{d}t = h(x(t_f),t_f) + \int_{t_0}^{t_f} g(x(t),u(t),t)\mathrm{d}t$$

第二步，我们证明使得每时每刻 $\mathcal{H}$ 取最小值的控制律 $u(t)$ 确实是最优控制律。设我们使用了不同于最优控制的控制律 $u'(t)$，系统的状态是 $x'(t)$，则：

$$-\frac{\partial V}{\partial t}(x'(t),t) = \min_{\xi} \mathcal{H}\left[x'(t), \xi, \frac{\partial V}{\partial x}(x'(t),t), t\right] \leq \mathcal{H}\left[x'(t), u'(t), \frac{\partial V}{\partial x}(x'(t),t), t\right]$$

使用与第一步中类似的推导，得到：

$$V(x'(t_0) = x_0, t_0) \leq V(x'(t_f),t_f) + \int_{t_0}^{t_f} g(x'(t),u'(t),t)\mathrm{d}t$$

左侧是从 t_0 时刻出发，以 x_0 为初始状态，控制律 $u(t)$ 下的性能指标，右侧是控制律 $u'(t)$ 下的性能指标。所以 $u(t)$ 确实是最优控制。

最后，我们要证明 $\frac{\partial V}{\partial x}(x(t),t)$ 就是协态变量，从而 HJB 方程给出了与庞特里亚金极值原理相同的 $\mathcal{H}$ 对 $u(t)$ 的极值条件。直接追踪其时间演化（仍以 $u(t)$ 记代表最优控制）：

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial V}{\partial x}(x(t),t)\right] &= \frac{\partial^2 V}{\partial x^2}(x(t),t)\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial^2 V}{\partial t \partial x}(x(t),t) \\
&= \frac{\partial^2 V}{\partial x^2}(x(t),t) \cdot f(x(t),u(t),t) + \frac{\partial^2 V}{\partial x \partial t}(x(t),t) \\
&= \frac{\partial^2 V}{\partial x^2}(x(t),t) \cdot f(x(t),u(t),t) - \frac{\partial}{\partial x}\left[g(x(t),u(t),t) + \frac{\partial V}{\partial x}(x(t),t) \cdot f(x(t),u(t),t)\right] \\
&= -\frac{\partial g}{\partial x}(x(t),u(t),t) - \frac{\partial V}{\partial x}(x(t),t) \cdot \frac{\partial f}{\partial x}(x(t),u(t),t) \\
&= -\frac{\partial \mathcal{H}}{\partial x}\left(x(t), u(t), \frac{\partial V}{\partial x}(x(t),t), t\right)
\end{aligned}$$

所以 $\frac{\partial V}{\partial x}(x(t),t)$ 的演化方程与 $p(t)$ 一致。边界条件 $V(x(t_f),t_f) = h(x(t_f),t_f)$ 直接给出：

$$p(x(t_f),t_f) = \frac{\partial h}{\partial x}(x(t_f),t_f)$$

t_f 时刻使用 HJB 方程给出：

$$0 = \frac{\partial h}{\partial x}(x(t_f),t_f) + \mathcal{H}(x(t_f),u(t_f),\frac{\partial V}{\partial x}(x(t_f),t_f),t_f)$$

这是前面动力学方程和性能指标含时问题的关于 $p(t)$ 和 $\mathcal{H}(x,u,p,t)$ 的边界条件。

1.4.4 简单例子：线性二次型调节器

现在我们举一个简单且常用的例子：线性二次型调节器（LQR）。考虑系统有状态变量 $x(t):[t_0,t_f]\to\mathbb{R}^N$，控制变量 $u(t):[t_0,t_f]\to\mathbb{R}^M$，系统的动力学方程是：

$$\dot{x}(t)=A(t)x(t)+B(t)u(t)\quad A(t):[t_0,t_f]\to\mathbb{R}^{N\times N},B(t):[t_0,t_f]\to\mathbb{R}^{N\times M}$$

系统终端时刻 t_f 固定，终端状态自由，控制变量无约束，最小化二次型的性能指标：

$$J[u]=\frac{1}{2}x^T(t_f)Hx(t_f)+\frac{1}{2}\int_{t_0}^{t_f}\left[x^T(t)Q(t)x(t)+u^T(t)R(t)u(t)\right]\mathrm{d}t$$

其中 $H\in\mathbb{R}^{N\times N}$，$Q(t):[t_0,t_f]\to\mathbb{R}^{N\times N}$，$R(t):[t_0,t_f]\to\mathbb{R}^{M\times M}$，$H,Q(t)$ 半正定，$R(t)$ 正定。首先使用庞特里亚金极值原理求解这个问题。构造哈密顿量：

$$\mathcal{H}(x,u,p,t)=\frac{1}{2}x^T(t)Q(t)x(t)+\frac{1}{2}u^T(t)R(t)u(t)+p^T(t)A(t)x(t)+p^T(t)B(t)u(t)$$

利用极值条件：

$$0=\frac{\partial\mathcal{H}}{\partial u}=R(t)u(t)+B^T(t)p(t)\Rightarrow u(t)=-R^{-1}(t)B^T(t)p(t)\tag{1.60}$$

这是 LQR 问题的开环控制律 $u(t)$。将其代入哈密顿量，使用正则方程组即可得到状态变量、协态变量的演化：

$$\mathcal{M}(t)=\begin{bmatrix}A(t)&-B(t)R^{-1}(t)B^T(t)\\-Q(t)&-A^T(t)\end{bmatrix}\quad\begin{bmatrix}\dot{x}(t)\\\dot{p}(t)\end{bmatrix}=\mathcal{M}(t)\begin{bmatrix}x(t)\\p(t)\end{bmatrix}$$

然而，这样的开环控制律是在使用之前预先计算好的，在实际使用中，系统常受到外界扰动导致系统状态偏离开环控制律下的 $x(t)$，此时再给系统施加预先计算好的开环控制律可能会导致“越来越错”的效果，系统的鲁棒性极差（没有反馈，就没有鲁棒性！这会导向一个被称为“鲁棒控制”的领域，专门设计控制器以最优化最坏情况下的控制器性能！）。因此，我们希望计算一个反馈控制律/闭环控制律，显然是要将 $p(t)$ 用 $x(t)$ 表示出来。该微分方程的解有如下形式：

$$\begin{bmatrix}x(t_f)\\p(t_f)\end{bmatrix}=\begin{bmatrix}\phi_{11}(t_f,t)&\phi_{12}(t_f,t)\\\phi_{12}(t_f,t)&\phi_{22}(t_f,t)\end{bmatrix}\begin{bmatrix}x(t)\\p(t)\end{bmatrix}$$

根据 $p(t)$ 的边界条件，有：

$$p(t_f)=Hx(t_f)\Rightarrow H\phi_{11}(t_f,t)x(t)+H\phi_{12}(t_f,t)p(t)=\phi_{21}(t_f,t)x(t)+\phi_{22}(t_f,t)p(t)$$

从这里解出：

$$p(t)=[\phi_{22}(t_f,t)-H\phi_{12}(t_f,t)]^{-1}[H\phi_{11}(t_f,t)-\phi_{21}(t_f,t)]x(t):=K(t)x(t)$$

这里的 $K(t)$ 可以直接解 ODEs 解出来，我们也可以导出 $K(t)$ 服从的演化方程。将 $p(t)$ 代入 $x(t),p(t)$ 的演化方程：

$$\dot{x}(t)=[A(t)-B(t)R^{-1}(t)B(t)K(t)]x(t)\quad\dot{K}(t)x(t)+K(t)\dot{x}(t)=[-Q(t)-A^T(t)K(t)]x(t)$$

消去 $\dot{x}(t)$，使得 $(\cdots)x(t)=0$ 中 $x(t)$ 前的系数得 0，可以得到 $K(t)$ 满足的 Ricatti 递归方程：

$$0=\dot{K}(t)+Q(t)-K(t)B(t)R^{-1}(t)K(t)+K(t)A(t)+A^T(t)K(t)\tag{1.61}$$

边界条件 $K(t_f)=H$。

这个问题也可以使用 HJB 方程求解。哈密顿量中的 $p(t)$ 换成 $\dfrac{\partial V}{\partial x}$：

$$\mathcal{H}\left(x(t),u(t),\frac{\partial V}{\partial x}(x(t),t),t\right)=\frac{1}{2}x^T(t)Q(t)x(t)+\frac{1}{2}u^T(t)R(t)u(t)+\left[\frac{\partial V}{\partial x}(x(t),t)\right][A(t)x(t)+B(t)u(t)]$$

由 $\mathcal{H}$ 取极值得到：

$$u(t)=-R^{-1}(t)B(t)\frac{\partial V}{\partial x}(x(t),t)$$

代入 HJB 方程：

$$\frac{\partial V}{\partial t}(x(t),t)+\frac{1}{2}x^T(t)Q(t)x(t)+\left[\frac{\partial V}{\partial x}(x(t),t)\right]A(t)x(t)-\frac{1}{2}\left[\frac{\partial V}{\partial x}(x(t),t)\right]^TB(t)R^{-1}(t)B^T(t)\frac{\partial V}{\partial x}(x(t),t)=0$$

直接解这个方程当然是很难的，需要猜测值函数的形式。直接猜测它也是二次型，即 $V(x(t),t)=\frac{1}{2}x^T(t)K(t)x(t)$，将其代入上式，可以得到与方程1.61相同的结果。

1.4.5 * 试验内容：最优控制与经典力学的关系；一类最优控制问题的几何化

看到这里，相信每个人都会有点疑惑：为什么我们将 $\mathcal{H}$ 叫做“哈密顿量”？为什么状态变量和协态变量竟然满足哈密顿正则方程组？这里给出一种也许可行的解释。首先我们说明其实有两种等效的极值原理。之前，我们构造的哈密顿量是：

$$\mathcal{H}=g(x(t),u(t),t)+p(t)\cdot f(x(t),u(t),t)$$

正则方程组、极值条件：

$$\dot{x}=\frac{\partial\mathcal{H}}{\partial p}=f\quad \dot{p}=-\frac{\partial\mathcal{H}}{\partial x}=-\frac{\partial g}{\partial x}-p\cdot\frac{\partial f}{\partial x}\quad u(t)=\arg\min_{u(t)}\mathcal{H}(x(t),u(t),p(t),t)$$

严格来说这个叫庞特里亚金极小值原理，还有另外一种极值原理，它的哈密顿量：

$$\mathcal{H}'=-g(x'(t),u'(t),t)+p'(t)\cdot f(x'(t),u'(t),t)$$

正则方程组和极值条件：

$$\dot{x}'=\frac{\partial\mathcal{H}'}{\partial p'}=f\quad \dot{p}'=-\frac{\partial\mathcal{H}'}{\partial x'}=+\frac{\partial g'}{\partial x'}-p'\cdot\frac{\partial f'}{\partial x'}\quad u'(t)=\arg\max_{u'(t)}\mathcal{H}'(x'(t),u'(t),p'(t),t)$$

这被称为严格来说这个叫庞特里亚金极大值原理。很容易证明它们二者是等价的，只需令 $x'(t)=x(t),p'(t)=-p(t),u(t)=u'(t)$，则 $\mathcal{H}'=-\mathcal{H}$，从而这个 $u(t)$ 在最小化 $\mathcal{H}$ 的同时将最大化 $\mathcal{H}'$，也就是说极小值原理和极大值原理给出相同的控制律；并且，此时两套正则方程组也是完全相同的。

我们在极小值原理的框架下试着找找最优控制和经典力学的关系。由于整个经典力学都是在标准变分的框架下推导的，没有使用针状变分，因此它的能力仅限于处理控制律无约束的最优控制问题。简便起见，考虑系统有状态变量 $x(t):[t_0,t_f]\to\mathbb{R}^N$，控制变量 $u(t):[t_0,t_f]\to\mathbb{R}^M$，终端状态、终端时刻自由或固定（我们暂时不管边界条件，先尝试做出 $x(t),p(t)$ 的演化方程），性能指标只有运行代价：$J[u]=\int_{t_0}^{t_f}g(x(t),u(t),t)\mathrm{d}t$。一个重要的特性是，经典力学是一种“命题作文”，系统演化的所有信息都包含在作用量中；最优控制是“完形填空”，性能指标和动力学方程共同指定了系统的演化。所以首先我们要给最优控制问题找个作用量/拉氏量。一种做法是：考虑 $2N+M$ 维系统，将其前 N 个广义坐标 $q_1,\cdots,q_N$ 对应到状态变量 $x_1(t),\cdots x_N(t)$；将其第 $N+1\sim N+M$ 个广义坐标对应到控制变量 $u_1(t),\cdots,u_M(t)$；将其第 $N+M+1\sim 2N+M$ 个广义坐标对应到协态变量 $p_1(t),\cdots p_N(t)$，使用如下拉氏量：

$$\mathcal{L}(q,\dot{q},t)=-g(q,t)+\sum_{i=1}^{N}q_{N+M+i}(\dot{q}_i-f_i(q))$$

做个勒让德变换给出哈氏量。先求正则动量：

$$p_i=\frac{\partial L}{\partial\dot{q}_i}=q_{N+M+i},\ p_{N+k}=\frac{\partial L}{\partial\dot{q}_{N+k}}=0\ (k=1,\cdots,M),\ p_{N+M+i}=\frac{\partial L}{\partial\dot{q}_{N+M+i}}=0\ (i=1,\cdots,N)$$

注意这里的 p 是正则动量，不是之前提到的协态变量（然而它们共用符号了）。我们之前把 q_{N+M+i} 设置成协态变量，现在可以看到它们就是与前 N 个正则坐标对应的正则动量。求出哈氏量：

$$\begin{aligned}\mathcal{H}&=\sum_{i=1}^{2M+N}p_i\dot{q}_i-\mathcal{L}\\&=\sum_{i=1}^{N}p_i\dot{q}_i-\sum_{i=1}^{N}p_i(\dot{q}_i-f_i(q,t))+g(q,t)\\&=g(q,t)+\sum_{i=1}^{N}p_if_i(q,t)\end{aligned}$$

这个就是我们在极大值原理中构造的哈密顿量。但是注意我们之前求出了 $p_{N+k}=0,p_{N+M+i}=0$，也就是说系统有 $N+M$ 个约束，根据 Dirac 的约束理论，我们需要在哈氏量中补充 $N+M$ 个乘子。令 $\phi_{N+k}(p,q)=$

$p_{N+k}\ (k=1,\cdots,M)$，$\phi_{N+M+j}(p,q)=p_{N+M+j}\ (j=1,\cdots N)$，加入乘子后的哈氏量是：

$$\mathcal{H}^{\dagger}=g(q,t)+\sum_{i=1}^{N}p_i f_i(q,t)+\sum_{k=1}^{M}\phi_{N+k}(p,q)\lambda_{N+k}(t)+\sum_{j=1}^{N}\phi_{N+M+j}(p,q)\lambda_{N+M+j}(t)$$

下面写出正则方程。关于下标 $1,\cdots,N$ 的正则方程是：

$$\dot{q}_i=\frac{\partial\mathcal{H}}{\partial p_i}=f_i(q,t)\quad \dot{p}_i=-\frac{\partial\mathcal{H}}{\partial q_i}=-\frac{\partial g(q,x)}{\partial q_i}-\sum_{j=1}^{N}p_j\frac{\partial f_j(q,t)}{\partial q_i}$$

这是状态变量、协态变量的演化方程。关于下标 $N+1,\cdots,N+M$ 的正则方程是：

$$\dot{q}_{N+k}=\frac{\partial\mathcal{H}}{\partial p_{N+k}}=\lambda_{N+k}(t)\quad \dot{p}_{N+k}=-\frac{\partial\mathcal{H}}{\partial q_{N+k}}=-\frac{\partial g(q,x)}{\partial q_{N+k}}-\sum_{j=1}^{N}p_j\frac{\partial f_j(q,t)}{\partial q_{N+k}}=0$$

这里前一半方程不给出任何信息，后一半方程作为次级约束给出最优控制律。关于下标 $N+M+1,\cdots 2N+M$ 的正则方程是：

$$\dot{q}_{N+M+i}=\frac{\partial\mathcal{H}}{\partial p_{N+M+i}}=\lambda_{N+M+i}(t)\quad \dot{p}_{N+M+i}=-\frac{\partial\mathcal{H}}{\partial q_{N+M+i}}=0$$

从上面这些方程，我们可以看出 $p_{N+k}q_{N+M+i},p_{N+M+i}$ 这三部分完全没用，通过求解 $\dot{p}_{N+k}=0$ 这 M 个方程，在一些条件下（隐函数定理），我可将控制律 q_{N+k} 以 q_i,p_i 显式写出。从而系统应当被视作在相空间的 $2N$ 维子流形上运动（换言之，原系统等价于一个 $2N$ 维相空间中的新系统，这个新系统的哈氏量是 $\mathcal{H}'=g(q,t)+\sum_{i=1}^{N}p_i f_i(q,t)$，$f_i(q,t)$ 中所有的控制变量使用 q_i,p_i 表示出），只有 q_i,p_i 是能有效描述系统状态的变量。

经典力学中的 HJB 方程和上述最优控制中的 HJB 方程有不同，经典力学中的 HJB 方程是关于经典作用量/第二类正则变换生成函数的方程，其自变量是路径的终点；最优控制中的 HJB 方程是关于值函数的方程，其自变量是路径的起点。为了从经典力学的角度做出最优控制中的 HJB 方程，我们需要研究从某一起点 (q,t) 出发，到达固定终点 (q_f,t_f) 的路径的作用量。考虑扰动路径的初始点、初始时间和路径本身，作用量的变化是：

$$\delta S(q,t)=\int_{t_0}^{t_f}\left(\frac{\partial\mathcal{L}}{\partial q}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}}\right)\right)\mathrm{d}t+\left(\frac{\partial\mathcal{L}}{\partial\dot{q}}\delta q(t)\right)\Big|_{t_0}^{t_f}-\mathcal{L}\delta t=-\left(\frac{\partial\mathcal{L}}{\partial\dot{q}}\delta q(t)\right)\Big|_{t_0}-\mathcal{L}\delta t$$

利用关系 $\delta q(t_0)=\mathrm{d}q-\dot{q}(t_0)\delta t$，$q'(t_0+\delta t)-q(t_0)=\mathrm{d}q$，可知：

$$\delta S=-(p\cdot\mathrm{d}q-p\cdot\dot{q}\mathrm{d}t)-\mathcal{L}\mathrm{d}t=-p\cdot\mathrm{d}q+\mathcal{H}\mathrm{d}t=\frac{\partial S}{\partial q}\cdot\mathrm{d}q+\frac{\partial S}{\partial t}\mathrm{d}t$$

注意，按照上文中拉氏量的定义，路径的经典作用量 $S=\int\mathcal{L}\mathrm{d}t$ 是值函数的相反数（我们在拉氏量中减去了运行代价）。因此：

$$\delta V=+p\cdot\mathrm{d}q-\mathcal{H}\mathrm{d}t=\frac{\partial V}{\partial q}\cdot\mathrm{d}q+\frac{\partial V}{\partial t}\mathrm{d}t$$

从中直接提取出最优控制的 HJB 方程：

$$\frac{\partial V}{\partial t}(q,t)=-\mathcal{H}\left(q,\frac{\partial V}{\partial q},t\right)$$

另一个有趣的问题是最优控制律下系统的轨迹可以被视作某个度规下的测地线吗？或者说我们能否找到一套程序化的方式将最优控制问题几何化？这里浅浅讨论一类问题，系统有状态变量 $x(t):[t_0,t_f]\to\mathbb{R}^N$，控制变量 $u_1(t),u_2(t),\cdots,u_M(t)$ 都是标量函数，系统的动力学方程 $\dot{x}=f_0(x)+\sum_{i=1}^{M}u_i f_i(x)$，$f_i(x):[t_0,t_f]\times\mathbb{R}^N\to\mathbb{R}^N$，性能指标 $J[u]=\int\frac{1}{2}\sum_{i=1}^{M}u_i^2\mathrm{d}t$。这个问题代表了一系列的能量最小化问题。这里我们从极大值原理的哈氏量出发开始讨论（这只是因为我是从一些简单例子开始尝试的，从极大值原理出发导致后面的一些例子中，$\mathcal{H}$ 将包含一个关于 p 的正定二次型，看起来比较顺眼），考虑庞特里亚金极大值原理的哈氏量：

$$\mathcal{H}=p^T\left(f_0(x)+\sum_{i=1}^{M}u_i f_i(x)\right)-\frac{1}{2}\sum_{i=1}^{M}u_i^2$$

最优控制律是 $u_i=p^T f_i(x)$，代入得到上文中提到的 $2N$ 维相空间中的哈氏量：

$$\mathcal{H}=p^T f_0(x)+\frac{1}{2}\sum_{i=1}^{M}p^T F_i(x)p,\quad F_i(x)=f_i(x)f_i^T(x)$$

从这里能看出其实最优控制和哈氏力学是自动兼容的，然而我们一般不在相空间中谈度规（相空间中只有正则辛形式），在拉氏力学中，我们可以在构型空间上配备度规，比如如果拉氏量形如 $\mathcal{L}=\frac{1}{2}m_{ij}\dot{x}^i\dot{x}^j-V(x)$，我们会有 Jacobi 度规 $g_{ij}=2(E-V(x))m_{ij}$，所以我们的第一个想法一定是试图把这个哈氏力学系统换到拉氏力学系统。然而，这个变换要求 $\sum_{i=1}^M F_i(x)$ 可逆。一般情况下这不可能，一种可行的方式是先补上一个 $p^T(\epsilon I)p$，之后再令 $\epsilon\to 0$。令 $\tilde{F}(x)=\sum_{i=1}^M F_i(x)+\epsilon I$，考虑哈氏量到拉氏量的变换：

$$\tilde{\mathcal{H}}=p^Tf_0(x)+\frac{1}{2}p^T\tilde{F}(x)p\Rightarrow\mathcal{L}=\frac{1}{2}(\dot{x}-f_0(x))^T\tilde{F}^{-1}(x)(\dot{x}-f_0(x))$$

这个拉氏量的形式可以归结为：

$$\mathcal{L}=\frac{1}{2}m_{ij}(x)\dot{x}^i\dot{x}^j+A_i(x)\dot{x}^i-V(x)$$

其实就是电磁场中经典带电粒子的拉氏量。我们声称这个拉氏量可以通过“升一维”的方式实现几何化：引入新坐标 s，以下度规的测地线在 $\{x^1,\cdots,x^N\}$ 上的投影与上述拉氏量给出的运动轨迹相同：

$$(\mathrm{d}l)^2=m_{ij}(x)\mathrm{d}x^i\mathrm{d}x^j+\phi(x)(\mathrm{d}s+A_i(x)\mathrm{d}x^i)^2$$

我们来证明这一点，同时将未定的标量函数 $\phi(x)$ 确定下来。由 E-L 方程，原拉氏量给出的运动方程是：

$$\frac{1}{2}\frac{\partial m_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j-\frac{\partial m_{kj}}{\partial x^l}\dot{x}^l\dot{x}^j-m_{kj}\ddot{x}^j+F_{ki}\dot{x}^i-\frac{\partial V}{\partial x^k}=0,\quad F_{ki}=\frac{\partial A_i}{\partial x_k}-\frac{\partial A_k}{\partial x_i}$$

仍然使用 E-L 方程求出 $(dl)^2=\cdots$ 度规的测地线，这需要构造拉氏量：

$$\bar{\mathcal{L}}=\frac{1}{2}m_{ij}(x)\dot{x}^i\dot{x}^j+\frac{1}{2}\phi(x)\left(\dot{s}+A_i(x)\dot{x}^i\right)^2$$

显然 s 是循环坐标，那么有守恒量：

$$\phi(x)(\dot{s}+A_i(x)\dot{x}^i)=\mathrm{Const}=Q$$

关于 x^k 的 E-L 方程给出：

$$\frac{1}{2}\frac{\partial m_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j-\frac{\partial m_{kj}}{\partial x^l}\dot{x}^l\dot{x}^j-m_{kj}\ddot{x}^j+QF_{ki}\dot{x}^i+\frac{1}{2}\frac{\partial\phi}{\partial x^k}\frac{Q^2}{\phi^2(x)}=0$$

对比可知 $Q=1$，以及

$$\frac{Q^2}{\phi^2(x)}\frac{\partial\phi(x)}{\partial x^k}=-\frac{\partial V(x)}{\partial x^k}\Rightarrow\frac{1}{2}\frac{1}{\phi(x)}=V(x)+C$$

C 是任意常数。至此我们完成了证明。看看新引入的坐标 s，由于 $\phi(x)(\dot{s}+A_i(x)\dot{x}^i)=1$，它不是独立的。$s$ 本身不出现在度规中，因此 s 的取值有规范自由性。在度规中出现了 $(\mathrm{d}s+A_i(x)\mathrm{d}x^i)$ 这样的项，如果我对磁场做规范变换，使得 $A_i(x)\to A_i(x)+\partial_i\Lambda(x)$，那么我可以同时做变换：$s\to s-\Lambda(x)$ 使得度规保持不变。注意，电磁场的规范变换与波函数/狄拉克旋量的相位变换一同出现，所以我完全可以认为我们的构型空间不是 $\mathbb{R}^N\times\mathbb{R}^1$，而是 $\mathbb{R}^N\times\mathbb{S}^1$，也就是说新加入的这个维度是个圆环，做变换 $s\to s+2\pi$ 后回到原点。我们讨论的是如何为电磁场中经典粒子的构型空间配备度规，但是这个理论有相对论协变版本（Kaluza-Klein 理论），这是人类历史上第一次统一引力和电磁相互作用的尝试，这种引入新的、被紧致化的维度的做法将通向弦论。

下面算一个例子。考虑非常经典的最优控制问题：

$$\dot{x}_1=x_2\quad\dot{x}_2=u\qquad\min\int\frac{1}{2}u^2\mathrm{d}t$$

按照极大值原理的做法，哈氏量是 $\mathcal{H}=p_1x_2+p_2u-\frac{1}{2}u^2$，最优控制律是 $u=p_2$，代入后得到哈密顿量 $\mathcal{H}=p_1x_2+\frac{1}{2}p_2^2$，加上我们之前提到的 $\epsilon-$ 修正，并变成拉氏量：

$$\tilde{\mathcal{H}}=\frac{1}{2}\epsilon p_1^2+p_1x_2+\frac{1}{2}(1+\epsilon)p_2^2\Rightarrow\mathcal{L}=\frac{1}{2}(\dot{x}_1-x_2)^2+\frac{1}{2(1+\epsilon)}\dot{x}_2^2$$

对比得：

$$m_{11}=\frac{1}{\epsilon},\ m_{22}=\frac{1}{(1+\epsilon)},\ A_1(x)=-\frac{x_2}{\epsilon},\ \phi(x)=-\frac{\epsilon}{x_2^2+C}$$

不妨取 $C=1$，构造度规：

$$(\mathrm{d}l)^2=\frac{1}{\epsilon}(\mathrm{d}x_1)^2+\frac{1}{(1+\epsilon)}(\mathrm{d}x_2)^2-\frac{\epsilon}{(x_2^2+1)}\left(\mathrm{d}s-\frac{x_2}{\epsilon}\mathrm{d}x_1\right)^2$$

仍然用 E-L 方程把它的测地线做出来，令：

$$\bar{\mathcal{L}}=\frac{1}{2\epsilon}(\dot{x}_1)^2+\frac{1}{2(1+\epsilon)}(\dot{x}_2)^2-\frac{\epsilon}{2(x_2^2+1)}\left(\dot{s}-\frac{x_2}{\epsilon}\dot{x}_1\right)^2$$

首先，s 是循环坐标，所以：

$$\frac{\partial\bar{\mathcal{L}}}{\partial\dot{s}}=-\frac{\partial\epsilon}{\partial(x_2^2+1)}\left(\dot{s}-\frac{x_2}{\epsilon}\dot{x}_1\right)=C_1=1$$

x_1 也是循环坐标，所以：

$$\frac{\partial\mathcal{L}}{\partial\dot{x}_1}=\frac{1}{\epsilon}\dot{x}_1+\frac{\epsilon}{(x_2^2+1)}\left(\dot{s}-\frac{x_2}{\epsilon}\dot{x}_1\right)\frac{x_2}{\epsilon}=C_2\Rightarrow\dot{x}_1=x_2+\epsilon C_2$$

对 x_2 的 E-L 方程给出：

$$-\frac{1}{1+\epsilon}\ddot{x}_2+\frac{\epsilon}{(x_2^2+1)^2}x_2\left(\dot{s}-\frac{x_2}{\epsilon}\dot{x}_1\right)^2+\frac{\epsilon}{(x_2^2+1)}\left(\dot{s}-\frac{x_2}{\epsilon}\dot{x}_1\right)\frac{\dot{x}_1}{\epsilon}\Rightarrow\ddot{x}_2=(1+\epsilon)C_2$$

取 $\epsilon\to 0$，立刻得到这个最优控制问题的解：

$$\dot{x}_1=x_2\quad\ddot{x}_2=C_2\Rightarrow x_1(t)=A(t-t_0)^3+B(t-t_0)^2+C(t-t_0)+D$$