

ECE 6553: Homework #2

Klaus Okkelberg

February 16, 2017

1. (a) Since τ and x_0 are given, $x(t)$ is fixed for $0 \leq t < \tau$. Thus, the cost only depends on the integral from τ to T :

$$J(\rho) = \int_0^T L(x(t)) dt = \int_0^\tau L(x(t)) dt + \int_\tau^T L(x(t)) dt = C + \int_\tau^T L(x(t)) dt,$$

since the finite jump ρ in $x(\tau)$ does not affect the integral. In essence, we are finding the optimal initial condition at time τ , i.e. $x(\tau^+)$. The Lagrangian is

$$\tilde{J}(\rho) = C + \int_\tau^T \left\{ L(x(t)) + \lambda^T(t)[f(x(t)) - \dot{x}(t)] \right\} dt$$

Perturbing ρ as $\rho \mapsto \rho + \varepsilon\nu$ causes a variation in $x(t) \mapsto x(t) + \varepsilon\eta(t)$:

$$\begin{aligned} \tilde{J}(\rho + \varepsilon\nu) &= C + \int_\tau^T \left\{ L(x(t) + \varepsilon\eta(t)) + \lambda^T(t)[f(x(t) + \varepsilon\eta(t)) - \dot{x}(t) - \varepsilon\dot{\eta}(t)] \right\} dt + o(\varepsilon) \\ &= C + \int_\tau^T \left\{ L(x) + \varepsilon \frac{\partial L}{\partial x}(x)\eta + \lambda^T \left[f(x) + \varepsilon \frac{\partial f}{\partial x}(x)\eta - \dot{x} - \varepsilon\dot{\eta} \right] \right\} dt + o(\varepsilon) \\ \tilde{J}(\rho + \varepsilon\nu) - \tilde{J}(\rho) &= \int_\tau^T \left\{ \varepsilon \frac{\partial L}{\partial x}(x)\eta + \lambda^T \left[\varepsilon \frac{\partial f}{\partial x}(x)\eta - \varepsilon\dot{\eta} \right] \right\} dt + o(\varepsilon) \\ \delta\tilde{J}(\rho; \nu) &= \int_\tau^T \left[\frac{\partial L}{\partial x}(x)\eta + \lambda^T \frac{\partial f}{\partial x}(x)\eta \right] dt - \int_\tau^T \lambda^T \dot{\eta} dt \\ &= \int_\tau^T \left[\frac{\partial L}{\partial x}(x)\eta + \lambda^T \frac{\partial f}{\partial x}(x)\eta \right] dt + \int_\tau^T \dot{\lambda}^T \eta dt - \lambda^T(T)\eta(T) + \lambda^T(\tau)\eta(\tau) \\ &= \int_\tau^T \left[\frac{\partial L}{\partial x}(x) + \lambda^T \frac{\partial f}{\partial x}(x) + \dot{\lambda}^T \right] \eta dt - \lambda^T(T)\eta(T) + \lambda^T(\tau)\eta(\tau) \end{aligned}$$

We want the directional derivative to be 0, so we pick

$$\begin{cases} \dot{\lambda}(t) = -\frac{\partial L^T}{\partial x}(x(t)) - \frac{\partial f^T}{\partial x}(x(t))\lambda(t) \\ \lambda(T) = 0 \end{cases} \quad (1)$$

This gives

$$\delta\tilde{J}(\rho; \nu) = \frac{\partial \tilde{J}}{\partial \rho} \nu = \lambda^T(\tau)\eta(\tau) = \lambda^T(\tau)\nu.$$

This is linear in ν so the FONC is

$$\frac{\partial \tilde{J}}{\partial \rho} = \lambda(\tau) = 0,$$

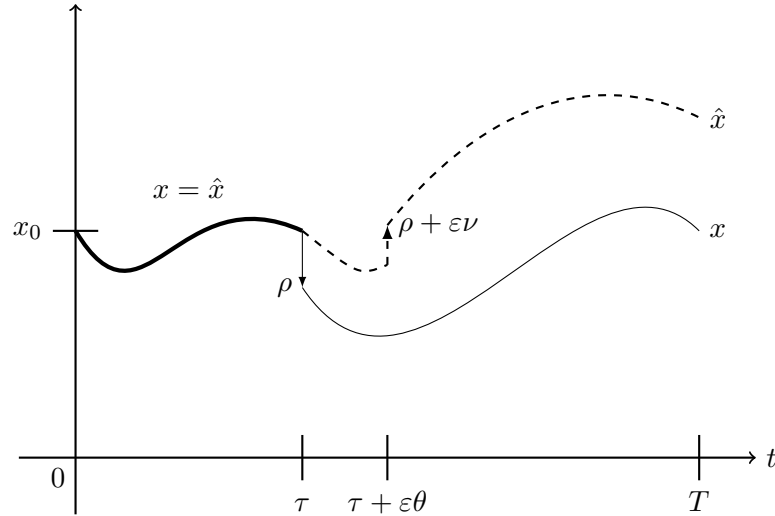
where $\lambda(t)$ satisfies (1).

- (b) Now, we perturb $\rho = \alpha w \mapsto (\alpha + \varepsilon)w = \alpha w + \varepsilon\nu$, where $w = \nu$. This causes a variation in $x(t) \mapsto x(t) + \varepsilon\eta(t)$, as before. However, the directional derivative only has to be 0 along $\nu = w$ instead of for all $\nu \in \mathbb{R}^n$. Thus, the FONC is

$$\delta\tilde{J}(\rho; \nu) = \lambda^T(\tau)w = 0.$$

The $\lambda(\tau) \in \mathbb{R}^n$ that satisfy this condition are vectors that lie in the hyperplane perpendicular to w , of which there are infinitely many. $\lambda(t)$ still has to satisfy (1).

2. (a) The thick line in the figure indicates the portion where $x(t) = \hat{x}(t)$, $0 \leq t < \tau$. The thin solid line is the original trajectory x , and the dashed line is the perturbed trajectory \hat{x} . The vertical lines with arrowheads indicate the state jumps.



- (b) Yes, direct application of Calculus of Variations would run into trouble. So far in class, we have only considered one perturbation. However, here both ρ and τ are being perturbed. This means there are two changes in trajectory, i.e. $\hat{x}(t) = x(t) + \varepsilon\eta_1(t) + \varepsilon\eta_2(t) + o(\varepsilon)$.

The FONCs for both perturbations can be computed using Calculus of Variations, but there would be two equations constraining the costate, one from each directional derivative. This means the costate is likely overconstrained and not solvable.

3. (a) The Lagrangian is

$$\tilde{J}(\tau) = \int_0^\tau [L(x(t)) + \lambda^T(t)(-\alpha - \dot{x})] dt + \int_\tau^1 [L(x(t)) + \lambda^T(t)(\alpha - \dot{x})] dt$$

Perturbing $\tau \mapsto \tau + \varepsilon\theta$ causes $x(t) \mapsto x(t) + \varepsilon\eta(t)$. Note that $\eta = \dot{\eta} = 0$ on $[0, \tau)$.

$$\begin{aligned} \tilde{J}(\tau + \varepsilon\theta) &= \int_0^\tau \{L(x) + \lambda^T[-\alpha - \dot{x}]\} dt \\ &\quad + \int_\tau^{\tau+\varepsilon\theta} \{L(x + \varepsilon\eta) + \lambda^T[-\alpha - \dot{x} - \varepsilon\dot{\eta}]\} dt \\ &\quad + \int_{\tau+\varepsilon\theta}^1 \{L(x + \varepsilon\eta) + \lambda^T[\alpha - \dot{x} - \varepsilon\dot{\eta}]\} dt + o(\varepsilon) \\ &= \int_0^\tau \{L(x) + \lambda^T[-\alpha - \dot{x}]\} dt \\ &\quad + \int_\tau^{\tau+\varepsilon\theta} \left\{ L(x) + \varepsilon \frac{\partial L}{\partial x}(x)\eta + \lambda^T[-\alpha - \dot{x} - \varepsilon\dot{\eta}] \right\} dt \\ &\quad + \int_{\tau+\varepsilon\theta}^1 \left\{ L(x) + \varepsilon \frac{\partial L}{\partial x}(x)\eta + \lambda^T[\alpha - \dot{x} - \varepsilon\dot{\eta}] \right\} dt + o(\varepsilon) \end{aligned}$$

$$\begin{aligned}\tilde{J}(\tau + \varepsilon\theta) - \tilde{J}(\tau) &= \underbrace{\int_{\tau}^{\tau+\varepsilon\theta} \left\{ \varepsilon \frac{\partial L}{\partial x}(x)\eta + \lambda^T[-\alpha - \alpha - \varepsilon\dot{\eta}] \right\} dt}_{I_1} \\ &\quad + \underbrace{\int_{\tau+\varepsilon\theta}^1 \left\{ \varepsilon \frac{\partial L}{\partial x}(x)\eta + \lambda^T[-\varepsilon\dot{\eta}] \right\} dt}_{I_2} + o(\varepsilon)\end{aligned}\tag{2}$$

Using the mean-value theorem, the first integral is

$$\begin{aligned}I_1 &= \int_{\tau}^{\tau+\varepsilon\theta} \left\{ \varepsilon \frac{\partial L}{\partial x}(x)\eta - \lambda^T[2\alpha + \varepsilon\dot{\eta}] \right\} dt \\ &= [(\tau + \varepsilon\theta) - \tau] \cdot [-2\lambda^T(\xi)\alpha] + o(\varepsilon) \\ &= -2\varepsilon\theta\lambda^T(\xi)\alpha + o(\varepsilon)\end{aligned}\tag{3}$$

where $\tau \leq \xi \leq \tau + \varepsilon\theta$. The second integral is

$$\begin{aligned}I_2 &= \int_{\tau}^1 \left\{ \varepsilon \frac{\partial L}{\partial x}(x)\eta - \varepsilon\lambda^T\dot{\eta} \right\} dt - \underbrace{\int_{\tau}^{\tau+\varepsilon\theta} \left\{ \varepsilon \frac{\partial L}{\partial x}(x)\eta + \lambda^T[-\varepsilon\dot{\eta}] \right\} dt}_{o(\varepsilon)} \\ &= \varepsilon \int_{\tau}^1 \left[\frac{\partial L}{\partial x}(x)\eta + \dot{\lambda}^T\eta \right] dt - \varepsilon\lambda^T(1)\eta(1) + \varepsilon\lambda^T(\tau)\underbrace{\eta(\tau)}_{=0} + o(\varepsilon) \\ &= \varepsilon \int_{\tau}^1 \left[\frac{\partial L}{\partial x}(x) + \dot{\lambda}^T \right] \eta dt - \varepsilon\lambda^T(1)\eta(1) + o(\varepsilon)\end{aligned}\tag{4}$$

Substituting (3) and (4) into (2) produces

$$\begin{aligned}\tilde{J}(\tau + \varepsilon\theta) - \tilde{J}(\tau) &= -2\varepsilon\theta\lambda^T(\xi)\alpha + \varepsilon \int_{\tau}^1 \left[\frac{\partial L}{\partial x}(x) + \dot{\lambda}^T \right] \eta dt - \varepsilon\lambda^T(1)\eta(1) + o(\varepsilon) \\ \delta\tilde{J}(\tau; \theta) &= \lim_{\varepsilon \rightarrow 0} \frac{\tilde{J}(\tau + \varepsilon\theta) - \tilde{J}(\tau)}{\varepsilon} \\ &= -2\theta\lambda^T(\tau)\alpha + \int_{\tau}^1 \left[\frac{\partial L}{\partial x}(x) + \dot{\lambda}^T \right] \eta dt - \lambda^T(1)\eta(1),\end{aligned}$$

since $\xi \rightarrow \tau$ as $\varepsilon \rightarrow 0$. Select the costate such that

$$\begin{cases} \dot{\lambda}(t) = -\frac{\partial L}{\partial x}(x(t)) = -\frac{\partial}{\partial x} \frac{1}{2}(x(t) - \alpha)^2 = \alpha - x(t) \\ \lambda(1) = 0 \end{cases}$$

With this choice of $\lambda(t)$,

$$\begin{aligned}\delta\tilde{J}(\tau; \theta) &= -2\alpha\lambda(\tau)\theta = \frac{\partial \tilde{J}}{\partial \tau}\theta \\ \frac{\partial \tilde{J}}{\partial \tau} &= -2\alpha\lambda(\tau) = 0 \\ \lambda(\tau) &= 0\end{aligned}$$

From the dynamics of x ,

$$x(t) = \begin{cases} -\alpha t + 1, & t \in [0, \tau) \\ \alpha(t - \tau) - \alpha\tau + 1, & t \in [\tau, 1] \end{cases}$$

By the fundamental theorem of calculus,

$$\begin{aligned} \lambda(1) - \lambda(\tau) &= \int_{\tau}^1 \dot{\lambda}(t) dt = \int_{\tau}^1 [\alpha - x(t)] dt \\ &= \int_{\tau}^1 [\alpha - \alpha(t - \tau) + \alpha\tau - 1] dt \\ &= \left[-\frac{1}{2}\alpha t^2 + (\alpha + 2\alpha\tau - 1)t \right]_{\tau}^1 \\ &= -\frac{1}{2}\alpha(1 - \tau^2) + (\alpha + 2\alpha\tau - 1)(1 - \tau) \\ &= (1 - \tau) \left[-\frac{\alpha}{2}(1 + \tau) + \alpha + 2\alpha\tau - 1 \right] \\ &= (1 - \tau) \left[\frac{\alpha}{2} - 1 + \left(\frac{3\alpha}{2} \right) \tau \right] \end{aligned}$$

Since $\lambda(1) - \lambda(\tau) = 0 - 0 = 0$, the optimal τ is given by the zeros of the expression, i.e. $\tau = 1$ or $\tau = (2 - \alpha)/3\alpha$. Since we need $\tau \leq 1$,

$$\tau^* = \begin{cases} \frac{2 - \alpha}{3\alpha}, & \alpha \geq \frac{1}{2} \\ 1, & \alpha < \frac{1}{2} \end{cases}$$

(b) The optimal switching times and the costs are shown in the table below.

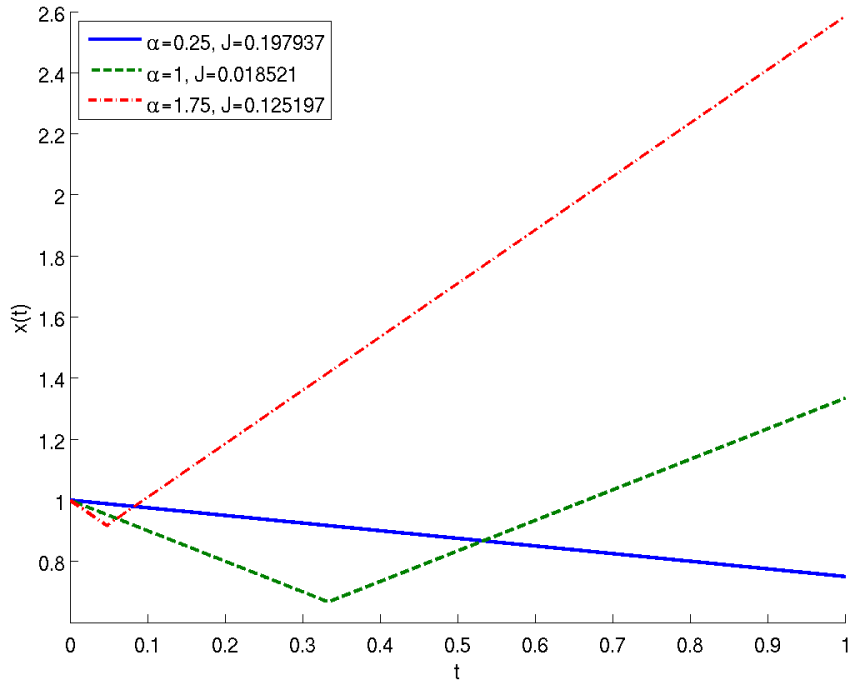
α	τ^*	$J(\tau^*)$
0.25	1	19/96
1	1/3	1/54
1.75	1/21	757/6048

The costs were calculated as follows:

$$\begin{aligned} \alpha = 0.25 \implies J(1) &= \int_0^1 \frac{1}{2}(-0.25t + 1 - 0.25)^2 dt \\ &= -\frac{1}{6 \cdot 0.75}(-0.25t + 0.75)^3 \Big|_0^1 \\ &= -\frac{1}{6 \cdot 0.25} [(-0.25 + 0.75)^3 - (0.75)^3] \\ &= \frac{19}{96} \approx 0.1979 \end{aligned}$$

$$\begin{aligned}
\alpha = 1 \implies J(1/3) &= \int_0^{1/3} \frac{1}{2}(-t + 1 - 1)^2 dt + \int_{1/3}^1 \frac{1}{2} \left(t - \frac{1}{3} - \frac{1}{3} + 1 - 1 \right)^2 dt \\
&= \frac{1}{6} t^3 \Big|_0^{1/3} + \frac{1}{6} \left(t - \frac{2}{3} \right)^3 \Big|_{1/3}^1 = \frac{1}{6} \left[\frac{1}{3^3} + \frac{1}{3^3} + \frac{1}{3^3} \right] \\
&= \frac{1}{54} \approx 0.01852 \\
\alpha = 1.75 \implies J(1/21) &= \int_0^{1/21} \frac{1}{2}(-1.75t + 1 - 1.75)^2 dt \\
&\quad + \int_{1/21}^1 \frac{1}{2} \left[1.75 \left(t - \frac{1}{21} \right) - \frac{1.75}{21} + 1 - 1.75 \right]^2 dt \\
&= -\frac{1}{6 \cdot 1.75} (-1.75t - 0.75)^3 \Big|_0^{1/21} + \frac{1}{6 \cdot 1.75} \left(1.75t - \frac{2(1.75)}{21} - 0.75 \right)^3 \Big|_{1/21}^1 \\
&= \frac{1}{6 \cdot 1.75} \left[\left(\frac{1.75}{21} + 0.75 \right)^3 - 0.75^3 + \left(1 - \frac{3.5}{21} \right)^3 - \left(-\frac{1.75}{21} - 0.75 \right)^3 \right] \\
&= \frac{757}{6048} \approx 0.1252
\end{aligned}$$

The following figure shows the plots and approximate costs from Matlab. We can see that the costs calculated by Matlab are close to the actual costs calculated by hand.



```

1 % Klaus Okkelberg
2 % ECE 6553, HW2
3 % P3b
4
5 % givens
6 alpha = [0.25, 1, 1.75];
7 x0 = 1;
8 % dynamics
9 f1 = -alpha;
10 f2 = alpha;
11 % optimal switching time
12 tau = min(1, (2-alpha)./(3*alpha));
13
14 % times
15 t = linspace(0,1,1e4)';
16 dt = t(2)-t(1);
17
18 % simulate system
19 x = zeros(numel(t),numel(alpha));
20 x(1,:) = x0;
21 for k = 1:numel(alpha)
22     for idx = 2:numel(t)
23         if t(idx)<tau(k)
24             x(idx,k) = x(idx-1,k) + f1(k)*dt;
25         else % t(idx)>=tau(k)
26             x(idx,k) = x(idx-1,k) + f2(k)*dt;
27         end
28     end
29 end
30 % approximate cost
31 J = sum( 1/2*bsxfun(@minus,x,alpha).^2 * dt, 1 );
32
33 % plot figures
34 figure
35 set(gcf,'DefaultLineLineWidth',2)
36 set(gcf,'DefaultAxesFontSize',12)
37 linespec = {'-', '—', '-.'};
38 hold all
39 for idx = 1:numel(alpha)
40     plot(t,x(:,idx),linespec{idx})
41 end
42 legend(arrayfun(@(x,y) sprintf('\alpha=%g, J=%f',x,y),alpha,J,'uni',0),...
43         'Location','NorthWest')
44 xlabel('t')
45 ylabel('x(t)')

```

4. (a) As shown in class, the costate equation only depends on the partial derivative of the Hamiltonian with respect to $x(t)$. The fact that the control is x_0 does not matter. Therefore, the costate has to satisfy

$$\begin{aligned}\dot{\lambda} &= -\frac{\partial H^T}{\partial x} = -\frac{\partial L^T}{\partial x} - \frac{\partial f^T}{\partial x}\lambda \\ &= -\left(\frac{\partial}{\partial x}\frac{1}{2}x^T Qx\right)^T - \left(\frac{\partial}{\partial x}Ax\right)^T \lambda \\ &= -Qx - A^T \lambda \\ \lambda(T) &= \frac{\partial \Psi^T}{\partial x} = 0 \\ \Rightarrow \begin{cases} \dot{\lambda} = -Qx - A^T \lambda \\ \lambda(T) = 0 \end{cases}\end{aligned}$$

- (b) A Hurwitz matrix is defined as a square matrix whose eigenvalues have strictly negative real parts. From the dynamics of $\dot{\lambda}$, we see that the costate is governed by $-A^T$, whose eigenvalues would have strictly positive real parts, since eigenvalues are given by

$$0 = \det(A - \lambda I) = -\det(-(A - \lambda I)^T) = -\det(-A^T - (-\lambda)I)$$

Therefore, the costate equation is unstable.

- (c) Since the costate equation is unstable, we need certain $-Qx$ to stabilize the system in order to satisfy the boundary condition $\lambda(0) = 0$. However, this is untrue for general x . Therefore, the numerical solution to this problem is sensitive to the initial $x_{0,0}$. This means the solution is not robust to the initial point chosen. Additionally, the gradient $\lambda(0)$ is likely to be large due to the instability, so the costate converges quickly for certain initial $x_{0,0}$. This means the numerical solution can be very efficient.
5. This problem is equivalent to the following unconstrained optimization problem, where the constraint has been substituted into the utility function:

$$\max_{T,c} F(T,c) = \gamma(T+c) - \frac{1}{2}\alpha T^2 - \frac{1}{2}\beta c^2$$

The FONCs are

$$\begin{aligned}\frac{\partial F}{\partial T} &= \gamma - \alpha T = 0 \Rightarrow T = \frac{\gamma}{\alpha} \\ \frac{\partial F}{\partial c} &= \gamma - \beta c = 0 \Rightarrow c = \frac{\gamma}{\beta}\end{aligned}$$

For this to be a maximum instead of a minimum, it is sufficient that the Hessian is negative definite:

$$\begin{bmatrix} \frac{\partial^2 F}{\partial T^2} & \frac{\partial^2 F}{\partial T \partial c} \\ \frac{\partial^2 F}{\partial c \partial T} & \frac{\partial^2 F}{\partial c^2} \end{bmatrix} = \begin{bmatrix} -\alpha & 0 \\ 0 & -\beta \end{bmatrix} \prec 0,$$

since the Hessian is Hermitian and has strictly negative eigenvalues. Therefore, the optimal grade is

$$g^* = \gamma(T+c) = \gamma\left(\frac{\gamma}{\alpha} + \frac{\gamma}{\beta}\right) = \gamma^2\left(\frac{\alpha+\beta}{\alpha\beta}\right)$$

6. The augmented cost is

$$\tilde{J}(u) = \int_0^T [L(x, u, t) + \lambda^T \Phi(x, u, t)] dt$$

Consider a perturbation $u \mapsto u + \varepsilon \nu$ that causes a change $x \mapsto x + \varepsilon \eta$:

$$\begin{aligned} \tilde{J}(u + \varepsilon \nu) &= \int_0^T [L(x + \varepsilon \eta, u + \varepsilon \nu, t) + \lambda^T \Phi(x + \varepsilon \eta, u + \varepsilon \nu, t)] dt \\ &= \int_0^T \left\{ L(x, u, t) + \varepsilon \frac{\partial L}{\partial x}(x, u, t) \eta + \varepsilon \frac{\partial L}{\partial u}(x, u, t) \nu \right. \\ &\quad \left. + \lambda^T \left[\Phi(x, u, t) + \varepsilon \frac{\partial \Phi}{\partial x}(x, u, t) \eta + \varepsilon \frac{\partial \Phi}{\partial u}(x, u, t) \nu \right] \right\} dt + o(\varepsilon) \\ \tilde{J}(u + \varepsilon \nu) - \tilde{J}(u) &= \int_0^T \left\{ \varepsilon \frac{\partial L}{\partial x}(x, u, t) \eta + \varepsilon \frac{\partial L}{\partial u}(x, u, t) \nu \right. \\ &\quad \left. + \lambda^T \left[\varepsilon \frac{\partial \Phi}{\partial x}(x, u, t) \eta + \varepsilon \frac{\partial \Phi}{\partial u}(x, u, t) \nu \right] \right\} dt + o(\varepsilon) \\ \delta \tilde{J}(u; \nu) &= \int_0^T \left\{ \frac{\partial L}{\partial x}(x, u, t) \eta + \frac{\partial L}{\partial u}(x, u, t) \nu + \lambda^T \left[\frac{\partial \Phi}{\partial x}(x, u, t) \eta + \frac{\partial \Phi}{\partial u}(x, u, t) \nu \right] \right\} dt \\ &= \int_0^T \left[\frac{\partial L}{\partial u}(x, u, t) + \lambda^T \frac{\partial \Phi}{\partial u}(x, u, t) \right] \nu dt + \int_0^T \left[\frac{\partial L}{\partial x}(x, u, t) + \lambda^T \frac{\partial \Phi}{\partial x}(x, u, t) \right] \eta dt \end{aligned}$$

Since for optimality, we want $\delta \tilde{J}(u; \nu) = 0$ for all ν and η , we need

$$\begin{cases} \frac{\partial L}{\partial u}(x, u, t) + \lambda^T \frac{\partial \Phi}{\partial u}(x, u, t) = 0, & \forall t \in [0, T] \\ \frac{\partial L}{\partial x}(x, u, t) + \lambda^T \frac{\partial \Phi}{\partial x}(x, u, t) = 0, & \forall t \in [0, T] \end{cases}$$