Calculus of variations and the Euler - Lagrange equations

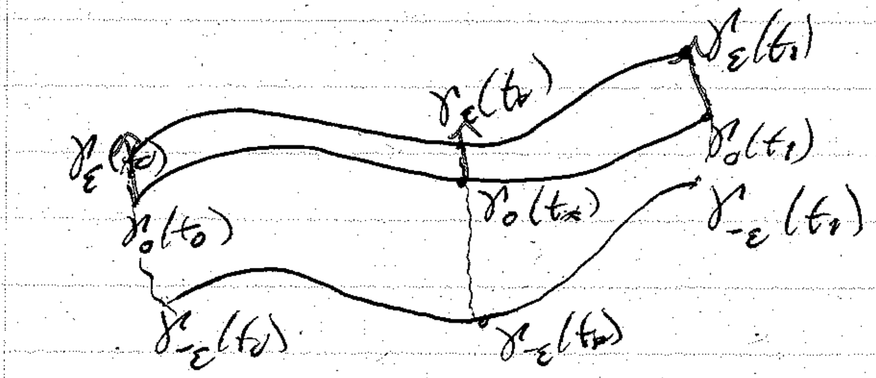
In the calculus of variations, we seek optimal paths - the best way from here to there, with best being interpreted as minimizing some cost.

In optimal control, we are designing controls that bring about the desired outcome with minimal cost. In Lagrangian mechanics, systems are proposed with the tacit philosophical assumption that "nature optimizes"; in fact, the Euler-Lagrange eqs. determine critical points - these could be saddles or maxima!

General variational set-up on $\mathbb{R}^n$:
Let $C(t_0, t_1)$ denote the set of smooth curves $\gamma : [t_0, t_1] \to \mathbb{R}^n$.

Given points $q_0, q_1 \in \mathbb{R}^n$, let $C(t_0, t_1; q_0, q_1) \subseteq C(t_0, t_1)$ be the subset of $C(t_0, t_1)$ consisting of curves γ satisfying $\gamma(t_0) = q_0$ and $\gamma(t_1) = q_1$.

Let γ_ε be a curve in $\mathcal{C}(t_0, t_1)$, $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$



For each $t \in [t_0, t_1]$, $\gamma_\varepsilon(t)$ is a curve in $\mathbb{R}^n$ parametrized by $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$.

$\frac{d}{d\varepsilon}\gamma_\varepsilon(t)\Big|_{\varepsilon=0} \in T_{\gamma_0(t)}\mathbb{R}^n \ (\simeq \mathbb{R}^n)$.

Given a vector field X on a neighborhood of $\gamma_0([t_0, t_1])$, there is a curve γ_ε in $\mathcal{C}(t_0, t_1)$ such that $\frac{d}{d\varepsilon}\gamma_\varepsilon(t)\Big|_{\varepsilon=0} = X(\gamma_0(t))$ $\forall\, t \in [t_0, t_1]$.

If γ_ε is a curve in $\mathcal{C}(t_0, t_1; q_0, q_1)$, then $\frac{d}{d\varepsilon}\gamma_\varepsilon(t_0)\Big|_{\varepsilon=0} = \underline{0} = \frac{d}{d\varepsilon}\gamma_\varepsilon(t_1)\Big|_{\varepsilon=0}$ (obvious, but important!).

We'll identify $T\mathbb{R}^n$ with $\mathbb{R}^n \times \mathbb{R}^n$ — finding local analogues of this identification is the crucial "trick" in extending the calculus of variations to non-linear manifolds "bare hands":

Given a function $L: \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, define the functional $\mathcal{L}: \mathcal{C}(t_0, t_1) \to \mathbb{R}$ by

$$\mathcal{L}(\gamma) := \int_{t_0}^{t_1} L(\underset{\sim}{\gamma}(t), \dot{\underset{\sim}{\gamma}}(t), t)\, dt.$$

What are the critical points of $\mathcal{L}$?

Consider a curve $\underset{\sim}{\gamma}_\varepsilon$ in $\mathcal{C}(t_0, t_1)$:

$$\frac{d}{d\varepsilon} \mathcal{L}(\underset{\sim}{\gamma}_\varepsilon)\Big|_{\varepsilon=0} = \frac{d}{d\varepsilon} \int_{t_0}^{t_1} L(\underset{\sim}{\gamma}_\varepsilon(t), \dot{\underset{\sim}{\gamma}}_\varepsilon(t), t)\, dt \Big|_{\varepsilon=0}$$

$$= \int_{t_0}^{t_1} \Big(\Big\langle \frac{\partial L}{\partial \underset{\sim}{q}}(\underset{\sim}{\gamma}_0(t), \dot{\gamma}_0(t), t), \delta\underset{\sim}{\gamma}(t) \Big\rangle$$

$$+ \Big\langle \frac{\partial L}{\partial \underset{\sim}{v}}(\underset{\sim}{\gamma}_0(t), \dot{\underset{\sim}{\gamma}}_0(t), t), \delta\dot{\underset{\sim}{\gamma}}(t) \Big\rangle \Big) dt$$

where $\delta\underset{\sim}{\gamma}(t) := \frac{d}{d\varepsilon} \underset{\sim}{\gamma}_\varepsilon(t)\big|_{\varepsilon=0}$.

Integrating the second term by parts gives

$$\int_{t_0}^{t_1} \Big\langle \frac{\partial L}{\partial \underset{\sim}{v}}(\underset{\sim}{\gamma}_0(t), \dot{\underset{\sim}{\gamma}}_0(t), t), \delta\dot{\underset{\sim}{\gamma}}(t) \Big\rangle dt$$

$$= \Big\langle \frac{\partial L}{\partial \underset{\sim}{v}}(\underset{\sim}{\gamma}_0(t_1), \dot{\underset{\sim}{\gamma}}_0(t_1), t_1), \delta\underset{\sim}{\gamma}(t_1) \Big\rangle$$

$$- \Big\langle \cdots t_0 \cdots t_0, t_0), \cdots (t_0) \Big\rangle$$

$$- \int_{t_0}^{t_1} \Big\langle \frac{d}{dt}\Big(\frac{\partial L}{\partial \underset{\sim}{v}}(\underset{\sim}{\gamma}_0(t), \dot{\underset{\sim}{\gamma}}_0(t), t) \Big), \delta\underset{\sim}{\gamma}(t) \Big\rangle dt$$

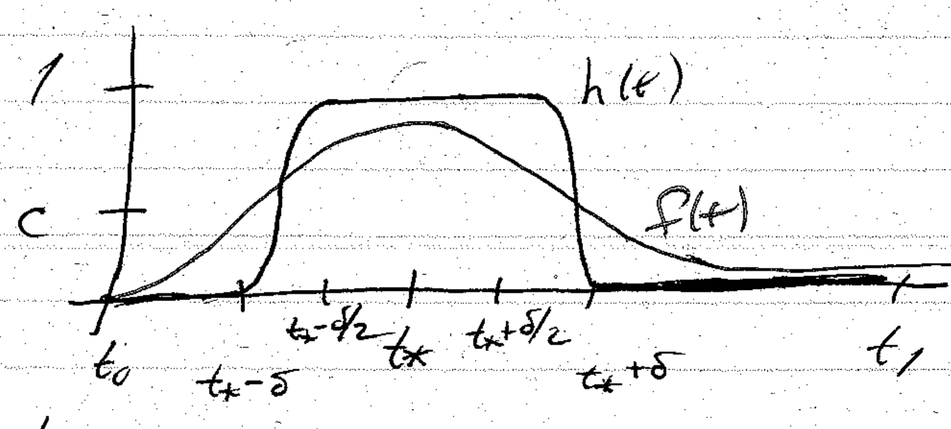
Use the following lemma:

If $f:[t_0,t_1]\to\mathbb{R}$ is continuous and $\int_{t_0}^{t_1} f(t)h(t) = 0$ for all smooth functions $h:[t_0,t_1]\to\mathbb{R}$ with $h(t_0)=h(t_1)=0$, then $f\equiv 0$.

Pf: Assume $\exists\, t_* \in (t_0,t_1)$ such that $f(t_*)\neq 0$, say $f(t_*)>0$.

Continuity of $f \Rightarrow \exists\, c>0,\ \delta>0$ such that $|t-t_*|<\delta \Rightarrow f(t)\geq c$. $\oplus$

There exists a smooth function $h:[t_0,t_1]\to\mathbb{R}$ such that
i) $h\geq 0$
ii) $|t-t_*|>\delta \Rightarrow h(t)=0$
iii) $|t-t_*|<\frac{\delta}{2} \Rightarrow h(t)=1$
(take δ suff. small that $t_*+\delta/2 < t_1$)



$$\int_{t_0}^{t_1} f(t)h(t)\,dt = \int_{t_*-\delta}^{t_*+\delta} f(t)h(t)\,dt$$

because of ii)

$$\geq c \int_{t_*-\delta}^{t_*+\delta} h(t)\,dt \quad \text{because of } (*)$$

$$\geq c\delta \quad \text{because of } i), iii).$$

$\lightning \Rightarrow f|_{(t_0,t_1)} \equiv 0$

Continuity of $f \Rightarrow f \equiv 0$.

Assume $\underline{\gamma}_0$ is a critical point of $\mathcal{L}$.
Consider first a variation $\underline{\gamma}_\varepsilon$ of $\underline{\gamma}_0$ in $\mathcal{C}(t_0, t_1; \underline{\gamma}_0(t_0), \underline{\gamma}_0(t_1))$:
The boundary terms are trivial, so

$$0 = \frac{d}{d\varepsilon}\mathcal{L}(\underline{\gamma}_\varepsilon)\Big|_{\varepsilon=0}$$

$$= \int_{t_0}^{t_1} \Big\langle \frac{\partial L}{\partial q}(\underline{\gamma}_0(t), \dot{\underline{\gamma}}_0(t), t) - \frac{d}{dt}\Big(\frac{\partial L}{\partial v}(\underline{\gamma}_0(t), \dot{\underline{\gamma}}_0(t), t)\Big), \frac{d}{d\varepsilon}\underline{\gamma}_\varepsilon(t)\Big|_{\varepsilon=0}\Big\rangle\, dt.$$

Given an arbitrary smooth function h with $h(t_0) = h(t_1) = 0$, we can construct a curve $\underline{\gamma}_\varepsilon(t)$ such that the j-th component of $\frac{d}{d\varepsilon}\underline{\gamma}_\varepsilon(t)\big|_{\varepsilon=0} = h(t)$ and the other components are zero.
Hence the lemma implies that

$$\frac{\partial L}{\partial q} = \frac{d}{dt}\frac{\partial L}{\partial v} \qquad \text{Euler-Lagrange equation.}$$

Now consider γ_ε in which the end points vary (if permissible):
Since γ_0 must satisfy the Euler-Lagrange equations,

$$0 = \frac{d}{d\varepsilon} \mathcal{L}(\gamma_\varepsilon)\Big|_{\varepsilon=0}$$
$$= \left\langle \frac{\partial L}{\partial \underline{v}}(\underline{\gamma}_0(t_1), \dot{\underline{\gamma}}_0(t_1), t_1), \delta\underline{\gamma}(t_1) \right\rangle$$
$$- \left\langle \frac{\partial L}{\partial \underline{v}}(\gamma_0(t_0), \dot{\gamma}_0(t_0), t_0), \delta\gamma(t_0) \right\rangle.$$

If only one end point may vary, the boundary conditions involve only one term.

Example: $L(q, \underline{v}, t) = \frac{m}{2}\|\underline{v}\|^2 - V(q)$
for some $m \in \mathbb{R}^+$, $V: \mathbb{R}^n \to \mathbb{R}$;
fixed end points:

$$\frac{\partial L}{\partial q}(q, \dot{q}, t) = -\nabla V(q) \quad \frac{\partial L}{\partial \underline{v}}(q, \dot{q}, t) = m\dot{q},$$

so E-L eqs are $m\ddot{q} = \frac{d}{dt} m\dot{q} = \frac{\partial L}{\partial q} = -\nabla V(q)$

Equations for a particle of mass m in a conservative force field with potential V.

Important point: The Euler-Lagrange equations determine a 2^{nd} order ODE for q only if L is appropriately non-degenerate:

$$\frac{d}{dt}\frac{\partial L}{\partial \underline{v}}(q(t), \dot{q}(t), t)$$

$$= \frac{\partial^2 L}{\partial \underline{v}\,\partial q}\dot{q} + \frac{\partial^2 L}{\partial \underline{v}^2}\ddot{q} + \frac{\partial^2 L}{\partial \underline{v}\,\partial t},$$

so $\frac{d}{dt}\frac{\partial L}{\partial \underline{v}} = \frac{\partial L}{\partial q}$ determines $\ddot{q}$

$\Longleftrightarrow$ $\frac{\partial^2 L}{\partial \underline{v}^2}(q, \dot{q}, t)$ is invertible

(We'll go into this more when considering canonical Hamiltonian systems and their relation to the E-L equations on manifolds.)

Optimal control

Consider a parametrized ODE on $\mathbb{R}^n$:

$$\dot{q} = X(q, \chi, t) \qquad \chi \in \mathbb{R}^k \quad \circledast$$

Given a cost function $C : \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R} \to \mathbb{R}$, we seek curves $(q(t), \chi(t))$, $t \in [t_0, t_1]$, such that $\int_{t_0}^{t_1} C(q(t), \chi(t), t)\, dt$ is minimized among all curves satisfying $\circledast$ and given boundary conditions.

We can minimize the integral of C on the (typically nonlinear) set of curves satisfying $\circledast$ and the boundary conditions or we can enforce $\circledast$ using a Lagrange multiplier and taking variations on the set of curves satisfying the boundary conditions.

Follow the latter approach.

Introduce the Lag. mult. $\lambda : [t_0, t_1] \to \mathbb{R}^n$

Given curves $\gamma : [t_0, t_1] \to \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^n$,

$\gamma(t) = (q(t), \chi(t), \lambda(t))$,

define
$$\mathcal{L}(\underline{\gamma}) := \int_{t_0}^{t_1} \Big(\langle \underline{\lambda}(t), \dot{q}(t) - \underline{X}(q(t), \underline{\chi}(t), t) \rangle + C(q(t), \underline{\chi}(t), t) \Big)\, dt.$$

Extrema of $\mathcal{L}$ must be critical points or, if q or $\underline{\chi}$ is constrained to lie in a closed subset of $\mathbb{R}^n$ or $\mathbb{R}^k$, boundary points.

Given a curve $\underline{\gamma}_\varepsilon$, with $\frac{d}{d\varepsilon} q_\varepsilon(t)\big|_{\varepsilon=0} = \delta q(t)$,
$\frac{d}{d\varepsilon} \underline{\chi}_\varepsilon(t)\big|_{\varepsilon=0} = \delta \underline{\chi}(t)$, etc,

$$\frac{d}{d\varepsilon} \mathcal{L}(\underline{\gamma}_\varepsilon)\Big|_{\varepsilon=0} =$$
$$\int_{t_0}^{t_1} \Big(\langle \delta\underline{\lambda}(t), \dot{q}(t) - \underline{X}(q(t), \underline{\chi}(t), t) \rangle + \langle \underline{\lambda}(t), \delta\dot{q}(t) - \frac{\partial \underline{X}}{\partial q}(q(t), \underline{\chi}(t), t)\, \delta q(t) - \frac{\partial \underline{X}}{\partial \underline{\chi}}(q(t), \underline{\chi}(t), t)\, \delta \underline{\chi}(t) \rangle$$
$$+ \langle \frac{\partial C}{\partial q}(q(t), \underline{\chi}(t), t), \delta q(t) \rangle + \langle \frac{\partial C}{\partial \underline{\chi}}(q(t), \underline{\chi}(t), t), \delta \underline{\chi}(t) \rangle \Big)\, dt.$$

Since $\lambda(t)$ is arbitrary, criticality w.r.t. variations of λ enforces $(*)$.

Integrating the term $\int_{t_0}^{t_1} \langle \lambda(t), \delta \dot{q}(t) \rangle dt$ by parts yields

$$\langle \lambda(t_1), \delta q(t_1) \rangle - \langle \lambda(t_0), \delta q(t_0) \rangle - \int_{t_0}^{t_1} \langle \dot{\lambda}(t), \delta q(t) \rangle dt,$$

so criticality w.r.t variations of q implies
$$\dot{\lambda} + \left(\frac{\partial X}{\partial q}(q, \chi, t)\right)^T \lambda = \frac{\partial c}{\partial q}(q, \chi, t),$$

If $q(t_j)$ may vary, then $\lambda(t_j) = 0$, $j = 0, 1$.

Finally, criticality w.r.t. variations of χ implies
$$\left(\frac{\partial X}{\partial \chi}(q, \chi, t)\right)^T \lambda = \frac{\partial c}{\partial \chi}(q, \chi, t).$$

Note: If this last (algebraic) equation can be solved for χ as a function of q, λ and t, then the solution $\chi(q, \lambda, t)$ can be substituted into the differential eqs for q and λ yielding a 1st order system of ODEs on $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$.

Example : $n=2,\ k=1$.

$\ddot{x} = \chi$, i.e. $\begin{array}{l}\dot{x}=v\\ \dot{v}=\chi\end{array}$, $C(x,v,\chi,t)=\frac{\chi^2}{2}$.

$$\mathcal{L}(\underset{\sim}{\gamma}) = \int_{t_0}^{t_1} \left(\langle \underset{\sim}{\lambda}, (v,\chi) \rangle + \frac{\chi^2}{2} \right) dt$$

$$\underset{\sim}{X}(\underbrace{x,v}_{q},\chi,t) = \cancel{(v,\chi)} \begin{pmatrix} v \\ \chi \end{pmatrix}$$

$$\Rightarrow \quad \frac{\partial \underset{\sim}{X}}{\partial q}(x,v,\chi,t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$\frac{\partial \underset{\sim}{X}}{\partial \chi}(\qquad) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\frac{\partial C}{\partial q} \equiv \underset{\sim}{0} \quad , \frac{\partial C}{\partial \chi} = \chi .$$

$$\Rightarrow \quad \chi = \frac{\partial C}{\partial \chi} = \left(\frac{\partial \underset{\sim}{X}}{\partial \chi}(q,\chi,t) \right)^T \underset{\sim}{\lambda} = \lambda_2$$

$$\underset{\sim}{0} = \frac{\partial C}{\partial q}(q,\chi,t) = \dot{\underset{\sim}{\lambda}} + \left(\frac{\partial \underset{\sim}{X}}{\partial q}(q,\chi,t) \right)^T \underset{\sim}{\lambda}$$

$$= \dot{\underset{\sim}{\lambda}} + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \underset{\sim}{\lambda} ,$$

i.e. $\dot{\lambda}_1 = 0$, $\dot{\lambda}_2 = -\lambda_1$,

so $\underset{\chi(t)}{\overset{\parallel}{\lambda_2(t)}} = \lambda_{20} - \lambda_{10} t$.

Hence $\dot{v} = \chi \Rightarrow v(t) = \int^t \chi(\tau)\, d\tau$

$\dot{x} = v \Rightarrow x(t) = \int^t v(\tau)\, d\tau$.

Special case: Add boundary conditions $x(t_0) = x_0$, $x(t_1) = x_1$, $v(t_0) = v(t_1) = 0$.

Let $T = t_1 - t_0$.

$$0 = v(t_1) = \int_0^T x(\tau)\,d\tau = T\left(\lambda_{20} - \frac{\lambda_{10}}{2}T\right)$$

$$\Rightarrow \lambda_{20} = \frac{\lambda_{10}}{2}T, \text{ and hence}$$

$$v(t+t_0) = \frac{\lambda_{10}}{2}\,t\,(T-t)$$

$$x(t+t_0) - x_0 = \int_0^t v(\tau + t_0)\,d\tau$$

$$= \frac{\lambda_{10}\,t^2}{2}\left(T/2 - t/3\right).$$

$$x_1 = x(t_1) = x_0 + \frac{\lambda_{10}T^3}{12}$$

$$\Rightarrow \lambda_{10} = \frac{12(x_1 - x_0)}{T^3}$$

What is the actual cost of the maneuver?

$$C(\gamma_{opt}) = \frac{1}{2}\int_{t_0}^{t_1} x(t)^2\,dt$$

$$= \frac{1}{2}\int_{0}^{t_1 - t_0} \left(\lambda_{10}\left(T/2 - t\right)\right)^2 dt$$

$$= \frac{\lambda_{10}^2}{2}\,\frac{T^3}{12}$$

$$= \frac{6(x_1 - x_0)^2}{T^3}$$

Note: For fixed x_0, x_1, $C(\gamma_{opt}) \to 0$ as $T \to \infty$.

Question: If the magnitude of the control satisfies $|\chi| \leq \chi_{max}$, which of these quadratic cost minimizing trajectories takes the least time? What is this shortest time?

$$\chi_{max} \geq |\chi(t+t_0)| = \frac{12}{T^3}|x_1-x_0|\,|T/2 - t|$$

$\forall\, t \in [0,T]$; max value of χ is achieved at end points.

$$\Rightarrow \chi_{max} \geq \frac{6}{T^2}|x_1 - x_0|.$$

Hence min time trajectory has $|\chi(t_0)| = |\chi(t_1)| = \chi_{max}$

$$T = \sqrt{\frac{6|x_1-x_0|}{\chi_{max}}}.$$

<u>Example 2</u>: $\ddot{x} = \chi$, $x(t_0) = x_0$, $x(t_1) = x_1$, $\dot{x}(t_0) = \dot{x}(t_1) = 0$, as above, but now allow t_1 to vary and take $C \equiv 1$, $|\chi| \leq \chi_{max}$

Time minimization problem!

$\frac{\partial X}{\partial q} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\frac{\partial X}{\partial \chi} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ as before,

but now $\frac{\partial C}{\partial q} \equiv 0$ and $\frac{\partial C}{\partial \chi} \equiv 0$.

Criticality conditions $\Rightarrow \underset{\sim}{\lambda} \equiv \underset{\sim}{0}$, but say nothing about χ — look for optimal χ at end points of the admissible interval $[-\chi_{max}, \chi_{max}]$.

Can show that the "bang-bang" solution:

$$\chi(t) = \begin{cases} \chi_{max} & 0 \le t-t_0 \le T/2 \\ -\chi_{max} & T/2 < t-t_0 \le T \end{cases}$$

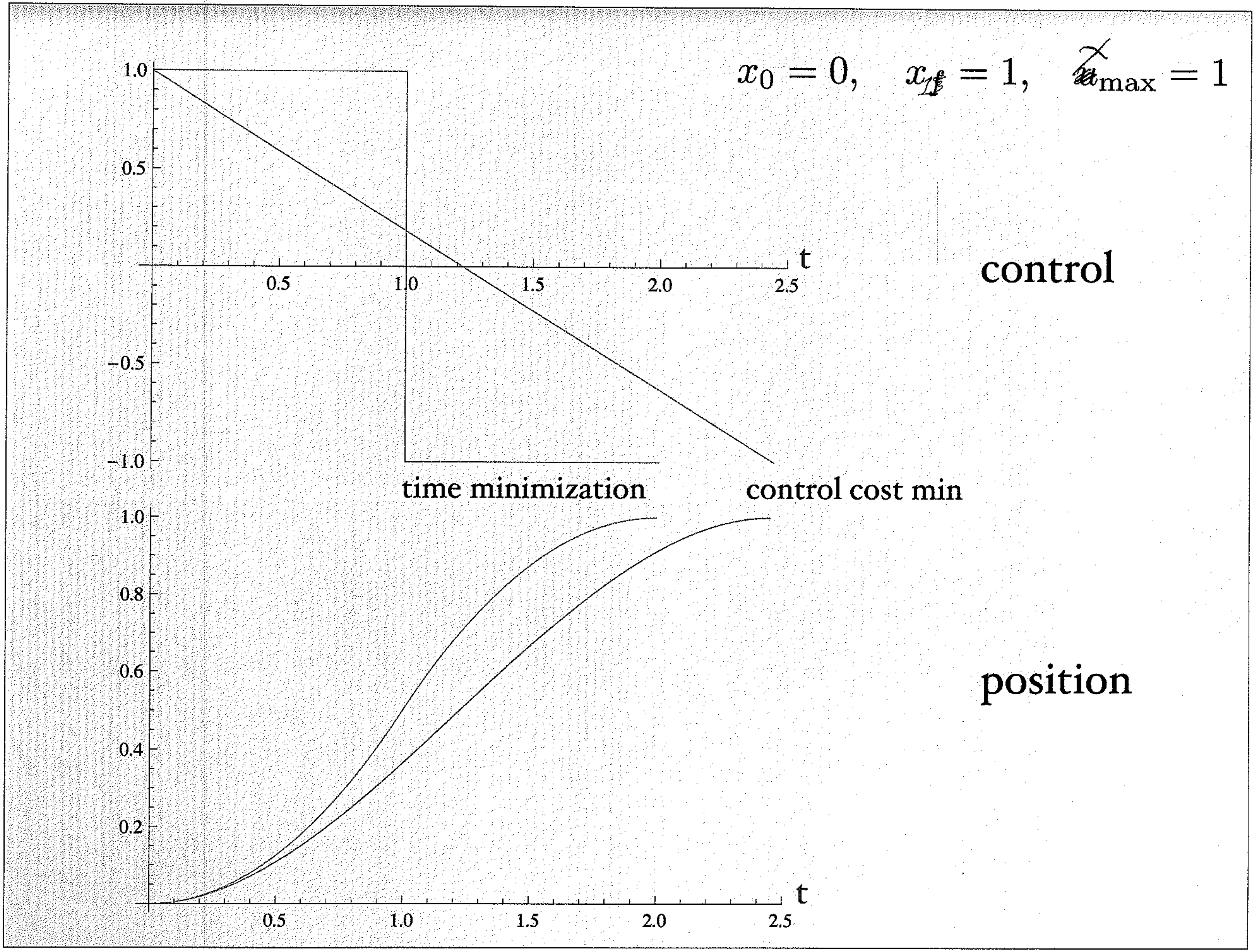
(assuming $x_1 > x_0$)

is optimal. (See, e.g. Kirk, *Optimal Control Theory*.)

If $\tau := \frac{t-t_0}{t_1-t_0}$, then $\frac{x-x_0}{\chi_{max}(t_1-t_0)^2} = \frac{1}{2}\begin{cases} \tau^2 & 0 \le \tau \le 1/2 \\ \frac{1}{2}-(1-\tau)^2 & 1/2 < \tau \le 1 \end{cases}$

is optimal; $t_1 - t_0 = 2\sqrt{\frac{x_1-x_0}{\chi_{max}}}$.

The fastest quadratic cost min. solution takes $\sqrt{3/2}$ times as long, but has $\frac{1}{\sqrt{6}}$ the quad. control cost $\int_{t_0}^{t_1} \frac{\chi^2}{2}\, dt$.



$x_0 = 0, \quad x_{1f} = 1, \quad \cancel{u}_{\max} = 1$
control
position
time minimization
control cost min
t
1.0
0.5
−0.5
−1.0
0.5
1.0
1.5
2.0
2.5
0.8
0.6
0.4
0.2