

Review:

Suppose y is a piecewise- C^1 curve and is a strong extremal

for the Basic VP.

piecewise- C^1 : everywhere C^1 except finitely many points
that is only C^0 (continuous).

1) it is a weak extremal between each two corner points.

$$\Rightarrow \text{must satisfy E-L eq.} \Rightarrow \begin{cases} y' = H_p \\ p' = -H_y \end{cases}$$

2) Weierstrass-Erdmann corner condition for strong extremum y :

$p = L_z$ and $H = p \cdot y' - L$ (momentum and Hamiltonian)
are continuous at corner points.

3) we can also generalize the Legendre's condition to
corner points for a strong minimum:

$$L_{zz}(x, y(x), y'(x^\pm)) \geq 0, \quad \forall x \in [a, b].$$

4) Weierstrass-excess condition for strong minimum y :

$$\text{define } E(x, y, z, w) \triangleq L(x, y, w) - L(x, y, z) - (w - z) \cdot L_z(x, y, z)$$

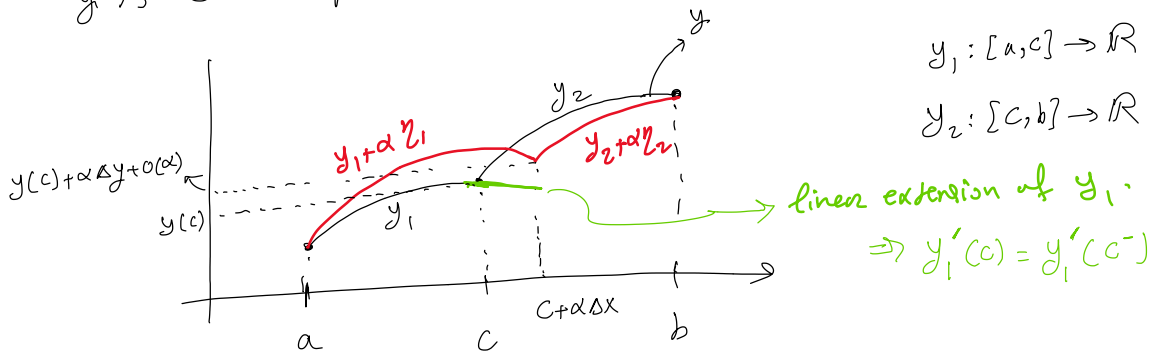
then $E(x, y(x), y'(x), w) \geq 0, \quad \forall x \in [a, b]$ non-corner points.

$$\forall w \in \mathbb{R}^n \text{ s.t. } |y' - w| < \varepsilon.$$

Geometrically, this means $L(x, y(x), \cdot)$ is locally convex at $y'(x)$.

→ let's try to argue the Weierstrass-Erdmann corner conditions:

y is C^1 except at one point $c \in [a, b]$.



→ $y_1 + \alpha \eta_1$ is the perturbed curve on $[a, c + \alpha \Delta x]$ where y_1 is extended linearly beyond c (green line) and

$$\underline{\eta_1(c) = 0}$$

→ $y_2 + \alpha \eta_2$ is defined similarly and $\underline{\eta_2(b) = 0}$

→ To get a piecewise C^1 curve:

$$\textcircled{I} \quad y_1(c + \alpha \Delta x) + \alpha \eta_1(c + \alpha \Delta x) = y_2(c + \alpha \Delta x) + \alpha \eta_2(c + \alpha \Delta x).$$

$$\triangleq y(c) + \alpha \underline{\Delta y} + o(\alpha).$$

where Δy is defined, and is independent of Δx .

* Note: we have perturbations η_1 , η_2 , Δx , Δy and they are all conditioned on these equalities to be admissible! we will simplify these later!

* Note: the entire curve is parameterized by the same α , and we denote it by $y(\cdot, \alpha): [a, b] \rightarrow \mathbb{R}$ which is C^1 except at $c + \alpha \Delta x$.

→ Lets consider the first variations of J with respect to $y(\cdot, \alpha)$:

$$J(y(\cdot, \alpha)) = \underbrace{\int_a^{c+\alpha\Delta x} L(x, y_1, y_1') dx}_{\triangleq J_1(y_1 + \alpha\eta_1)} + \underbrace{\int_{c+\alpha\Delta x}^b L(x, y_2, y_2') dx}_{\triangleq J_2(y_2 + \alpha\eta_2)}$$

Use Leibniz rule to get:

$$\delta J_1|_{y_1}(\eta_1) = \frac{d}{d\alpha} \Big|_{\alpha=0} J_1(y_1 + \alpha\eta_1)$$

$$= \int_a^c \left(L_y|_{y_1} \eta_1 + L_{y_1'}|_{y_1} \eta_1' \right) dx + L(c, y(c), y'(c)) \cdot \Delta x$$

$$= \int_a^b \left(L_y(x, y_1, y_1') - \frac{d}{dx} L_{y_1'}(x, y_1, y_1') \right) \eta_1 dx$$

$$+ L_2(c, y_1(c), y_1'(c)) \eta_1(c) + L(c, y_1(c), y_1'(c)) \cdot \Delta x$$

Integration by parts and

$$\left\{ \begin{array}{l} \eta_1(c) \neq 0 \\ y_1'(c) = y_1'(c^-) \end{array} \right.$$

Similarly,

$$\delta \mathcal{J}_2|_{y_2}(\eta_2) = \int_c^b \left(L_y(x, y_2, y_2') - \frac{d}{dx} L_z(x, y_2, y_2') \right) \cdot \eta_2 dx \\ - L_z(c, y(c), y'(c^+)) \eta_2(c) - L(c, y(c), y'(c^+)) \Delta x.$$

So :

$$0 = \frac{d}{d\alpha} \Big|_{\alpha=0} \mathcal{J}(y(\cdot, \alpha)) = \frac{d}{d\alpha} \Big|_{\alpha=0} \left(\mathcal{J}_1(y_1 + \alpha \eta_1) + \mathcal{J}_2(y_2 + \alpha \eta_2) \right) \\ = \delta \mathcal{J}_1|_{y_1}(\eta_1) + \delta \mathcal{J}_2|_{y_2}(\eta_2),$$

(II)

$\forall$ admissible perturbations $(\eta_1, \eta_2, \Delta x, \Delta y)$.

$\rightarrow$ Choose the subset of perturbations with $\Delta x = \Delta y = \eta_1(c) = \eta_2(c) = 0$,

then we must have E-L eq for y_1 on $[a, c]$ and y_2 on $[c, b]$.

$\rightarrow$ but, now that y_1 and y_2 should satisfy E-L. eq. for all perturbations, condition (II) reads:

$$(III) \quad L_z(c, y(c), y'(c^-)) \eta_1(c) - L_z(c, y(c), y'(c^+)) \eta_2(c) \\ + L(c, y(c), y'(c^-)) \Delta x - L(c, y(c), y'(c^+)) \Delta x = 0 \\ \forall \text{ admissible } (\eta_1, \eta_2, \Delta x, \Delta y).$$

→ Now, we need to characterize the admissible set better from (I):

look at first variations of (I) w.r.t. α at 0:

$$\underbrace{y'(c) = y'_1(c^-)}_{\hookrightarrow} \left[y'(c^-) \Delta x + \eta_1(c) = y'(c^+) \Delta x + \eta_2(c) = \Delta y \right]$$

this is necessary for perturbations to be admissible,

but is enough to eliminate η_1, η_2 from (III) to get:

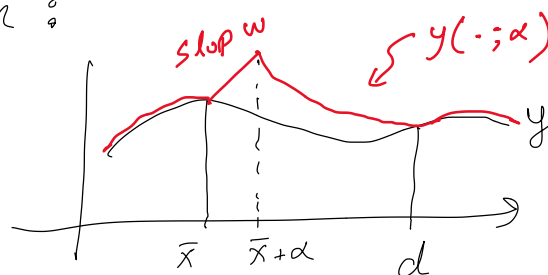
$$- L_2(x, y(x), y'(x)) \Big|_{c^-}^{c^+} \cdot \Delta y + \left(L_2(x, y(x), y'(x)) \cdot y'(x) - L(x, y(x), y'(x)) \right) \Big|_{c^-}^{c^+} \cdot \Delta x = 0$$

$\forall$ admissible $\Delta x, \Delta y$ which are independent and totally arbitrary!

$$\Rightarrow \int_{c^-}^{c^+} L_2 = 0 \text{ and } (L_2 \cdot y' - L) \Big|_{c^-}^{c^+} = 0$$

let's argue about the Weierstrass excess condition.

suppose y is a strong minimum. pick $\bar{x} \in [a, b]$ that is not a corner point. we will design a perturbation of y as $y(\cdot; \alpha)$ that is close to y in 0-norm:



$$y(x; \alpha) = \begin{cases} y(x) & x \leq \bar{x} \\ y(\bar{x}) + w(x - \bar{x}) & \bar{x} \leq x \leq \bar{x} + \alpha \\ y(x) + \frac{d-x}{d-(\bar{x}+\alpha)} (y(\bar{x}) + w\alpha - y(\bar{x}+\alpha)) & \bar{x} + \alpha \leq x \leq d \\ y(x) & d \leq x \end{cases}$$

Note that $y(x; 0) = y(x) \quad \forall x \in [a, b]$. Now, we look at variation

$\lim_{\alpha \downarrow 0^+} \frac{d}{d\alpha} J(y(\cdot; \alpha)) = 0$ vanishing for all admissible perturbations.

$$\frac{d}{d\alpha} J(y(\cdot; \alpha)) = \frac{d}{d\alpha} \left(\underbrace{\int_{\bar{x}}^{\bar{x}+\alpha} L(x, y(x; \alpha), y_x(x; \alpha)) dx}_{\text{I}} + \underbrace{\int_{\bar{x}+\alpha}^d L(x, y(x; \alpha), y_x(x; \alpha)) dx}_{\text{II}} \right)$$

$$\text{I} \rightarrow = \frac{d}{d\alpha} \int_{\bar{x}}^{\bar{x}+\alpha} L(x, y(\bar{x}) + w(x - \bar{x}), w) dx = L(\bar{x} + \alpha, y(\bar{x}) + w\alpha, w)$$

$$\text{II} \rightarrow = -L(\bar{x} + \alpha, y(\bar{x} + \alpha; \alpha), y_x(\bar{x} + \alpha; \alpha)) + \int_{\bar{x}+\alpha}^d L_y(x, y(x; \alpha), y_x(x; \alpha)) y_x(x; \alpha) dx + \underbrace{\int_{\bar{x}+\alpha}^d L_z(x, y(x; \alpha), y_x(x; \alpha)) y_{x\alpha}(x; \alpha) dx}_{\text{III}}$$

$$\textcircled{\text{III}} \rightarrow = L_Z(x, y(x; \alpha), y_x(x; \alpha)) \cdot y_\alpha(x; \alpha) \Big|_{\bar{x} + \alpha}^d$$

Integration

by part

$$y_\alpha(x; \alpha) = \frac{d}{d\alpha} y_\alpha(x; \alpha)$$

$$- \int_{\bar{x} + \alpha}^d \frac{d}{dx} [L_Z(x, y(x; \alpha), y_x(x; \alpha))] \cdot y_\alpha(x; \alpha) dx$$

$$\Rightarrow \lim_{\alpha \downarrow 0^+} \frac{d}{d\alpha} J(y(\cdot; \alpha)) = L(\bar{x}, y(\bar{x}), w) - L(\bar{x}, y(\bar{x}; 0), y_x(\bar{x}; 0))$$

$$+ L_Z(x, y(x; 0), y_x(x; 0)) \cdot y_\alpha(x; 0) \Big|_{\bar{x}}^d$$

$$+ \int_{\bar{x}}^d \left(L_y(x, y(x; 0), y_x(x; 0)) - \frac{d}{dx} [L_Z(x, y(x; 0), y_x(x; 0))] \right) y_\alpha(x; 0) dx$$

Notice: $y(x; 0) = y(x)$, $y_x(x; 0) = y'(x)$ and

y is strong minimum $\Rightarrow y$ is weak minimum $\Rightarrow$ E-L eq. is satisfied.

$\Rightarrow$ the integral is identically zero!

Let's simplify the rest of terms:

$$y(d; \alpha) = y(d), \forall \alpha \Rightarrow y_\alpha(d, \alpha) = 0$$

$$y(\bar{x} + \alpha, \alpha) = y(\bar{x}) + w\alpha \Rightarrow y_\alpha(\bar{x} + \alpha, \alpha) = y_x(\bar{x} + \alpha, \alpha) + y_\alpha(\bar{x} + \alpha; \alpha) = w$$

$$\Rightarrow y_\alpha(\bar{x} + \alpha; \alpha) = w - y_x(\bar{x} + \alpha; \alpha)$$

So,

$$L_Z(x, y(x; 0), y_x(x; 0)) \cdot y_\alpha(x; 0) \Big|_{\bar{x}}^d = -L_Z(\bar{x}, y(\bar{x}), y'(\bar{x})) \cdot [w_\alpha y'(\bar{x})]$$

$$\Rightarrow \lim_{\alpha \rightarrow 0^+} \frac{d}{d\alpha} J(y(\cdot; \alpha)) = L(\bar{x}, y(\bar{x}), w) - L(\bar{x}, y(\bar{x}), y'(\bar{x})) - L_z(\bar{x}, y(\bar{x}), y'(\bar{x})) \cdot [w - y'(\bar{x})] = E(\bar{x}, y(\bar{x}), y'(\bar{x}), w) \geq \underline{0}$$

weierstrass excess function. 

and $\forall \bar{x}$ (n-n-corner point)
 $\forall w \in R$ (small enough).

The optimal control problem is:

system: $\dot{x} = f(t, x, u)$, $x(t_0) = x_0$ given.

t : time

x_0, t_0 : initial state and time

X : states

x_f, t_f : final/terminal state and time.

u : input

$x = x(t)$, $u = u(t)$ are functions of time

regularity conditions on f :

(1) f is continuous in t and u and c' in x .

(1) f is continuous in t and u .

(3) u is piece-wise continuous in t .

(3) can be replaced by the locally Lipschitz condition:

→ ...

$\forall D \text{ bounded } \subset \mathbb{R} \times \mathbb{R}^n \times U$, $\exists L$ s.t.

$$|f(t, x_1, u) - f(t, x_2, u)| \leq L |x_1 - x_2| \quad \forall (t, x_1, u), (t, x_2, u) \in D.$$

→ Note that f need not be differentiable w.r.t. u .

Cost functional:

$$J(u) := \int_{t_0}^{t_f} L(t, x(t), u(t)) dt + K(t_f, x_{t_f})$$

$L: \mathbb{R} \times \mathbb{R}^n \times U \rightarrow \mathbb{R} \rightsquigarrow$ "running cost"

$K: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R} \rightsquigarrow$ "terminal cost".

Note: To relate this back to our VP problem, note that this is more general than the case where $L_{vp} = L_{vp}(t, x, \dot{x})$ because $\dot{x} = f(t, x, u)$ and thus

$$L_{vp} = L_{vp}(t, x, f(t, x, u)) \triangleq L(t, x, u).$$

But, we cannot necessarily go in the opposite direction.

Note: the terminal cost can be turned into a running cost:

$$\begin{aligned} K(t_f, x_{t_f}) &= K(t_0, x_{t_0}) + \int_{t_0}^{t_f} \frac{d}{dt} K(t, x_t) dt \\ &= \underbrace{K(t_0, x_{t_0})}_{\text{constant}} + \int_{t_0}^{t_f} \underbrace{\left[K_t(t, x_t) + K_x(t, x_t) \cdot f(t, x_t, u_t) \right]}_{\triangleq L(t, x, u)} dt \end{aligned}$$

Conversely, we introduce a new state variable x^L with

$$\dot{x}^L = L(t, x, u), \quad x^L(t_0) = 0, \quad \text{and then}$$

$$J(u) = \int_{t_0}^{t_f} L(t, x, u) dt = x^L(t_f) \triangleq K\left(t_f, \begin{bmatrix} x \\ x^L \end{bmatrix}(t_f)\right)$$

with the augmented dynamics

$$\frac{d}{dt} \begin{bmatrix} x \\ x^L \end{bmatrix} = \begin{bmatrix} f(t, x, u) \\ L(t, x, u) \end{bmatrix}.$$

Note: we can get rid of time dependence by introducing a new state $x_{n+1} = t \Rightarrow \dot{x}_{n+1} = 1$ and then.

$$J(u) = \int_{t_0}^{t_f} L(x(t), u(t)) dt + K(x_f), \quad x \in \mathbb{R}^{n+1}$$

$$\text{and } \frac{d}{dt} x_t = \begin{bmatrix} f(x(t), u(t)) \\ 1 \end{bmatrix}.$$

Target set:

$\rightarrow x_0, t_0$ are always fixed.

$\rightarrow x_f, t_f$ can be fixed, free, or belong to a set. Generally, we define the "target set" to be a closed subset:

$$S \subset [t_0, \infty) \times \mathbb{R}^n$$

and define t_f as

$$t_f = \inf \left\{ t : (t, x_t) \in S \right\}$$

" - "

read it as "the first hitting time of S "

→ For now, $t_f = \infty$ is not allowed, but later we will consider these cases called "Infinite-horizon optimal control".

Examples: $S = [t_0, \infty) \times \{x_1\} \rightarrow \begin{cases} t_f \text{ is "free-time"} \\ x_f = x_1 \text{ is "fixed-endpoint"} \end{cases}$

$S = [t_0, \infty) \times M$ ← a manifold/surface

$S = \{t_1\} \times \mathbb{R}^n \rightarrow \begin{cases} t_f = t_1 \text{ is "fixed-time"} \\ x_f \text{ is "free-endpoint"} \end{cases}$

$S = \{t_1\} \times \{x_1\} \rightarrow$ fixed-time, fixed-endpoint.

$S = [t_0, \infty) \times \mathbb{R}^n \rightarrow$ free-time, free-endpoint

(advanced) $S = \{(t, g(t)) : t \in [t_0, \infty)\} \rightarrow$ moving target set for some continuous $g: \mathbb{R} \rightarrow \mathbb{R}^n$

↳ but for these cases, we can use the trick of introducing $x_{n+1} = t$ and reduce to $\tilde{S} = [t_0, \infty) \times M$ with $M \triangleq \left\{ x \in \mathbb{R}^{n+1} : x = \begin{bmatrix} g(x_{n+1}) \\ x_{n+1} \end{bmatrix} \right\}$.

In Summary:

$$\min_{u(\cdot)} J(u) = \int_{t_0}^{t_f} L(t, x, u) dt + K(t_f, x_{t_f})$$

$$\text{s.t.} \quad \dot{x} = f(t, x, u), \quad x_{t_0} = x_0$$

$$u(t) \in U, \quad \forall t.$$

$$(t_f, x_{t_f}) \in S \subset [t_0, \infty) \times \mathbb{R}^n$$