Optimal Control Theory

Optimal Control Problem Statement

- A generic optimal control problem:

$$\begin{aligned} \min \quad & \int_{t_0}^{t_f} L(x,u,t)\mathrm{d}t \\ \text{s.t.} \quad & \dot{x} = f(x,u,t) \quad \leftarrow \text{dynamical constraint} \\ & \psi(x_0, x_f, t_0, t_f) = 0 \quad \leftarrow \text{boundary constraint} \\ & x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m, \quad \psi \in \mathbb{R}^q \end{aligned}$$

- Three types of cost functions:
  - Lagrange — Mayer — Bolza

$$J = \int_{t_0}^{t_f} L(x,u,t)\mathrm{d}t \qquad J = K(x_f, \underset{\uparrow \, t_f}{T}) \qquad J = K(x_f, T) + \int_{t_0}^{t_f} L(x,u,t)\mathrm{d}t$$

*any cost function can be converted to Lagrange form[1]

1. L. Cesari, Optimization—Theory and Applications (Section 1.9), Springer-Verlag, 1983

Conversion to unconstrained optimization problem

- Solving constrained optimization is hard (analytically & numerically)—— *Convert the generic optimal control problem into an unconstrained optimization problem:

$$\begin{aligned} \min \quad & \int_{t_0}^{t_f} L(x,u,t)\mathrm{d}t \\ \text{s.t.} \quad & \dot{x} = f(x,u,t) \\ & \psi(x_0, x_f, t_0, t_f) = 0 \end{aligned}$$

$\dot{x} - f = 0$

$$\min \quad J = \nu^\top \psi + \int_{t_0}^{t_f} L + \lambda^\top (f - \dot{x})\ \mathrm{d}t$$

costate

$\lambda \in \mathbb{R}^n$, $\nu \in \mathbb{R}^q$: Lagrange multiplier

  - The same idea as the approach to constrained static optimization problems

- The optimal solution must minimize the augmented problem
  - Seek necessary conditions of optimality

Optimality Necessary Conditions

- Optimization problem:

$$\min \quad J = \nu^\top \psi + \int_{t_0}^{t_f} H - \lambda^\top \dot{x} \ \mathrm{d}t$$

  – where a scalar quantity is introduced for convenience:

$$H(x, u, \lambda, t) = L(x, u, t) + \lambda^\top f(x, u, t)$$

"control Hamiltonian" or simply "Hamiltonian"

- Necessary conditions via calculus of variations:
  – For the cost J to be minimum, it is necessary that the variation of J is zero, i.e.,

$$\mathrm{d}J = 0 \text{ (stationary condition)}$$

  – Expanding dJ, we have

$$\mathrm{d}J = \nu^\top \psi_{x_0} \mathrm{d}x_0 + \nu^\top \psi_{x_f} \mathrm{d}x_f + \nu^\top \psi_{t_0} \mathrm{d}t_0 + \nu^\top \psi_{t_f} \mathrm{d}t_f + \psi^\top \mathrm{d}\nu + \int_{t_0}^{t_f} H_x \mathrm{d}x + H_u \mathrm{d}u + H_\lambda \mathrm{d}\lambda - \lambda^\top \mathrm{d}\dot{x} - \dot{x}^\top \mathrm{d}\lambda \ \mathrm{d}t - \{H(t_0) - [\lambda(t_0)]^\top \dot{x}(t_0)\}\mathrm{d}t_0 + \{H(t_f) - [\lambda(t_f)]^\top \dot{x}(t_f)\}\mathrm{d}t_f$$

$$= \nu^\top \psi_{x_0} \mathrm{d}x_0 + \nu^\top \psi_{x_f} \mathrm{d}x_f + \{\nu^\top \psi_{t_0} - H(t_0) + [\lambda(t_0)]^\top \dot{x}(t_0)\}\mathrm{d}t_0 + \{\nu^\top \psi_{t_f} + H(t_f) - [\lambda(t_f)]^\top \dot{x}(t_f)\}\mathrm{d}t_f + \psi^\top \mathrm{d}\nu + \int_{t_0}^{t_f} H_x \mathrm{d}x + H_u \mathrm{d}u + (H_\lambda - \dot{x}^\top)\mathrm{d}\lambda \ \mathrm{d}t - \int_{t_0}^{t_f} \lambda^\top \mathrm{d}\dot{x} \ \mathrm{d}t$$

$$f_x = \frac{\partial f}{\partial x}, \quad f_u = \frac{\partial f}{\partial u}, \ldots$$

Optimality Necessary Conditions (cont'd)

- Integrating by parts,

$dx_0 = d[x(t_0)] + \dot{x}(t_0)dt_0$

$dx_f = d[x(t_f)] + \dot{x}(t_f)dt_f$

$$\int_{t_0}^{t_f} \lambda^\top \mathrm{d}\dot{x}\ \mathrm{d}t = [\lambda(t_f)]^\top \mathrm{d}[x(t_f)] - [\lambda(t_0)]^\top \mathrm{d}[x(t_0)] - \int_{t_0}^{t_f} \dot{\lambda}^\top \mathrm{d}x\ \mathrm{d}t = [\lambda(t_f)]^\top [\underline{\mathrm{d}x_f - \dot{x}(t_f)\mathrm{d}t_f}] - [\lambda(t_0)]^\top [\underline{\mathrm{d}x_0 - \dot{x}(t_0)\mathrm{d}t_0}] - \int_0^T \dot{\lambda}^\top \mathrm{d}x\ \mathrm{d}t$$

$$\mathrm{d}J = [\nu^\top \psi_{x_0} + \lambda(t_0)^\top]\mathrm{d}x_0 + [\nu^\top \psi_{x_f} - \lambda(t_f)^\top]\mathrm{d}x_f - [\nu^\top \psi_{t_0} - H(t_0)]\mathrm{d}t_0 + [\nu^\top \psi_{t_f} + H(t_f)]\mathrm{d}t_f + \psi^\top \mathrm{d}\nu + \int_{t_0}^{t_f} (H_x + \dot{\lambda}^\top)\mathrm{d}x + H_u \mathrm{d}u + (H_\lambda - \dot{x}^\top)\mathrm{d}\lambda\ \mathrm{d}t$$

- dJ needs to be zero for any possible variations, implying

$$\dot{\lambda}^\top = -H_x, \quad H_u = 0, \quad \dot{x}^\top = H_\lambda,$$

$$\psi = 0 \text{ (if } \mathrm{d}\nu \neq 0), \quad \nu^\top \psi_{x_0} + \lambda_0^\top = 0 \text{ (if } \mathrm{d}x_0 \neq 0), \quad \nu^\top \psi_{x_f} - \lambda(t_f)^\top = 0 \text{ (if } \mathrm{d}x_f \neq 0), \quad \nu^\top \psi_{t_0} - H(t_0) = 0 \text{ (if } \mathrm{d}t_0 \neq 0), \quad \nu^\top \psi_{t_f} + H(t_f) = 0 \text{ (if } \mathrm{d}t_f \neq 0)$$

- Necessary conditions:
  - State dynamics
    $$\dot{x}^\top = H_\lambda = f^\top$$
  - Costate dynamics
    $$\dot{\lambda}^\top = -H_x = -L_x - \lambda^\top f_x$$
    f_{EOM}
  - Optimal Control
    $$H_u = L_u + \lambda^\top f_u = 0$$
  - The others are called *transversality conditions* and determine the initial/final conditions of multipliers

Example

$$J = \int_{t_0}^{t_f} L \, dt \qquad \text{s.t.} \quad \dot{x} = f(x, u, t)$$

$$\psi(\cdots) = 0$$

$$L = \frac{1}{2} \vec{x} \cdot \vec{x} \qquad B = \begin{bmatrix} 0_{3\times 3} \\ I_3 \end{bmatrix}$$

$$\begin{bmatrix} 0_{3\times 1} \\ \vec{u} \end{bmatrix}$$

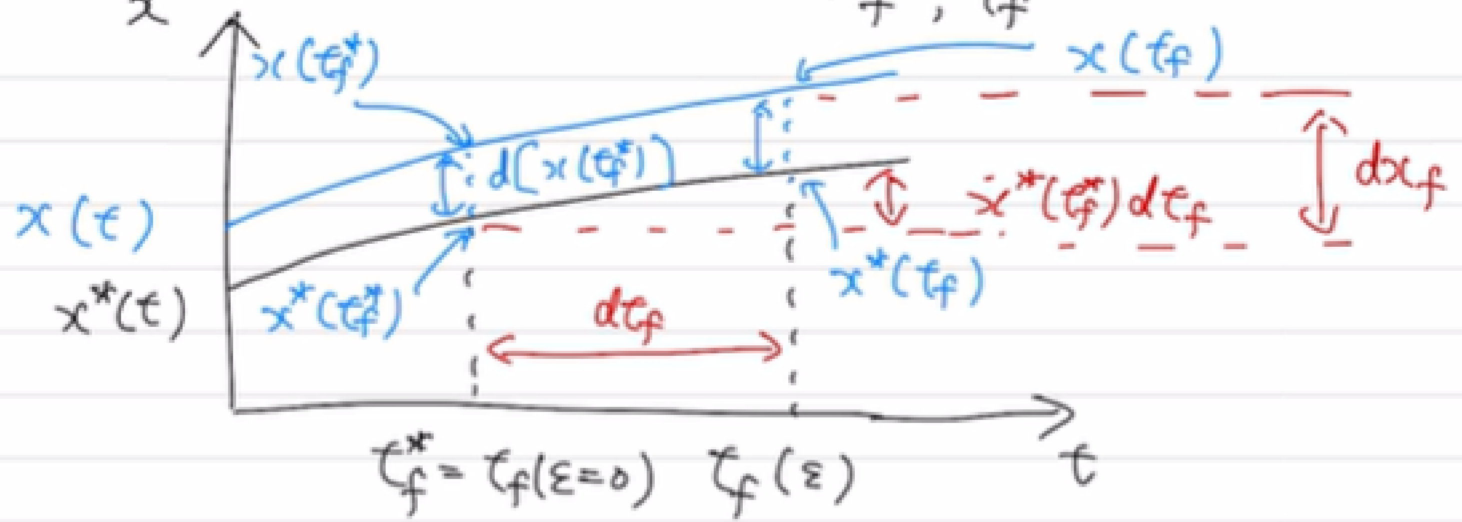
$$f = f_0(x) + B\vec{u} \qquad u \in \mathbb{R}^{3\times 1}$$

- $H = L + \lambda^T f = \frac{1}{2} \vec{x} \cdot \vec{x} + \vec{\lambda}^T \left(\vec{f}_0 + B\vec{u} \right)$

- $0 = \frac{\partial H}{\partial u} = \vec{\lambda}^T B$

Variation of dx_f $\quad dx_f$

x_f , t_f



$$d[x(t_f^*)] = x(t_f^*) - x^*(t_f^*)$$

$$dx_f = d[x(t_f^*)] + \dot{x}^*(t_f^*)\,dt_f \;(+\ \text{H.O.T.})$$

Pontryagin's Minimum Principle

- Pontryagin's Minimum Principle
  - First formulated by Pontryagin in 1950's
  - Also known as: Maximum Principle or Minimum Principle

- Necessary conditions according to Pontryagin:
  - State dynamics

    $$\dot{x}^\top = H_\lambda = f^\top$$

    Costate dynamics

    $$\dot{\lambda}^\top = -H_x = -L_x - \lambda^\top f_x$$

    Optimal Control

    $$H(x^*, u^*, \lambda, t) \leq H(x^*, u, \lambda, t)$$

    or $u^* = \arg\min_{u \in \mathcal{U}} H(x^*, u, \lambda, t)$
  - Same transversality conditions
  - *This makes many practical optimal control problems solvable

Some additional comments

- Control Hamiltonian H on an optimal trajectory is constant if H is not an explicit function of time

$\dot{x} = f$, $H_\lambda = f^T$

$H(x,u,\lambda,t) = L(x,u,t) + \lambda^T f(x,u,t)$

$$\dot{H} = H_t + H_x\dot{x} + H_u\dot{u} + H_\lambda\dot{\lambda} = H_t + H_u\dot{u} + (\underbrace{H_x + \dot{\lambda}^\top}_{H_x = -\dot{\lambda}^T})f = H_t + \underbrace{H_u}_{0}\dot{u} = H_t$$

- Notes on transversality condition
  - Generic form:

$$\psi = 0 \text{ (if } \mathrm{d}\nu \neq 0), \quad \nu^\top\psi_{x_0} + \lambda_0^\top = 0 \text{ (if } \mathrm{d}x_0 \neq 0), \quad \nu^\top\psi_{x_f} - \lambda(t_f)^\top = 0 \text{ (if } \mathrm{d}x_f \neq 0), \quad \nu^\top\psi_{t_0} - H(t_0) = 0 \text{ (if } \mathrm{d}t_0 \neq 0), \quad \nu^\top\psi_{t_f} + H(t_f) = 0 \text{ (if } \mathrm{d}t_f \neq 0)$$

  - Some typical cases (not an exhaustive list):
    - Fixed boundary condition
    - Free boundary condition
    - Boundary condition being function of the state

Solution Method for Optimal Control Problems

Applying Optimal Control Theory

- Optimality necessary conditions for generic problems:
  - State dynamics $\dot{x} = f(\cdot)$
  - Costate dynamics $\dot{\lambda} = g(\cdot) \quad g(x, \lambda, u, t) \triangleq -H_x^\top$
  - Optimal Control $u^* = \arg\min_{u \in \mathcal{U}} H(x^*, u, \lambda, t)$
  - Transversality condition

$$\psi = 0 \text{ (if } d\nu \neq 0), \quad \nu^\top \psi_{x_0} + \lambda_0^\top = 0 \text{ (if } dx_0 \neq 0), \quad \nu^\top \psi_{x_f} - \lambda(t_f)^\top = 0 \text{ (if } dx_f \neq 0), \quad \nu^\top \psi_{t_0} - H(t_0) = 0 \text{ (if } dt_0 \neq 0), \quad \nu^\top \psi_{t_f} + H(t_f) = 0 \text{ (if } dt_f \neq 0)$$

- To solve the optimal control problem:
  - Can the optimal control analytically be obtained? assume so
  - Optimal trajectory of control, state, and costate:

$u^*(t) = h(x^*(t), \lambda^*(t), t)$, $x^*(t) = x_0 + \int_{t_0}^{t} f(x^*(\tau), u^*(\tau), \tau)\, d\tau$

$\lambda^*(t) = \lambda^*(0) + \int_{t_0}^{t} g(x^*(\tau), \lambda^*(\tau), \tau)\, d\tau$

  - What are the other conditions? What are unknowns?

λ_0^* and sometimes x_0^*, t_f^* unknown   transversality conds

Indirect Method for TPBVP

- Two-point boundary value problem (TPBVP)
  - Need to satisfy the final piece of necessary conditions
  $$\Psi(Z) = \begin{bmatrix} \psi \\ \vdots \end{bmatrix} \qquad Z = \begin{bmatrix} \lambda_0 \\ \nu \\ \vdots \end{bmatrix}$$
  - i.e., find a Z that satisfies $\Psi(Z) = 0$ following the optimal state-costate dynamics:
  $$u^*(t) = h(x^*(t), \lambda^*(t), t) \qquad x^*(t) = x_0 + \int_0^t f(x^*(\tau), u^*(\tau), \tau)\mathrm{d}\tau, \qquad \lambda^*(t) = \lambda_0 + \int_0^t g(x^*(\tau), \lambda^*(\tau), u^*(\tau), \tau)\mathrm{d}\tau,$$
  - An infinite-dimensional optimization problem transformed into a TPBVP
    - called *indirect method*
    - another category called *direct method* parameterizes the infinite-dimensional problem and directly solves it via parameter optimization
  - Analytical solution available?

Generally, no $\longrightarrow$ NL root finding algorithm

↳ shooting method. Ψ is shooting function

Indirect method

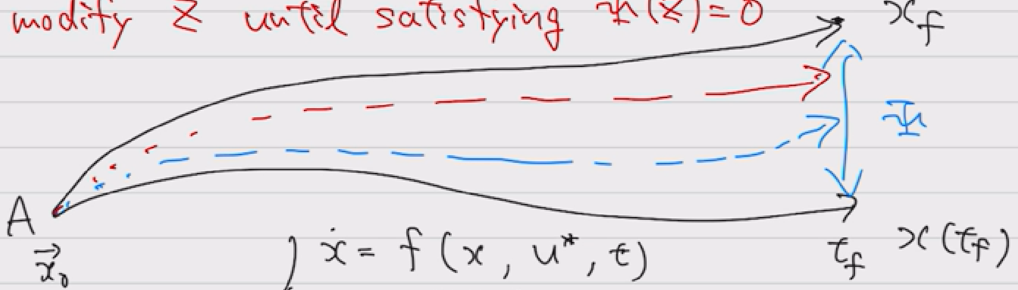
single - shooting method.

- unknowns : $\vec{\lambda}_0$, t_f $\qquad \vec{z} = \begin{bmatrix} \vec{\lambda}_0 \\ t_f \end{bmatrix}$

- find optimal $\vec{\lambda}_0^*$, t_f^* s.t. satisfy $\Psi(\vec{z}) = 0$

- guess : $\vec{\lambda}_0$, t_f

- modify z until satisfying $\Psi(z) = 0$



$$\begin{cases} \dot{x} = f(x, u^*, t) \\ \dot{\lambda} = -\left[\frac{\partial H}{\partial x}\right]^T \end{cases}$$

guess $\begin{cases} \vec{\lambda}_0 \\ t_f \end{cases}$

- nonlinear root-finding algorithm.
  - Ex: fsolve in Matlab.

$\rightarrow$ solve $\vec{z}$ for $\Psi(\vec{z}) = 0$

- $\frac{\partial \Psi}{\partial z}$: fsolve calculates this sensitivity numerically (finite differentiation)
- you can calculate $\frac{\partial \Psi}{\partial z}$ and provide to fsolve. $\rightarrow$ improved convergence

Applications:
Optimal Control with Linear Dynamics

General linear system: Time-varying LQR

- Finite-horizon LQR with time-varying system
  - Problem statement:
    - General linear time-varying system: $\dot{x} = A(t)x + B(t)u \qquad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m$
    - Objective: $J = \frac{1}{2}\int_0^{t_f} x^\top Q x + u^\top R u dt$

Fixed t_f, no terminal constraints
Q: PSD matrix, R: PD matrix (both symmetric)
(Q and R can be time-varying)

  - Solution (derived in note):
    - Optimal control: $u^* = -R^{-1}B^\top K x$
    - K matrix: $\dot{K} = -KA + KBR^{-1}B^\top K - Q - A^\top K, \quad K(t_f) = 0 \qquad K \in \mathbb{R}^{n \times n}$

Riccati differential equations

① Solve for K(t) by int $\dot{K}$ backward in time
② use K(t) to calc $u^*(t)$, hence $x^*(t)$ → propagate from $t_0 = 0$ to t_f

LQR

$$\begin{cases} \dot{x} = \underbrace{Ax + Bu}_{f(x,u)} \cdots \text{LTV} \\ J = \frac{1}{2}\int_0^{t_f} x^T Q x + u^T R u \, dt \end{cases}$$

$\hookrightarrow \quad L = \frac{1}{2}\left(x^T Q x + u^T R u\right)$

Hamiltonian

$$H = L + \lambda^T f = \frac{1}{2}\left(x^T Q x + u^T R u\right) + \lambda^T (Ax + Bu)$$

Optimal control

$\min_u H \quad \leftarrow$ Pontryagin.

$$\left.\begin{cases} \frac{\partial H}{\partial u} = 0 \\ \frac{\partial^2 H}{\partial u^2} > 0 \quad (\text{p.d.}) \end{cases}\right\} \rightarrow \text{Legendre - Clebsch condition.}$$

$$\begin{cases} \cdot\ \frac{\partial H}{\partial u} = u^T R + \lambda^T B = 0 \quad \Longleftrightarrow \quad u^* = -R^{-1} B^T \lambda \\ \cdot\ \frac{\partial^2 H}{\partial u^2} = R > 0 \quad (\text{assumption}) \end{cases}$$

R^{-1} exist because $R > 0$

Costate trajectory

$$\dot{\lambda}^T = -\frac{\partial H}{\partial x} = -x^T Q - \lambda^T A$$

$$\Longleftrightarrow \quad \dot{\lambda} = -Qx - A^T \lambda$$

Q : symmetric

Transversality condition

$dx_0 = 0, \quad dt_0 = 0, \quad dx_f : \text{free}, \quad dt_f = 0 \Rightarrow \lambda(t_f) = 0$

→ what is λ_0 that satisfies $\lambda(t_f) = 0$ under $\dot{\lambda}$ and u^*

(cont'd)

- costate dynamics: $\dot{\lambda} = -Qx - A^T\lambda$

- assume $\lambda = Kx$, then $u^* = -R^{-1}B^TKx$

then, costate dynamics become:

$$\dot{K}x + K\dot{x} = -Qx - A^TKx \qquad \begin{cases} \dot{x} = Ax + Bu \\ u^* = -R^{-1}B^TKx \end{cases}$$

$$\dot{K}x + K(Ax + B(-R^{-1}B^TKx)) = -Qx - A^TKx$$

$$\Leftrightarrow (\dot{K} + KA - KBR^{-1}B^TK + Q + A^TK)x = 0$$

This implies that

$$\dot{K} + KA - KBR^{-1}B^TK + Q + A^TK = 0$$

$$\Leftrightarrow \dot{K} = -KA + KBR^{-1}B^TK - Q - A^TK$$

- converted $\dot{\lambda} \rightarrow \dot{K}$, which eliminates the dependency on x.

- To integrate $\dot{K}$, we need initial condition for K.

$\rightarrow$ use transversality condition: $\lambda(t_f) = 0$

$$\lambda(t_f) = K(t_f)\underbrace{x(t_f)}_{\text{free}} = 0 \quad \text{then} \quad K(t_f) = 0$$

- therefore, we can calculate $K(t)$ by integrating

$\dot{K}$ backward in time from t_f to $t_0 = 0$ with $K(t_f) = 0$

Relative Orbit Transfer

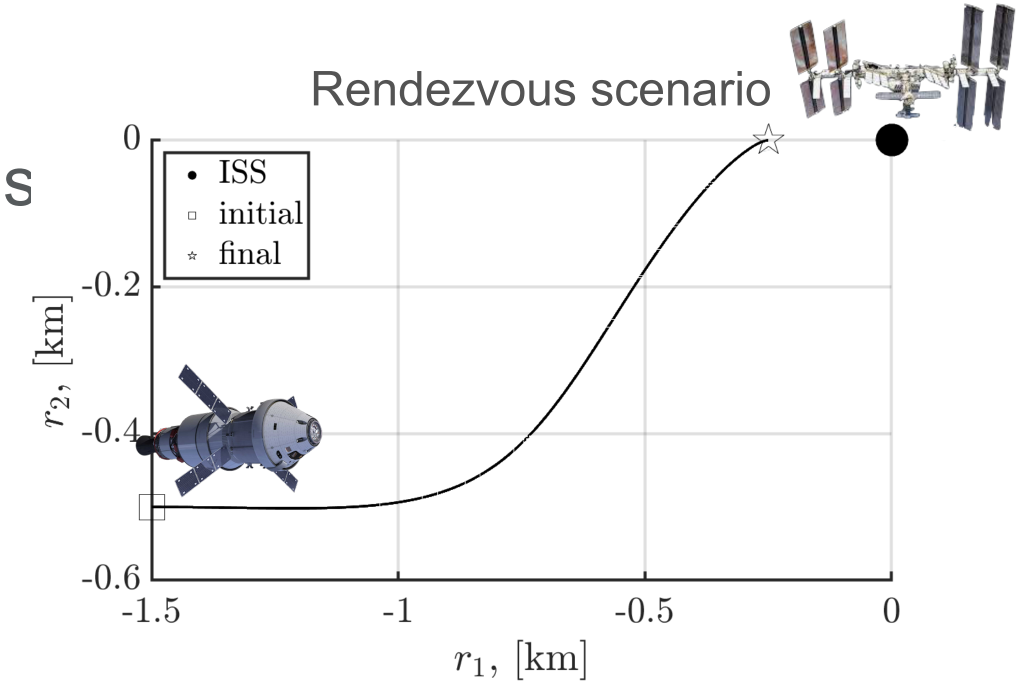
- Rendezvous with ISS in 2-D CWH equations
  - Problem statement:
    - Variables: $x = [r_1 \quad r_2 \quad v_1 \quad v_2]^\top, \quad u = [u_1 \quad u_2]^\top$
    - Objective: $J = \int_0^{t_f} \underbrace{u \cdot u}_{\|u\|_2^2} dt$
    - Equations of motion: CWH equation

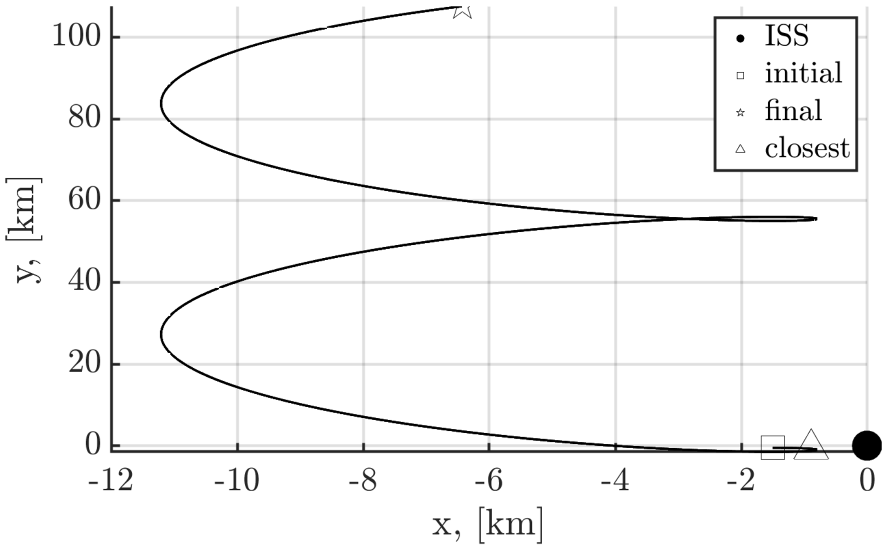
$$\dot{x} = Ax + Bu, \quad A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 3n^2 & 0 & 0 & 2n \\ 0 & 0 & -2n & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0_{2\times 2} \\ I_2 \end{bmatrix}$$

    - Terminal constraints:

$$t_f = 20 \text{ min} \qquad x_0 = \begin{bmatrix} -1.5 \text{ km} \\ -0.5 \text{ km} \\ 3.0 \text{ m/s} \\ 0.0 \text{ m/s} \end{bmatrix} \qquad x_f = \begin{bmatrix} -0.25 \text{ km} \\ 0.0 \text{ km} \\ 0.2 \text{ m/s} \\ 0.0 \text{ m/s} \end{bmatrix}$$

Rendezvous scenario

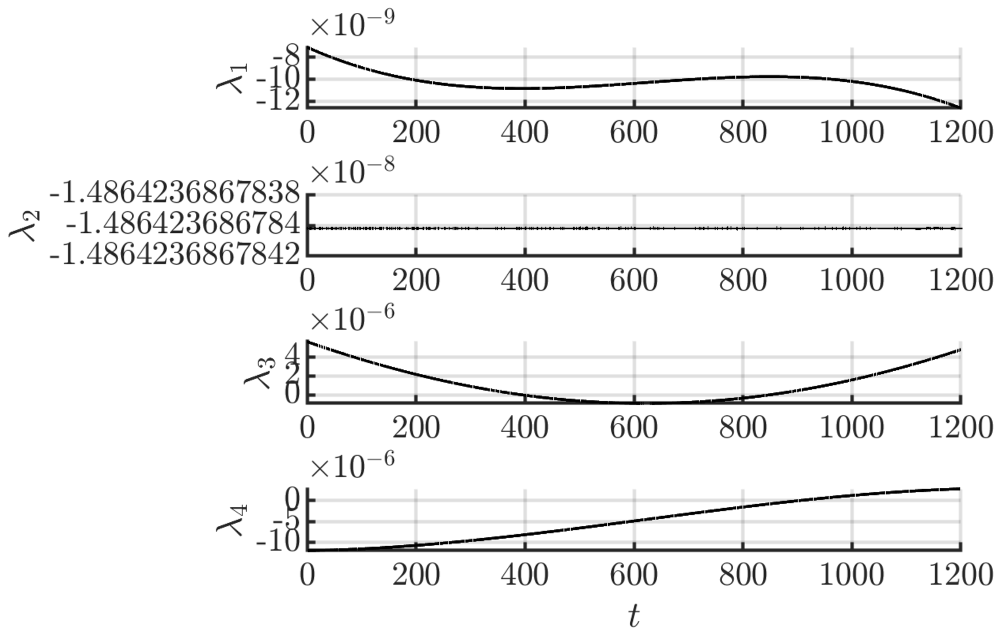


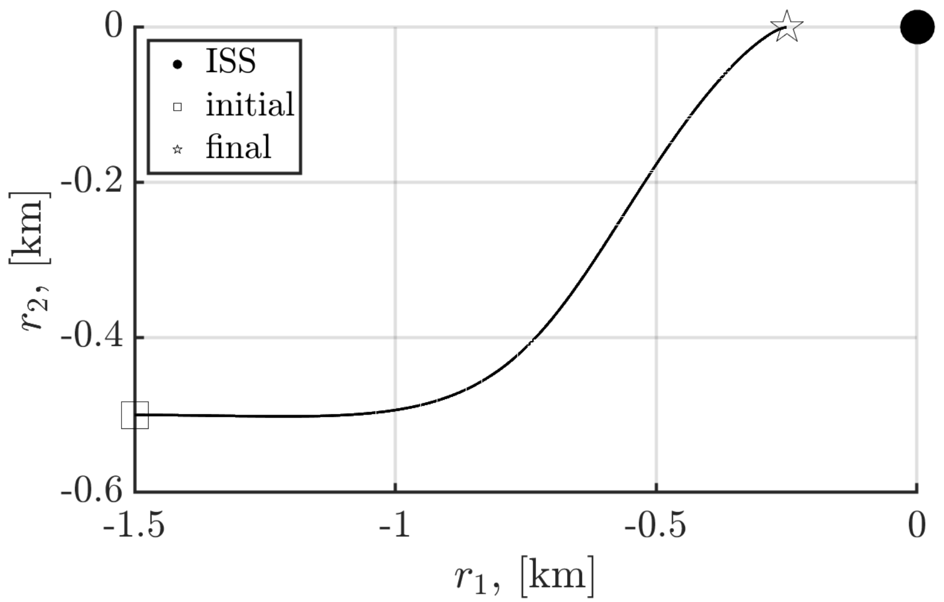


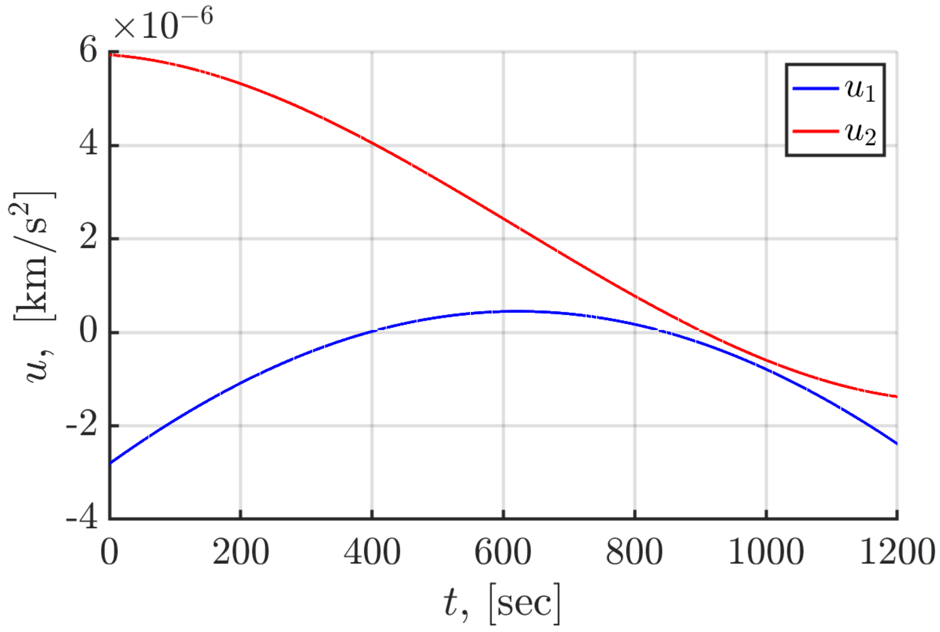
Natural trajectory (no control)

Relative Orbit Transfer (cont'd)

- We can analytically solve the problem:
  - optimal state/costate trajectory and control
  - …Derive in note
- Plots of the optimal trajectories & control:







· Hamiltonian

$$H = L + \lambda^T f = \|\vec{u}\|_2^2 + \lambda^T (Ax + Bu)$$

· State dynamics

$$\dot{x}^T = \frac{\partial H}{\partial \lambda} = f^T = (Ax + Bu)^T$$

· Costate dynamics

$$\dot{\lambda}^T = -\frac{\partial H}{\partial x} = -(\lambda^T A) \iff \dot{\lambda} = -A^T \lambda$$

· Optimal control

$$u^* = \underset{u}{\mathrm{argmin}}\ H = \underset{u}{\mathrm{argmin}}\ \|\vec{u}\|_2^2 + \lambda^T (Ax + Bu)$$

$$\begin{cases} \cdot\ \dfrac{\partial H}{\partial u} = 0 \iff 2u^T + \lambda^T B = 0 \iff u^* = -\dfrac{1}{2} B^T \lambda \\ \cdot\ \dfrac{\partial^2 H}{\partial u^2} > 0 \iff 2 I_2 > 0 \end{cases}$$

Transversality condition

$$dx_0 = dx_f = dt_0 = dt_f = 0 \Rightarrow \Psi = \begin{bmatrix} x_0 - x_{ini} \\ x_f - x_{tar} \end{bmatrix} = 0$$

$\Rightarrow$ we will start propagation from $x_0 = x_{ini}$

$\Rightarrow$ we can eliminate the first constraint : $\Psi_1 = \underbrace{x_f - x_{tar} = 0}_{4\text{-D}}$

unknown : $\lambda_0 \in \mathbb{R}^4$

cont'd

analytical sol'n for $\Phi(z) = 0$ $\quad z = \lambda_0$

Augmented state $\quad X = \begin{bmatrix} x \\ \lambda \end{bmatrix} \in \mathbb{R}^8, \quad X_0 = \begin{bmatrix} x_0 \\ \lambda_0 \end{bmatrix}$

- $\dot{X} = \begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} Ax + Bu \\ -A^T\lambda \end{bmatrix} = F(X, u)$

- $F(X, u^*) = \begin{bmatrix} Ax - \frac{1}{2}BB^T\lambda \\ -A^T\lambda \end{bmatrix} = \underbrace{\begin{bmatrix} A & -\frac{1}{2}BB^T \\ 0 & -A^T \end{bmatrix}}_{A_{aug}} \underbrace{\begin{bmatrix} x \\ \lambda \end{bmatrix}}_{X}$

$$= \underbrace{A_{aug}}_{\text{time-invariant}} X$$

Solve for $X(t)$

$$\begin{cases} \dot{x} = f \\ \dot{\Phi} = \frac{\partial f}{\partial x}\Phi \end{cases}$$

$$\begin{cases} X(t) = \Phi(t, t_0) X_0 \\ \dot{\Phi}(t, t_0) = \frac{\partial F}{\partial X}\Phi(t, t_0) = \underbrace{A_{aug}}_{\text{constant}}\Phi(t, t_0) \end{cases}$$

$$\Phi(t, t_0) = \exp[A_{aug}\cdot(t - t_0)]\underbrace{\Phi(t_0, t_0)}_{I}$$

$$= \exp[A_{aug}\cdot(t - t_0)] \in \mathbb{R}^{8\times 8}$$

$$\triangleq \begin{bmatrix} \phi_{11}(t) & \phi_{12}(t) \\ \phi_{21}(t) & \underbrace{\phi_{22}(t)}_{\in \mathbb{R}^{4\times 4}} \end{bmatrix}$$

- $\begin{bmatrix} x(t) \\ \lambda(t) \end{bmatrix} = \Phi(t, t_0) X_0 = \begin{bmatrix} \phi_{11}(t) & \phi_{12}(t) \\ \phi_{21}(t) & \phi_{22}(t) \end{bmatrix} \begin{bmatrix} x_0 \\ \lambda_0 \end{bmatrix} = \begin{bmatrix} \phi_{11}x_0 + \phi_{12}\lambda_0 \\ \phi_{21}x_0 + \phi_{22}\lambda_0 \end{bmatrix}$

(cont'd)

Solve TPBVP

$$X(t_f) = \Phi(t_f, t_0) X_0$$
$$X_0 = \Phi^{-1} X(t_f)$$

$$\begin{bmatrix} x(t_f) \\ \lambda(t_f) \end{bmatrix} = \begin{bmatrix} \phi_{11}(t_f) x_0 + \phi_{12}(t_f) \lambda_0 \\ \phi_{21}(t_f) x_0 + \phi_{22}(t_f) \lambda_0 \end{bmatrix} = \begin{bmatrix} x_{tar} \\ \text{arbitrary} \end{bmatrix}$$

$\Rightarrow$ $x_{tar} = \phi_{11}(t_f) x_0 + \phi_{12}(t_f) \lambda_0$

$\Leftrightarrow$ $\lambda_0 = \left[\phi_{12}(t_f)\right]^{-1} \left[x_{tar} - \phi_{11}(t_f) x_0\right]$

$\Rightarrow$ we can also propagate optimal $x(t)$, $\lambda(t)$, $u(t)$ analytically.

$$\begin{cases} x^*(t) = \phi_{11}(t) x_0 + \phi_{12}(t) \lambda_0 \\ \lambda^*(t) = \underbrace{\phi_{21}(t) x_0}_{=0} + \phi_{22}(t) \lambda_0 \\ u^*(t) = -\frac{1}{2} B^T \lambda^*(t) \end{cases}$$

Rocket Thrust Steering

- (Very) simplified rocket steering problem
  - Problem statement—*Launch into an orbit from flat Earth*
    - Equations of motion: $\begin{cases} \dot{r}_1 = v_1 \\ \dot{r}_2 = v_2 \end{cases}, \quad \begin{cases} \dot{v}_1 = a\cos\theta \\ \dot{v}_2 = a\sin\theta - g \end{cases}$
      - control: steering angle θ
      - constant acceleration
      - uniform gravity assumption
    - Objective—minimum-time orbit injection: $J = t_f - t_0$
    - Terminal constraints: $\begin{cases} r_1(t_f) : \text{free}, \\ r_2(t_f) = h_f, \end{cases}, \quad \begin{cases} v_1(t_f) = v_f \\ v_2(t_f) = 0 \end{cases}$ target circular orbit
  - Solution (derived in note)
    - Optimal control: $\tan\theta^* = \dfrac{-\lambda_4}{-\lambda_3}, \quad \text{or} \quad \theta^* = \arctan 2(-\lambda_4, -\lambda_3)$ *Bilinear tangent steering law*
    - Costate trajectory: $\lambda_1(t) = c_1(=0), \quad \lambda_2(t) = c_2, \quad \lambda_3(t) = -c_1 t + c_3, \quad \lambda_4(t) = -c_2 t + c_4$

$$H(t) = 1 + \lambda^\top f = 0$$

Rocket steering

min-time orbit injection.

Hamiltonian

$$H = L + \lambda^T f = 1 + \lambda^T \begin{bmatrix} v_1 \\ v_2 \\ a\cos\theta \\ a\sin\theta - g \end{bmatrix}$$

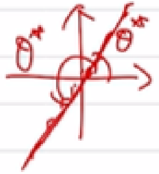
$$= 1 + \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 a\cos\theta + \lambda_4 (a\sin\theta - g)$$

Optimal control

Legendre - Clebsch condition.

$$\begin{cases} \dfrac{\partial H}{\partial \theta} = 0 & \Rightarrow \dfrac{\partial H}{\partial \theta} = -\lambda_3 a\sin\theta + \lambda_4 a\cos\theta = 0 \quad —① \\[2ex] \dfrac{\partial^2 H}{\partial \theta^2} > 0 & \Rightarrow \dfrac{\partial^2 H}{\partial \theta^2} = -\lambda_3 a\cos\theta - \lambda_4 a\sin\theta > 0 \quad —② \end{cases}$$

① $\Rightarrow$ $\lambda_4 = \tan\theta\, \lambda_3$ $\Rightarrow$ $\tan\theta^* = \dfrac{\lambda_4}{\lambda_3}$ ← quad. ambiguity



② $\Rightarrow$ $-\lambda_3 \cos\theta^* - \lambda_4 \sin\theta^* > 0$

① $\Rightarrow$ $\sin\theta^* = \dfrac{\pm\lambda_4}{\sqrt{\lambda_3^2 + \lambda_4^2}}$, $\cos\theta^* = \dfrac{\pm\lambda_3}{\sqrt{\lambda_3^2+\lambda_4^2}}$

② $\Rightarrow$ $\mp \dfrac{(\lambda_3^2 + \lambda_4^2)}{\sqrt{\lambda_3^2+\lambda_4^2}} > 0$

$\Rightarrow$ $\sin\theta^* = \dfrac{-\lambda_4}{\sqrt{\lambda_3^2+\lambda_4^2}}$, $\cos\theta^* = \dfrac{-\lambda_3}{\sqrt{\lambda_3^2+\lambda_4^2}}$

$\therefore$ $\theta^* = \mathrm{arctan2}\,(-\lambda_4, -\lambda_3)$

$\left(\theta^* = \arctan \dfrac{-\lambda_4}{-\lambda_3}\right)$

Costate dynamics

$$\dot{\lambda}^T = -\frac{\partial H}{\partial x}$$

$$\Rightarrow \begin{cases} \dot{\lambda}_1 = -\frac{\partial H}{\partial r_1} = 0 \quad , \quad \dot{\lambda}_2 = -\frac{\partial H}{\partial r_2} = 0 \\ \dot{\lambda}_3 = -\frac{\partial H}{\partial v_1} = -\lambda_1 \quad , \quad \dot{\lambda}_4 = -\frac{\partial H}{\partial v_2} = -\lambda_2 \end{cases}$$

costate trajectory

$$\begin{cases} \cdot\ \lambda_1 = C_1 \quad , \quad \lambda_2 = C_2 \\ \cdot\ \lambda_3 = -C_1 t + C_3 \quad , \quad \lambda_4 = -C_2 t + C_4 \end{cases} \quad \forall t \in (t_0, t_f)$$

Transversality conditions

$$\begin{cases} r_1(t_f) : \text{free} \\ r_2(t_f) : h_f \end{cases} \quad \begin{cases} v_1(t_f) = v_f \\ v_2(t_f) = 0 \end{cases} \quad \begin{cases} t_0 : \text{given} \\ t_f : \text{free} \end{cases}$$

t_f: free

$$\hookrightarrow \underbrace{\nu^T \frac{\partial \psi}{\partial t_f}}_{=0} + H(t_f) = 0 \iff H(t_f) = 0$$

$r_1(t_f)$: free

$$\underbrace{\nu^T \frac{\partial \psi}{\partial r_1(t_f)}}_{=0} - \lambda_1(t_f) = 0 \iff \lambda_1(t_f) = 0$$

$$\Rightarrow \text{shooting fn}: \ \Psi = \begin{bmatrix} r_2(t_f) - h_f \\ v_1(t_f) - v_f \\ v_2(t_f) \\ \lambda_1(t_f) \\ H(t_f) \end{bmatrix} = 0 \quad \text{for unknowns: } \underbrace{\lambda(t_0)}, \ t_f$$

(or, $C_1 \cdots C_4$)

Applications:
Optimal Low-thrust Orbit Transfer

PURDUE UNIVERSITY School of Aeronautics and Astronautics

Low-thrust Optimal Orbit Transfer Problem

- Problem statement:

$$\min_{x,u,t_0,t_f} \int_{t_0}^{t_f} L(u)\mathrm{d}t$$

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

$$x_0 : \text{given}, \quad \psi = x_f - x_{\text{tar}}(t_f) = 0$$

$$x = \begin{bmatrix} r \\ v \end{bmatrix}, \quad f = f_0(x,t) + Bu, \quad \|u\|_2 \leq u_{\max}$$

$$f_0(x,t) = \begin{bmatrix} v \\ -\frac{\mu}{\|r\|_2^3} r + a_{\text{dist}}(x,t) \end{bmatrix}, \quad B = \begin{bmatrix} 0_{2\times 2} \\ I_2 \end{bmatrix}$$

$$r \in \mathbb{R}^2, \quad v \in \mathbb{R}^2, \quad u \in \mathbb{R}^2$$

- Typical cost functions:
  - Minimum time-of-flight (ToF): $L(u) = 1 \quad \Leftrightarrow \quad \int_{t_0}^{t_f} 1\mathrm{d}t = t_f - t_0$
  - Minimum energy: $L(u) = \|u\|_2^2 = u \cdot u$
  - Minimum fuel consumption: $L(u) = \|u\|_2$

Minimum-time Low-thrust Orbit Transfer

- Control Hamiltonian: $H = L + \lambda^\top f = 1 + \lambda^\top f$
- Optimality necessary conditions:
  - State dynamics
    $$\dot{x} = f$$
  - Costate dynamics
    $$\dot{\lambda}^\top = -H_x = -\lambda^\top f_x$$
  - Optimal Control
    $$u^* = \arg\min_{u \in \mathcal{U}} H = \arg\min_{\|u\|_2 \le u_{\max}} \left[1 + \lambda^\top f(x^*, u)\right]$$
  - Analytical optimal control law:
    $$u^* = \arg\min_{\|u\|_2 \le u_{\max}} \lambda^\top f(x^*, u) = \arg\min_{\|u\|_2 \le u_{\max}} \lambda^\top f_0(x, t) + \lambda^\top B u = \arg\min_{\|u\|_2 \le u_{\max}} \lambda^\top B u$$
    $$u^* = \frac{p}{\|p\|_2} u_{\max}, \quad p = -B^\top \lambda$$
  - Transversality conditions:

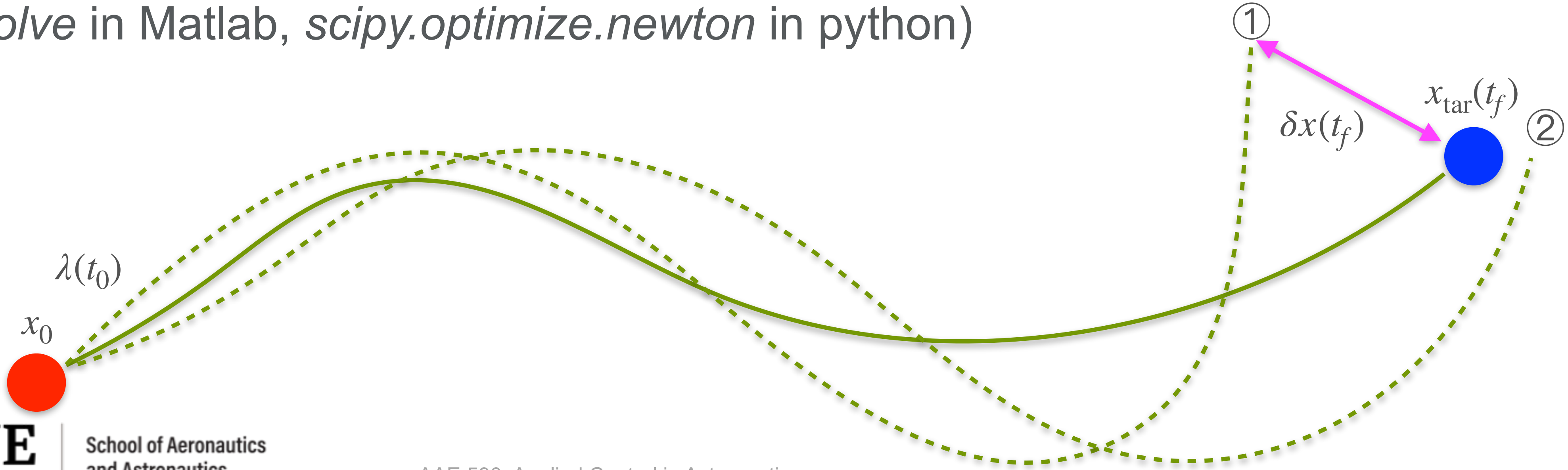
$$\psi = 0 \text{ (if } d\nu \neq 0), \quad \nu^\top \psi_{x_0} + \lambda_0^\top = 0 \text{ (if } dx_0 \neq 0), \quad \nu^\top \psi_{x_f} - \lambda(t_f)^\top = 0 \text{ (if } dx_f \neq 0), \quad \nu^\top \psi_{t_0} - H(t_0) = 0 \text{ (if } dt_0 \neq 0), \quad \nu^\top \psi_{t_f} + H(t_f) = 0 \text{ (if } dt_f \neq 0)$$

Minimum-time Low-thrust Orbit Transfer (cont'd)

- How many unknowns for how many equations?

$$Z = \begin{bmatrix} \lambda(t_0) \\ t_f \end{bmatrix} \qquad \Psi = \begin{bmatrix} x_f - x_{\mathrm{tar}}(T) \\ H(t_f) - [\lambda(t_f)]^\top [\dot{x}_{\mathrm{tar}}(t_f)] \end{bmatrix} \qquad x_f = x_0 + \int_{t_0}^{t_f} f(x(t), u^*(t), t)\mathrm{d}t, \quad u^* = \frac{p}{\|p\|_2} u_{\max}$$

- Indirect method to solve TPBVP
  - Find a Z that satisfies $\Psi(Z) = 0$ via a nonlinear root-finding algorithm (e.g., *fsolve* in Matlab, *scipy.optimize.newton* in python)



PURDUE UNIVERSITY School of Aeronautics and Astronautics

Minimum-energy Low-thrust Orbit Transfer (no control constraint)

- Control Hamiltonian: $H = L + \lambda^\top f = \|u\|_2^2 + \lambda^\top f$
- Optimality necessary conditions:
  - State dynamics

    $\dot{x} = f$

    Costate dynamics

    $\dot{\lambda}^\top = -H_x = -\lambda^\top f_x$

    Optimal Control

    $$u^* = \arg\min H$$
    $$= \arg\min \, [\|u\|_2^2 + \lambda^\top f(x^*, u)]$$

  - Analytical optimal control law:

    $$H_u = 0 \wedge H_{uu} \succ 0 \;\Leftrightarrow\; 2u^{*\top} + \lambda^\top B = 0 \;\wedge\; 2I_2 \succ 0 \;\Leftrightarrow\; u^* = -\frac{1}{2}B^\top \lambda = \frac{1}{2}p$$

  - Transversality conditions:

    $$\psi = 0$$

Minimum-energy Low-thrust Orbit Transfer (w/ control constraint)

- Control Hamiltonian: $H = L + \lambda^\top f = \|u\|_2^2 + \lambda^\top f$
- Optimality necessary conditions:
  - State dynamics $\dot{x} = f$
  - Costate dynamics $\dot{\lambda}^\top = -H_x = -\lambda^\top f_x$
  - Optimal Control $u^* = \arg\min_{\|u\|_2 \le u_{\max}} [\|u\|_2^2 + \lambda^\top f(x^*, u)]$
  - Analytical optimal control law:
    - Unconstrained optimal control $u^* = p/2$ may violate the magnitude constraint
    - Re-parameterize the control and apply PMP:

$$u = \Gamma \hat{u}, \quad \text{where} \quad \Gamma \in [0, u_{\max}], \quad \|\hat{u}\|_2 = 1$$

$$\{\Gamma^*, \hat{u}^*\} = \arg\min_{\Gamma, \hat{u}} H = \arg\min_{\Gamma} \min_{\hat{u}} \Gamma^2 (1 - p^\top \hat{u})$$

$$\Rightarrow \quad u^* = \Gamma^* \hat{u}^*, \quad \hat{u}^* = \frac{p}{\|p\|_2}, \quad \Gamma^* = \begin{cases} \frac{1}{2}\|p\|_2 & (\|p\|_2 \le 2u_{\max}) \\ u_{\max} & (\|p\|_2 > 2u_{\max}) \end{cases}$$

  - Transversality conditions:

$$\psi = 0$$

Minimum-fuel Low-thrust Orbit Transfer

- Control Hamiltonian: $H = L + \lambda^\top f = \|u\|_2 + \lambda^\top f$
- Optimality necessary conditions:
  - State dynamics $\dot{x} = f$
    Costate dynamics $\dot{\lambda}^\top = -H_x = -\lambda^\top f_x$
    Optimal Control $u^* = \arg \min_{\|u\|_2 \le u_{\max}} [\|u\|_2 + \lambda^\top f(x^*, u)]$
  - Analytical optimal control law:

$$u^* = \arg \min_{\|u\|_2 \le u_{\max}} \|u\|_2 - p^\top u \qquad u = \Gamma \hat{u}, \quad \text{where} \quad \Gamma \in [0, u_{\max}], \quad \|\hat{u}\|_2 = 1$$

$$\{\Gamma^*, \hat{u}^*\} = \arg \min_{\Gamma} \min_{\hat{u}} \Gamma(1 - p^\top \hat{u}) \qquad \hat{u}^* = \frac{p}{\|p\|_2}, \qquad \Gamma^* = \arg \min_{\Gamma} \Gamma(1 - \|p\|_2)$$

$$u^* = \Gamma^* \hat{u}^*, \quad \hat{u}^* = \frac{p}{\|p\|_2}, \quad \Gamma^* = \begin{cases} u_{\max} & (\|p\|_2 > 1) \\ 0 & (\|p\|_2 < 1) \end{cases}$$

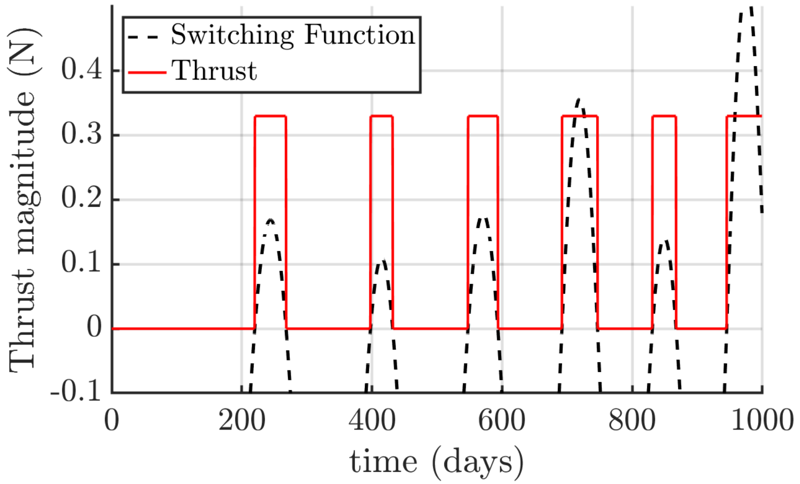
  - Transversality conditions: $\psi = 0$

Minimum-fuel Low-thrust Orbit Transfer (cont'd)

- Issue:
  - Bang-bang control profile (discontinuous)

$$\Gamma* = \begin{cases} u_{\max} & (\|p\|_2 > 1) \\ 0 & (\|p\|_2 < 1) \end{cases} = \frac{u_{\max}}{2}\left[1+\mathrm{sign}(S)\right], \quad S = \|p\|_2 - 1$$

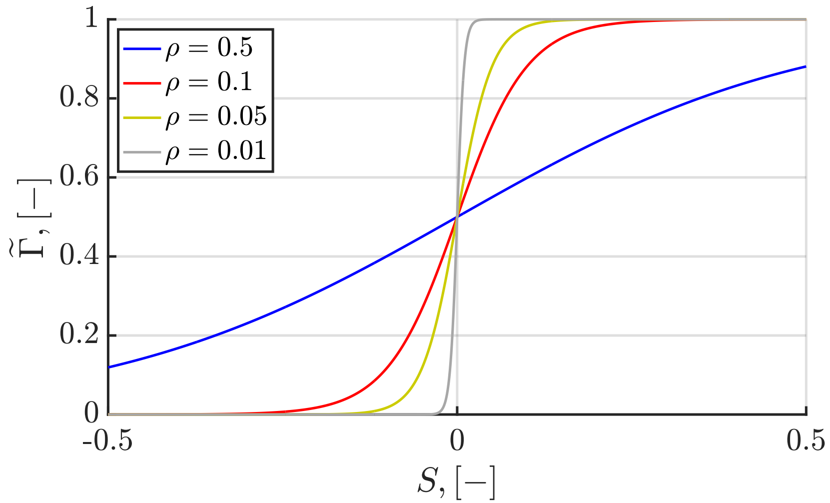
  - Causes numerical issues in nonlinear-root finding (gradient-based; Newton's method)



Earth-Venus optimal transfer example
(Y.Sidhoum & K.Oguri, 2023)

- Remedies:
  - Homotopic approach
  - Smoothing (e.g., hyperbolic tangent)

$$\tilde{\Gamma}* = \frac{u_{\max}}{2}\left[1+\tanh\left(\frac{S}{\rho}\right)\right]$$

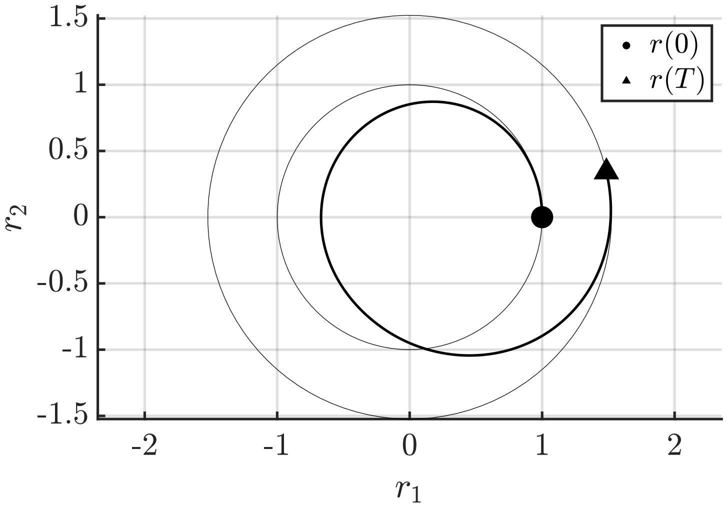


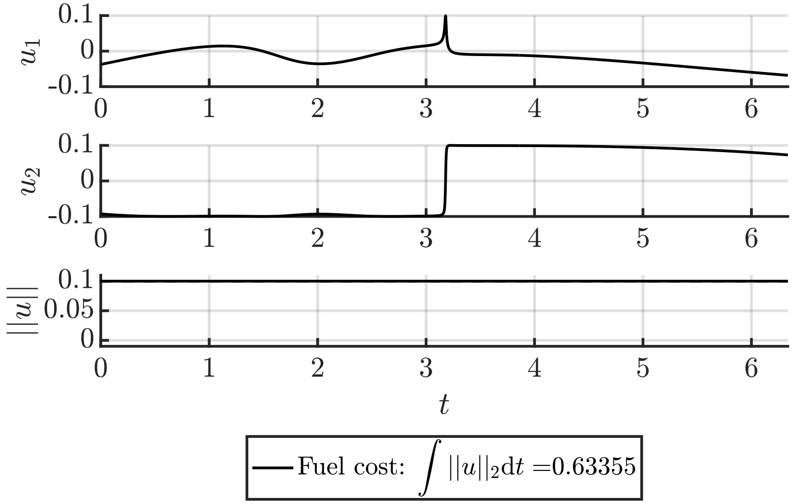
*for performance comparison of the two popular remedies, see Y.Sidhoum & K.Oguri, On the Performance of Different Smoothing Methods for Indirect Low-thrust Trajectory Optimization, 2023

Optimal orbit transfer example: Earth-Mars transfer

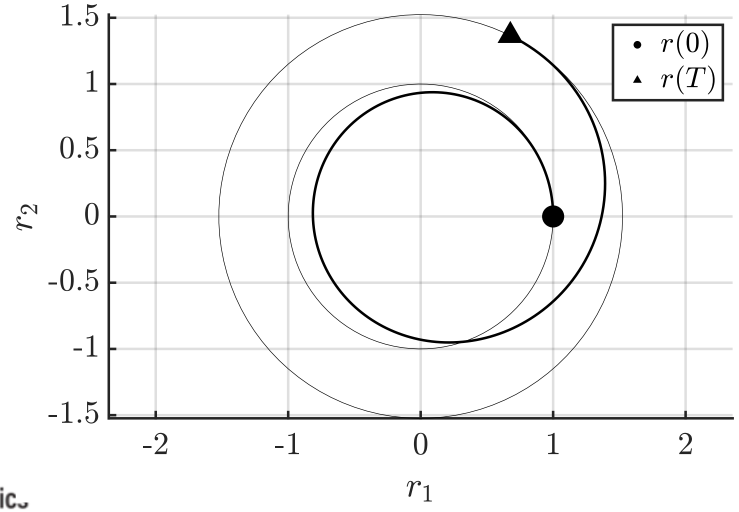
Earth-Mars transfer

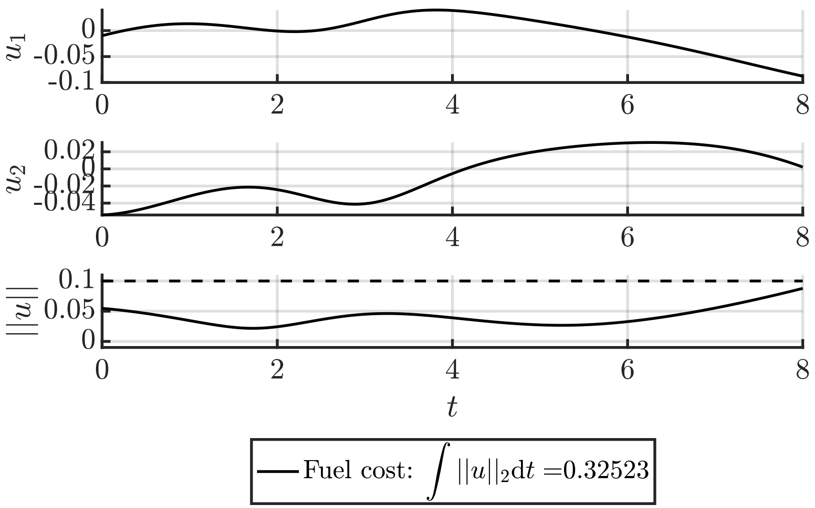
min-time problem:
- ToF: 6.34 (non-dim)
- Fuel cost: 0.63 (non-dim)





min-energy problem:
- ToF: 8.0 (non-dim)
- Fuel cost: 0.33 (non-dim)



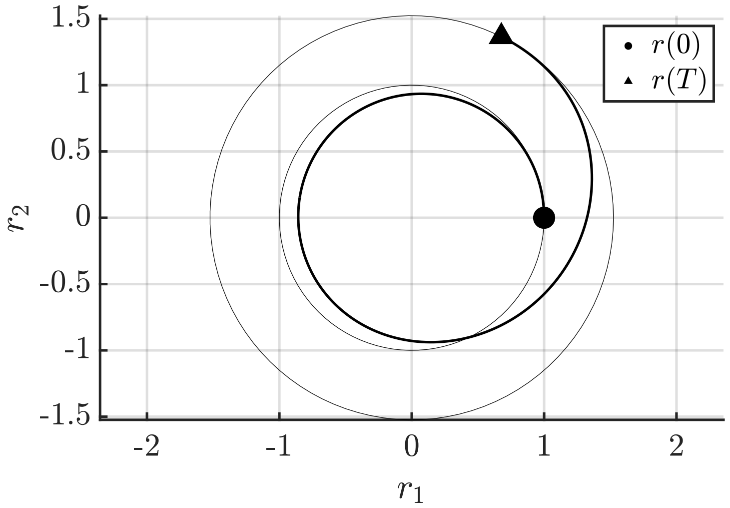


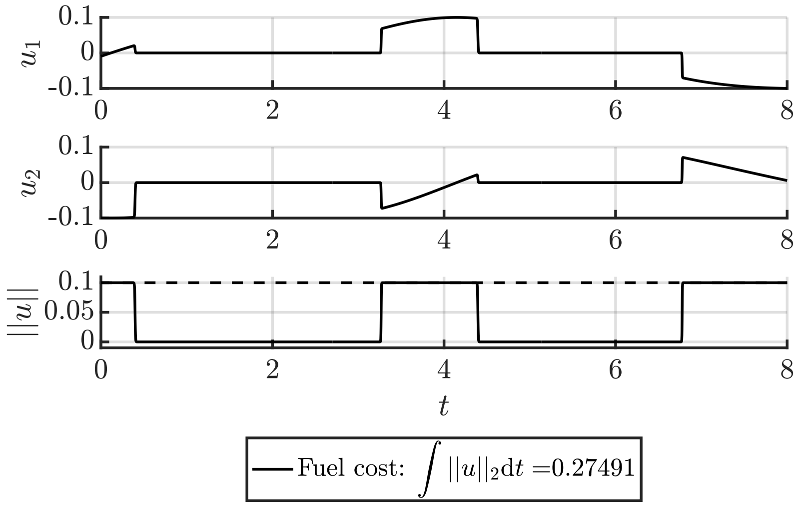
PURDUE UNIVERSITY School of Aeronautics and Astronautics

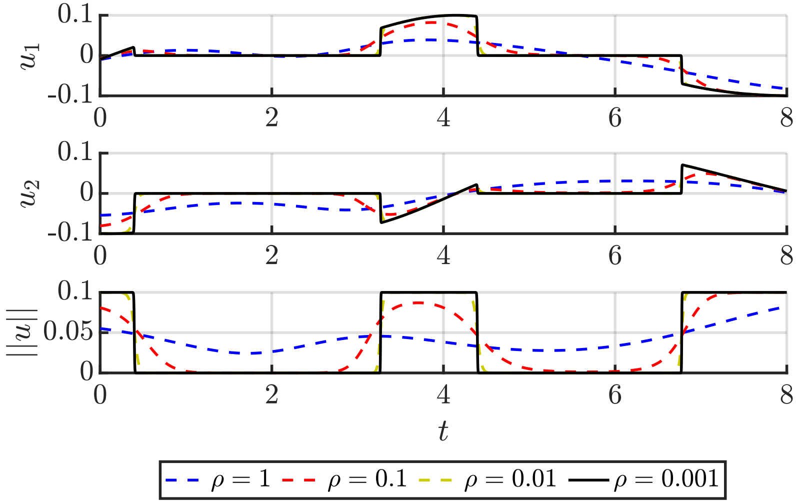
Optimal orbit transfer example: Earth-Mars transfer (cont'd)

Earth-Mars transfer

min-fuel problem
(hyperbolic tangent smoothing):
- ToF: 8.0 (non-dim)
- Fuel cost: 0.27 (non-dim)







Implementation tips for better convergence

- Scaling / non-dimensionalization
  - All variables (e.g., state, costate, time) should be scaled to be in the order of 1
    - They can be scaled back to the original units after optimization
  - A common scaling for interplanetary transfers: 1 unit length = AU, GM = 1
    - Scaling for time and velocity can be determined accordingly
- Costate initial guess
  - Indirect method is highly sensitive to the initial costate guess; many studies exist
- Providing analytic derivatives
  - Can dramatically improve convergence, although require more analytical efforts
- Solver settings
  - Use appropriate tolerances (which should be greater than the integration tolerance); give large enough max iterations/function evaluations

Advanced topics

- Relevant studies:
  - Solution methods: initial costate guess techniques[1], multiple shooting[2], forward-backward shooting[3], hybrid indirect/direct method[4]
  - Different low-thrust propulsion systems: variable Isp engine[5], solar sail[6]
  - Guidance applications: neighboring optimal control[7]
  - Optimal control theory: survey paper[8]
  - …

1. H.Yan and H.Wu, Initial Adjoint-Variable Guess Technique and Its Application in Optimal Orbital Transfer, AIAA JGCD 1999
2. Y.Meng, et.al., Low-Thrust Minimum-Fuel Trajectory Optimization Using Multiple Shooting Augmented by Analytical Derivatives, AIAA JGCD 2019
3. Y.Sidhoum & K.Oguri, Indirect Forward-Backward Shooting for Low-thrust Trajectory Optimization in Complex Dynamics, AAS SFM 2023
4. B.Pierson and C.Kluever, Three-Stage Approach to Optimal Low-Thrust Earth-Moon Trajectories, AIAA JGCD 1994
5. L.Casalino and G.Colasurdo, Optimization of Variable-Specific-Impulse Interplanetary Trajectories, AIAA JGCD 2004
6. K.Oguri, et.al., Solar Sailing Primer Vector Theory: Indirect Trajectory Optimization with Practical Mission Considerations, AIAA JGCD 2022
7. H. Seywald and E.Clif, Neighboring Optimal Control Based Feedback Law for the Advanced Launch System, AIAA JGCD 1994
8. R.Hartl, et.al., A Survey of the Maximum Principles for Optimal Control Problems with State Constraints, SIAM Review 1995

Summary

- Optimal orbit transfers
  - Direct application of optimal control theory (calculus of variations + Pontryagin)
  - Lawden's primer vector theory: derived via PMP, costate analytically characterizes the optimal low-thrust profile
  - Convert continuous-time optimal control problem (infinite-dimensional optimization problem) to TPBVP and solve via nonlinear root-finding algorithm; called *indirect method*
  - Low-thrust orbit transfers with minimum- time/energy/fuel objectives demonstrated
- Notes:
  - Deriving an analytical optimal controller may not be easy depending on the problem (e.g., different propulsion system, complex constraints, different objective functions)
  - Numerical solution to nonlinear TPBVP yields local optimum (not global)
  - Indirect method largely reduces the optimization variables; much less variables compared with direct method-based parameter trajectory optimization — trade off with the high sensitivity to the initial costate ("good" initial guesses are necessary for convergence)