Lecture notes for MAE 274 (Optimal Control)
Mechanical and Aerospace Eng. Dept.,
University of California Irvine

Solmaz Kia
http://solmaz.eng.uci.edu/Teaching/mae274.html

Contents

Note 1

The treatment corresponds to selected parts from Chapters 1 and 2 of [1] and Chapter 1 of [2]. The presentation is at times informal. For rigorous treatments, students should consult the aforementioned references and the other listed texts in the class syllabus.

1.1 Introduction

Control theory is a branch of applied mathematics that involves basic principles underlying the analysis and design of (control) systems/processes. Systems can be engineered physical systems (e.g., air conditioner, aircraft, CD player etc.), economic systems, biological systems and so on. Every system is usually driven by its internal dynamics and forces or other forms of external influences that are either actively (intentionally) or passively exerted on it. To control means that one uses the active input channels of the system to influence the behavior of the system in a desirable way: for example, in the case of an air conditioner, the aim is to control the temperature of a room and maintain it at a desired level, while in the case of an aircraft, we wish to control its altitude at each point of time so that it follows a desired trajectory.

Systems are normally described by underdetermined differential equations. This means that there is some freeness in the choice of the variables satisfying the differential equation. To comprehend the underdetermined-ness, consider the underdetermined algebraic equation is $x + u = 10$, where x, u are positive integers. There is freedom in choosing, say u, and once u is chosen, then x is uniquely determined. In the same manner, consider the differential equation

$$\dot{x} = f(x(t), u(t)), \quad x(t_i) = x_i, \quad t \geq t_i, \tag{1.1.1}$$

where $x(t) \in \mathbb{R}^n$, and $u(t) \in \mathbb{R}^m$.

The objective in control theory

- Choose the control inputs to achieve stabilization and regulation of the system state variables. For instance, we might want the state x to track some desired reference state x_d, and there must be stability under external disturbances.

- Impose performance on system behavior. The objective of optimal control is to determine the control signals that will cause a process to satisfy the physical constraints and at the same time minimize (or maximize) some performance criterion.

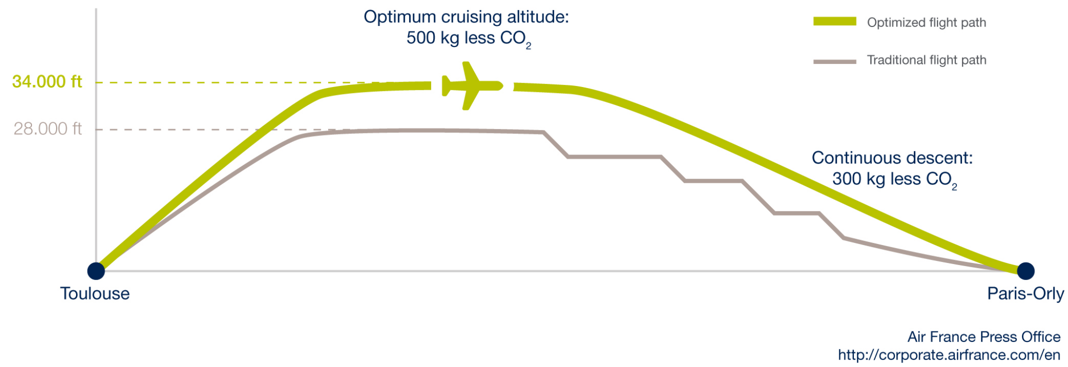


Figure 1: An example of optimized response

You are familiar with some performance measures such as rise time, settling time, peak overshoot, gain and phase margin and bandwidth for SISO systems. In this course, we will learn tools that allow us to enforce performance measures that are more closely tied to the physical response of the system, e.g., minimum energy, minimum fuel etc. Also, these techniques are applicable to MIMO systems, as well.

1.2 Optimal control formulation

The following three elements constitute the optimal control formulation. Each of these elements will be discussed in the proceeding subsections.

- model (a mathematical description) of the process/system to be controlled
- mathematical description of the (physical) constraints of the system
- a performance measure and its mathematical description

In this section we also discuss the form of the optimal control.

1.2.1 Model of the system

In this class we mainly deal with time-invariant systems whose model is described by

$$\dot{x} = f(x(t), u(t)), \tag{1.2.1}$$

where $x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} \in \mathbb{R}^n$ is the state vector and $u(t) = \begin{bmatrix} u_1(t) \\ x_2(t) \\ \vdots \\ u_m(t) \end{bmatrix} \in \mathbb{R}^m$ is the control vector of the system.

Consider system (1.2.1) over some time interval $t \in [t_0, t_f]$.

Definition 1 (Control history)**.** A history of control input values during the interval $[t_0, t_f]$ is denoted by $\mathbf{u}$ and is called a *control history*, or simply a *control*. □

Definition 2 (State trajectory)**.** A history of state values during the interval $[t_0, t_f]$ is denoted by $\mathbf{x}$ and is called a *state trajectory*. □

Controllability and observability are two important properties of a system in controller design. In the most of the problems we consider throughout this course, the goal is to find a controller that transfers a system from an arbitrary initial state to the origin while minimizing a performance measure. Therefore, the controllability of the system is a necessary condition for the existence of such a solution. In closed-loop controller design we use the output of the system to design the controller. The output of the system can be only a subset of the state of the system or function of a combination of them. In order to control all the states of the system, the states should be observable from the knowledge about output and control input of the system. For full treatment of the controllability and observability concepts refer to books on Linear System Theory or your notes from MAE270A class.

Definition 3 (Controllability)**.** If there is a finite time $t_1 \geq t_0$ and a control $u(t)$, $t \in [t_0, t_1]$, which transfers the state x_0 to the origin at time t_1, the state x_0 is said to be *controllable at time* t_0. If all values of x_0 are controllable for all t_0, the system is *completely controllable*, or simply *controllable*. □

Definition 4 (Observability)**.** If by observing the output $y(t)$ during the finite time interval $[t_0, t_1]$ the state $x(t_0) = x_0$ can be determined, the state x_0 is said to be *observable at time* t_0. If all states x_0 are observable for every t_0, the system is called *completely observable*, or simply *observable*. □

When the system function f is linear, that is $f(x, u) = Ax(t) + Bu(t)$ for some $A\mathbb{R}^{n\times n}$ and $B\mathbb{R}^{n\times m}$, the system is said to be linear. The trajectories of linear system

$$\dot{x} = Ax + Bu, \quad x(t_0) = x_0,$$

are described by

$$x(t) = \mathrm{e}^{A(t-t_0)}x_0 + \int_{t_0}^{t_f} \mathrm{e}^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau, \quad t \geq t_0,$$

A linear system is controllable if and only if

$$\operatorname{rank}(\mathcal{C}) = \operatorname{rank}\Big(\begin{bmatrix} B & AB & A^2B & \cdots & A^{n-1}B \end{bmatrix}\Big) = n.$$

A linear system is observable if and only if

$$\operatorname{rank}(\mathcal{O}) = \operatorname{rank}\left(\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}\right) = n.$$

Continuous-time system (1.2.1) over some time interval $[t_0, t_f]$ can be approximated by a discrete system by considering N equally spaced time increments in the time interval as follows

$$\frac{x(t+\Delta t) - x(t)}{\Delta t} \approx f(x(t), u(t)),$$

or

$$x(t+\Delta t) = x(t) + \Delta\, f(x(t), u(t)).$$

Using the shorthand notation $x(k) = x(k\Delta t)$ we can write then

$$x(k+1) = x(k) + \Delta\, f(x(k), u(k)),$$

which we will denote by

$$x(k+1) = f_D(x(k), u(k)).$$

When the system function f_D is linear, that is $f(x(k), u(k)) = A(k)x(k) + B(k)u(k)$ for some $A(k)\mathbb{R}^{n\times n}$ and $B(k)\mathbb{R}^{n\times m}$, $k \in \mathbb{Z}_{\geq 0}$, the system is said to be linear. If the system and control matrix are constant for all $k \in \mathbb{Z}_{\geq 0}$, the system is said to be linear time-invariant

$$x_k = Ax_k + Bu_k, \quad x(0) = x_0.$$

The trajectories of this system is described by

$$x_k = A^k x_0 + \sum\nolimits_{i=0}^{k-1} A^{k-i-1} B u_i, \quad k \in \mathbb{Z}_{\geq 0}.$$

1.2.2 Constraints on system state and control

The control and states of a system can be constrained for various reasons. A common form of control constraint is the saturation constrains which is due to the limited control authority of physical system. This constraint is described as

$$\mathsf{u}_i^{\min} \leq u_i(t) \leq \mathsf{u}_i^{\max}, \quad i \in \{1, \dots, m\}, t \in [0, t_f].$$

The states of a system can be constrained, as well. For example in control of an aircraft, the state describing the angle of attack of the system should be constraint by the stall angle (stall angle is an angle which beyond it the flow on the wing is separated and the control over aircraft can be lost). If the angle of attack is a state of the aircraft this constraint can be described by $\alpha^{\min} \leq \alpha(t) \leq \alpha^{\max}$ for all $t \in [0, t_f]$. If the states of the aircraft are pitch angle $\theta(t)$ and the flight path angle (slope angle) $\gamma(t)$ then given $\alpha(t) = \theta(t) - \gamma(t)$ (see Fig. 2), the state constraint can be described as $\alpha^{\min} \leq \theta(t) - \gamma(t) \leq \alpha^{\max}$ for all $t \in [0, t_f]$. Notice that the initial and final conditions $x(t_0) = x_0$ and $x(t_f) = x_f$ are also form of state constraints. The constraint also can be a coupled equation of both state and control variable of a system. For example, consider a car whose equations of motion are described by $\dot{x}_1 = x_2$ and $\dot{x}_2 = u$, where x_1 and x_2 are position and velocity and u is the acceleration control of the car. The fuel consumption of the car is proportional to both acceleration and speed of the car with constants, respectively, γ_1 and γ_2. If the car starts with F gallon of fuel initially, any optimal policy designed for this car while its moving on a straight line from a point A at time t_0 to point B at time t_f is constrained by

$$\int_{t_0}^{t_f} (\gamma_1 |u(t)| + \gamma_2 |x_2(t)|) \mathrm{d}t \leq F.$$

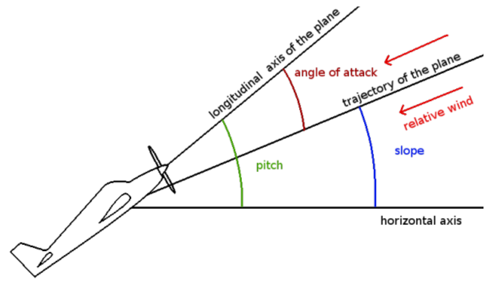


Figure 2: Angle is of an aircraft in vertical plane.

Definition 5 (Admissible control)**.** A control history which satisfies the control constraints during the entire time interval $[t_0, t_f]$ is called an *admissible control*. We denote the set of admissible controls by $\mathcal{U}$ and the notation $\mathbf{u} \in \mathcal{U}$ indicates that the control history $\mathbf{u}$ is admissible. □

Definition 6 (Admissible trajectory)**.** A state trajectory $\mathbf{x}$ which satisfies the state variable constraints during the entire time interval $[t_0, t_f]$ is called an *admissible trajectory*. We denote the set of admissible trajectories by $\mathcal{X}$ and the notation $\mathbf{x} \in \mathcal{X}$ indicates that the state trajectory $\mathbf{x}$ is admissible. □

Admissibility is an important concept. For physical systems acting beyond admissible ranges can result in catastrophic outcomes for system. From technical perspective, admissibility can be useful as it restricts our search region from the entire space to parts defined in admissible ranges.

1.2.3 Performance measure

A performance measure is a mathematical description of the desired behavior we wish the system under study to exhibit. Sometimes capturing this desired behavior is straightforward, e.g., “transfer the system from point A to point B as quickly as possible” clearly means that the performance measure is the time elapsed to go from point A to point B starting at some time t_0. In other times, description of the desired behavior may be qualitative, and leaves some room for the designer to decide how to design the performance measure, “Maintain the position and velocity of the system close to zero with a small expenditure of control energy”.

Some of the common forms of the performance measure are as follows.

- **Minimum-time problem**: To transfer a system from arbitrary initial state $x(t_0) = x_0$ to a specified target set $\mathcal{S}$ in minimum time

$$J = t_f - t_0 = \int_{t_0}^{t_f} \mathrm{d}t, \tag{1.2.2}$$

  where t_f is the first instant of time when $\mathbf{x}(t)$ and $\mathcal{S}$ intersect.

  For discrete-time systems, minimum-time performance can be cast as

$$J = N = \sum_{k=0}^{N-1} 1.$$

- **Terminal control problem**: to minimize the deviation of the final state of a system from its desired value $r(t_f) \in \mathbb{R}^n$

$$J = \sum\nolimits_{i=1}^{n}(x_i(t_f) - r_i(t_f))^2 = (x(t_f) - r(t_f)^\top(x(t_f) - r(t_f)) = \|x(t_f) - r(t_f)\|^2. \quad (1.2.3)$$

Notice that the error is squared as both positive and negative deviations are undesirable. Given the system model and the constrains, $x(t_f) = r(t_f)$ may not be accomplished. In this case, we may wish to put more weight or penalty on the deviation of certain state more than others. We can realize such a wish by inserting a symmetric positive semi-definite $n \times n$ matrix H to obtain

$$J = (x(t_f) - r(t_f))^\top H\,(x(t_f) - r(t_f)) = \|x(t_f) - r(t_f)\|_H^2. \quad (1.2.4)$$

Notice that the elements of H should be adjusted to normalize the numerical values encountered. A good example of this kind of adjustment is given in [1] as follows. Consider the ballistic missile shown in Fig. 3. The position of the missile at time t is specified by the spherical coordinate $l(t)$, $\alpha(t)$ and $\theta(t)$, which are, respectively, the distance from origin of the coordinate system, the elevation angle and azimuth angle. If $l(tf) = 5000$ miles, an azimuth error of 0.01 radian (0.57 degree) results in missing the target S by 50 miles!. If the performance measure is

$$J = h_{11}(l(t_f) - 5000)^2 + h_{22}(\theta(t_f))^2,$$

then we would select $h_{22} = (50/0.01)^2\,h_{11}$ to weight equally deviation in range and azimuth. Alternatively, the variables θ and l could be normalized in which case we will use $h_{11} = h_{22}$.

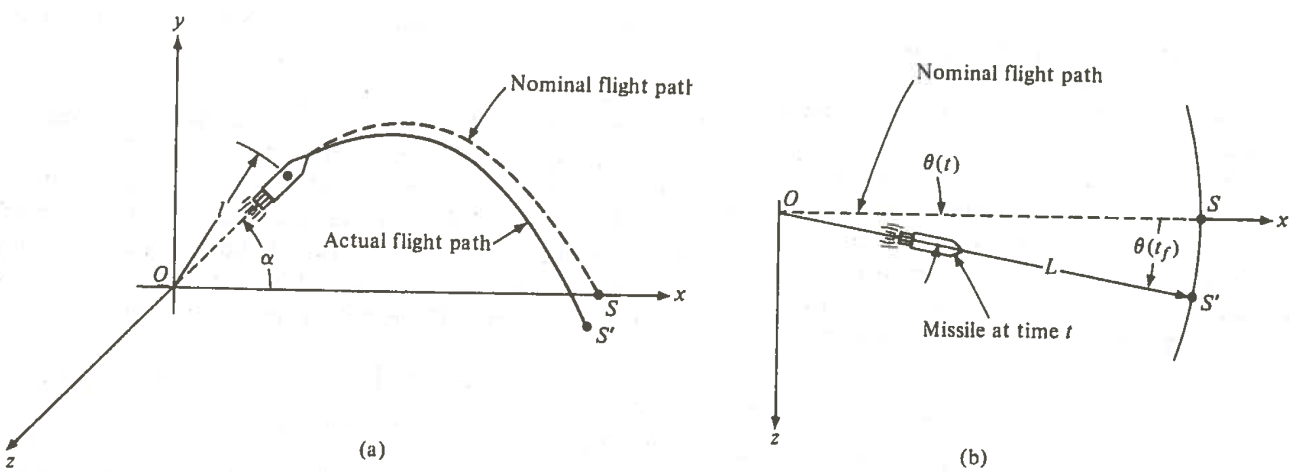


Figure 3: A ballistic missile aimed at target S (photo courtesy of [1]).

- **Minimum-control-effort problems**: to transfer a system from an arbitrary initial state $x(t_0) = x_0$ to a specified target S, with a minimum expenditure of control effort. Given any physical application, the form of the cost function depends on how one interprets " minimum control effort". For example, for a space craft with control input $u(t)$ as the thrust of the engine whose magnitude is proportional to the fuel consumption of the engine, the minimum-control-effort cost can be cast as

$$J = \int_{t_0}^{t_f} |u(t)| \, \mathrm{d}t.$$

For a discrete-time system with single input u_k, to drive the system from x_0 to a desired final state x at a fixed time N using minimum fuel, we could use

$$J = \sum\nolimits_{k=0}^{N-1} |u_k|.$$

As another example, consider a electric network without energy storage element. Let a voltage source $u(t)$ control this network. The minimum source energy dissipation energy for this problem can be with a cost function

$$J = \int_{t_0}^{t_f} u^2(t) \, \mathrm{d}t.$$

For several control inputs, we can write the cost function as

$$J = \int_{t_0}^{t_f} u^\top(t) R u(t) \, \mathrm{d}t = \int_{t_0}^{t_f} \|u(t)\|_R^2 \, \mathrm{d}t,$$

where $R \geq$ is a weighting matrix with real elements. The elements of R can be function of time if we wish to vary the weighting on control-effort expenditure over $[t_0, t_f]$.

For a discrete-time system, to drive the system from x_0 to a desired final state x_N at a fixed time N with minimum energy, we could use

$$J = \frac{1}{2} x_N^\top H x_N + \frac{1}{2} \sum\nolimits_{k=0}^{N-1} (x_k^\top Q x_k + u_k^\top R u_k).$$

where $H, Q, R \geq 0$ are positive semi-definite weighting matrices.

- **Tracking problem**: to maintain the system state $x(t)$ as close as possible to the desired state $r(t)$ in the interval $[t_0, t_f]$. The performance measure in this case is

$$J = \int_{t_0}^{t_f} (x(t) - r(t))^\top(t) Q (x(t) - r(t)) \, \mathrm{d}t = \int_{t_0}^{t_f} \|x(t) - r(t)\|_Q^2 \, \mathrm{d}t, \tag{1.2.5}$$

where $Q \geq$ is a weighting matrix with real elements. The elements of Q can be function of time if we wish to vary the weighting on closeness to the reference signal over $[t_0, t_f]$.

If we are minimizing the cost function subject to set of constraints that includes bounded inputs, e.g., $|u_i(t)| \leq 1$ for $i \in \{1, \ldots, m\}$, then the cost function (1.2.5) is a reasonable performance measure. However, if the control is not bounded, using the cost function (1.2.5) may result in impulses in control and its derivatives. To remove the hard control bounds from problem formulation or conserve energy while maintaining tracking it is customary to modify (1.2.5) as follows

$$J = \int_{t_0}^{t_f} (\|x(t) - r(t)\|_{Q(t)}^2 + \|u(t)\|_{R(t)}^2) \, \mathrm{d}t,$$

In some applications it is especially important that the states be close to their desired value at final time. In this case, the cost function is described by

$$J = \|x(t_f) - r(t_f)\|_H^2 + \int_{t_0}^{t_f} (\|x(t) - r(t)\|_{Q(t)}^2 + \|u(t)\|_{R(t)}^2)\,\mathrm{d}t.$$

where $H \geq 0$ is the weighting matrix with real elements.

- **Regulator problem**: A regulator problem is a special case of tracking problem where the reference signal is $r(t) = 0$ for $t \in [t_0, t_f]$.

All the performance measures discussed above are special cases of the general form

$$J = \underbrace{h(x(t_f), t_f)}_{\text{terminal cost}} + \underbrace{\int_{t_0}^{t_f} g(x(t), u(t), t)\mathrm{d}t}_{\text{running cost}}. \tag{1.2.6}$$

For discrete-time systems, the general form of the cost function can be cast as

$$J = \underbrace{\phi(x_N, N)}_{\text{terminal cost}} + \underbrace{\sum\nolimits_{k=0}^{N-1} L^k(x_k, u_k)\mathrm{d}t}_{\text{running cost}}. \tag{1.2.7}$$

1.2.4 The optimal control problem

Our objective in solving an optimal control is to find an admissible $u^\star$ that causes $\dot{x} = f(x(t), u(t), t)$ to follow an admissible $x^\star$ that minimize

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

- $u^\star$: optimal control $\quad$ $x^\star$: optimal trajectory

$$\begin{aligned} J^\star &= h(x^\star(t_f), t_f) + \int_{t_0}^{t_f} g(x^\star(t), u^\star(t), t)\mathrm{d}t \\ &\leq h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x(t), u(t), t)\mathrm{d}t, \quad \mathbf{u} \in \mathcal{U},\ \mathbf{x} \in \mathcal{X}. \end{aligned}$$

- We are looking for global minimum
  - Find all local minimum, and pick the smallest as global minimum
- Solution is not unique
  - con: complicates computational procedures
  - pro: choose among multiple possibilities accounting for other measures

1.2.5 Form of the optimal control

Let $u^{\star}(t)$ denote the optimal control that minimizes the performance measure (cost function) subject to system and control constraints.

Definition 7 (Optimal control law)**.** If there exists a functional form

$$u^{\star}(t) = f(x(t), t), \tag{1.2.8}$$

that describes the optimal control at time $t \in [t_0, t_f]$, then the function f is called the *optimal control law*, or the *optimal policy*.

Notice here that the reason for not using $x^{\star}(t)$, the optimal state, in (1.2.8) is to emphasis that the control law is optimal for all admissible $x(t)$, not just for some special state value at time t. Function f is a rule that specifies the optimal control at time t for any (admissible) stat value at that time. If

$$u^{\star}(t) = K\,x(t),$$

where $K \in \mathbb{R}^{m\times n}$, the control law is called linear time-invariant feedback of the state.

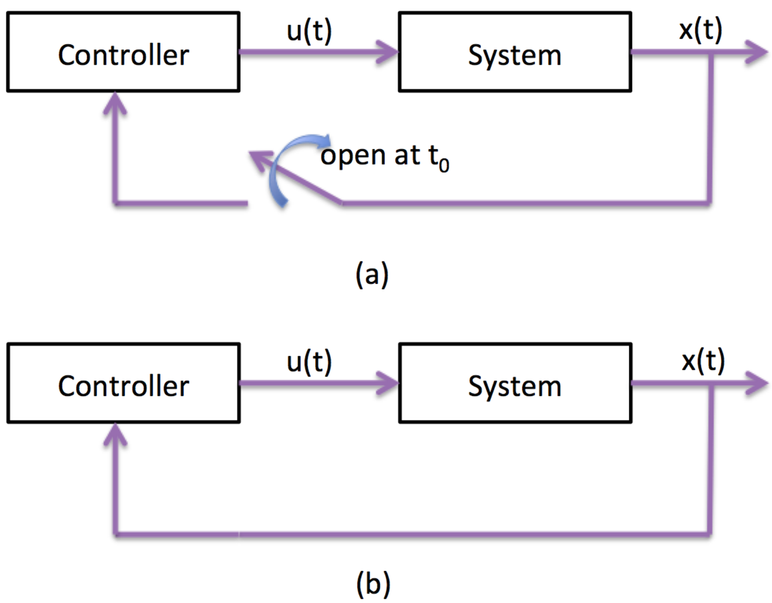


Figure 4: (a) Open-loop optimal control, (b) optimal control

Definition 8 (open-loop control)**.** If the optimal control is determined as a function of time for a specified initial state $x(t_0)$, i.e.,

$$u^{\star}(t) = f(x(t_0), t), \quad t \in [t_0, t_f],$$

then the optimal control is said to be in *open-loop* form.

Therefore, the optimal open-loop control is optimal only for a *particular* initial state value, whereas, if the optimal control *law* is known, the optimal control history starting from any state value can be generated. Fig. 4 demonstrates the conceptual difference between an optimal control law and an open-loop control. Notice that the mere presence of connection from the states to a controller does not, in general, guarantee an *optimal* control law.

Although closed-loop control solutions are normally more desirable in optimal control, there are cases that open-loop control may be feasible and an adequate answer. For example, in the radar tracking of a satellite, once the orbit is set very little can happen to cause an undesired change in the trajectory parameters. In this situation a pre-programmed control for the radar antenna might well be used.

A typical example of feedback control is in the classic servomechanism problem where the actual and desired outputs are compared and any deviation produces a control signal that attempts to reduce the discrepancy to zero.

Note 2

The treatment corresponds to selected parts from Chapter 1 of [3] and Chapters 7 and 8 of [4]. The presentation is at times informal. For rigorous treatments, students should consult the aforementioned references and the other listed texts in the class syllabus.

2.1 Parameter optimization

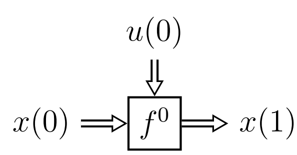
In this section we discuss parameter optimization for static problems, i.e., when time is not a parameter. The discussion is preparatory to dealing with time-varying systems in subsequent lectures. To introduce important concepts, mathematical style, and notation, the parameter minimization problem is formulated and conditions for local optimality are determined. By local optimality we mean that optimality can be verified about a small neighborhood of the optimal point. First, the notions of first- and second order local necessary conditions for unconstrained parameter minimization problems are derived. Next, the notion of first- and second-order local necessary conditions for parameter minimization problems is extended to include algebraic equality constraints.

An example of static problems in optimal control is controller design for multi-stage systems over finite horizon.

- Single-stage system: consider a system with n states $x \in \mathbb{R}^n$ initially at known state $x(0) = x_0 \in \mathbb{R}^n$. A choice of a m dimensional input $u(0) \in \mathbb{R}^m$, takes the state of the system to stage $x(1)$. The transition equation is described as follows

$$x(1) = f^0(x(0), u(0)), \quad x(0) = x_0 \in \mathbb{R}^n, \tag{2.1.1}$$

which is shown schematically as follows



Optimal control problem for this system can be described as: we wish to choose $u(0)$ to minimize a performance index of the form

$$J(u) = \phi(x(1)) + L^0(x(0), u(0)),$$

subject to equality constraint (2.1.1).

- Multi-stage system with no terminal cost: consider a multi-stage system with n states $x \in \mathbb{R}^n$ initially at known state $x(0) = x_0 \in \mathbb{R}^n$. At every stage $i \in \{0, \cdots, N-1\}$, a choice of a m dimensional input $u(i) \in \mathbb{R}^m$, takes the state of the system from $x(i)$ to stage $x(i+1)$. The transition equations are described as follows

$$x(i) = f^i(x(i), u(i)), \quad x(0) = x_0 \in \mathbb{R}^n \quad i \in \{0, \cdots, N-1\}, \tag{2.1.2}$$

which can be represented schematically as follows

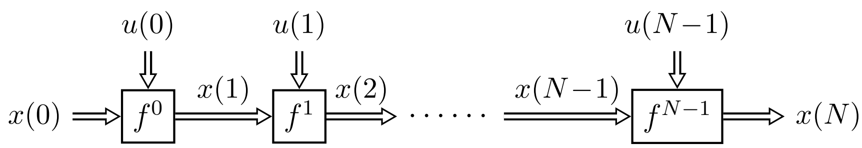


Figure 5: Schematic representation of a multi-stage system over N steps.

Optimal control problem for this system can be described as: we wish to choose sequence of control inputs $\{u(0), u(1), \cdots, u(N-1)\}$ that minimizes (or maximizes) a performance index of the form

$$J(u) = \phi(x(N)) + \sum\nolimits_{i=1}^{N-1} L^i(x(i), u(i)),$$

subject to equality constraints (2.1.2).

2.1.1 Unconstrained static optimization

Consider $F(u)$, where $F : \mathbb{R}^m \to \mathbb{R}$ is a differentiable cost function and $u \in \mathbb{R}^m$ is the control or decision vector, be a scalar performance index. It is desired to determine the value of u that results in a minimum value of $F(u)$, i.e, to determine

$$u^\star = \underset{u \in \mathbb{R}^m}{\operatorname{argmin}}\; F(u). \tag{2.1.3}$$

In an investigation of the general problem (2.1.3) we distinguish two kinds of solutions: *local minimum points* and *global* minimum point. We also distinguish between *strong (strict)* and *weak* minimum points (see Fig. 8).

Definition 9 (Local minimum)**.** A point $u^\star \in \mathbb{R}^m$ is said to be a *(weak) local minimum point* of F over $\mathbb{R}^m$ if there exists an $\epsilon > 0$ such that $F(u) \geq F(u^\star)$ for all $u \in \mathbb{R}^m$ within a distance ϵ of $u^\star$ (that is, $u \in \mathbb{R}^m$ and $\|u - u^\star\| < \epsilon$). If $F(u) > F(u^\star)$ for all $u \in \mathbb{R}^m$, $u \neq u^\star$, within a distance ϵ of $u^\star$, then $u^\star$ is said to be a *strong (strict) local minimum point* of F over $\mathbb{R}^m$. □

Definition 10 (Global minimum)**.** A point $u^\star \in \mathbb{R}^m$ is said to be a *(weak) global minimum point* of F over $\mathbb{R}^m$ if $F(u) \geq F(u^\star)$ for all $u \in \mathbb{R}^m$. If $F(u) > F(u^\star)$ for all $u \in \mathbb{R}^m$, then $u^\star$ is said to be a *strong (strict) global minimum point* of F over $\mathbb{R}^m$. □

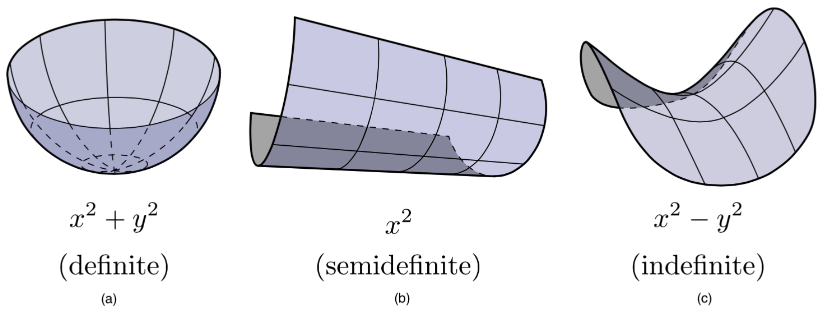


Figure 6: (a) strong local minimum, (b) weak local minimum, (c) saddle point.

We are generally interested in finding global minimum point of a problem, however, the practical reality, from both theoretical and computational viewpoints, in many problems we are made to be content with local minimum points. For example, in deriving necessary and sufficient conditions for optimality we use differential calculus to investigate how the value of the function is changing nearby a minimum point candidate. Or when we use iterative numerical algorithms to search for minimum point, comparison of values of nearby points is all we have, therefore, the point we can identify is a local minimum point. Global conditions and global solutions can, as a rule, only be found if the problem posses certain convexity properties that essentially guarantees that any local minimum is a global minimum.

Next, we derive the first order and second order necessary and sufficient optimality conditions for $u^\star$ in unconstrained optimization problem (2.1.3). These conditions may not be used directly to compute the minimum points but, nevertheless, these conditions form a foundation for the theory, and guidelines to construct iterative algorithms that are used to solve problem (2.1.3).

Let $F(u)$ be twice differentiable. For any point u, $F(u)$ in some small neighborhood of u can be approximated by Taylor series expansion

$$F(u + \mathrm{d}u) = F(u) + F_u(u)^\top \mathrm{d}u + \frac{1}{2}\mathrm{d}u^\top F_{uu}(u)\mathrm{d}u + O(3)$$

, where

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}, \quad F_u = \frac{\partial F}{\partial u} = \begin{bmatrix} \frac{\partial F}{u_1} \\ \vdots \\ \frac{\partial F}{u_m} \end{bmatrix}, \quad F_{uu} = \frac{\partial^2 F}{\partial u \partial u} = \begin{bmatrix} \frac{\partial^2 F}{\partial u_1 \partial u_1} & \cdots & \frac{\partial^2 F}{\partial u_1 \partial u_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial u_m \partial u_1} & \cdots & \frac{\partial^2 F}{\partial u_m \partial u_m} \end{bmatrix}.$$

Definition 11 (critical or stationary point)**.** A stationary or critical point of F is a point that the increment $\mathrm{d}F$ is zero for all increment $\mathrm{d}u$, i.e.,

$$F_u = \frac{\mathrm{d}F}{\mathrm{d}u} = 0. \tag{2.1.4}$$

□

- **First-order optimality condition**: consider the first order (first two terms) of Taylor approximation. Given the ambiguity of sign of the term $F_u(u)^\top \mathrm{d}u$, we can only avoid $F(u + \mathrm{d}u) < F(u)$ if $F_u = 0$.

- Therefore, a *necessary condition* for a point to be minimum is $F_u = 0$.
- Notice that $F_u(u) = 0$ is a necessary and sufficient condition for a point u to be a stationary point, but it is only a necessary condition for u to be a minima. A stationary point can be a maximum or a saddle point. We need to investigate higher order Taylor series approximations to obtain further information.
- Example: Consider the problem

$$\text{minimize } F(u_1, u_2) = u_1^2 - u_1 u_2 + u_2^2 - 3u_2.$$

Setting $F_u = 0$ gives

$$\begin{aligned} 2\,u_1 - u_2 &= 0, \\ -\,u_1 + 2\,u_2 - 3 &= 0, \end{aligned}$$

which gives the unique solution $u_1 = 1$, $u_2 = 2$ as unique global minimum point of F in $\mathbb{R}^2$.

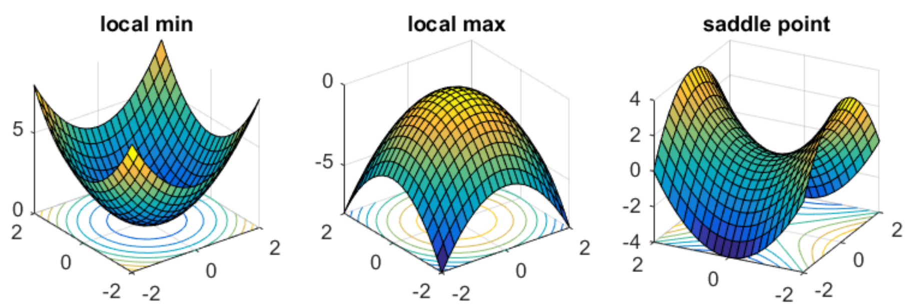


Figure 7: (a) strong minimum, (b) strong maximum , (c) saddle point.

- **Second-order optimality conditions**: suppose that we are at a critical point, that is $F_u(u^\star) = 0$. The Taylor series expansion of $F(u)$ around this point is

$$F(u^\star + \mathrm{d}u) = F(u^\star) + \frac{1}{2}\mathrm{d}u^\top F_{uu}(u^\star)\mathrm{d}u + O(3).$$

  - Using second order approximation, *a necessary condition* for a stationary point to be a minimum point is $\mathrm{d}u^\top F_{uu}(u^\star)\mathrm{d}u \geq 0$ for $\mathrm{d}u$ in all directions, which is equivalent to F_{uu} being positive semi-definite. When $F_{uu} = 0$ the higher order terms can play a role to still make $F(u^\star + \mathrm{d}u) \geq F(u^\star)$.
  - F_{uu} being positive definite ($F_{uu} > 0$), is *a sufficient condition* for the critical point to be a local minima. Notice that when $F_{uu}(u^\star) > 0$, we can write $\frac{1}{2}\mathrm{d}u^\top F_{uu}(u^\star)\mathrm{d}u \geq a\|\mathrm{d}u\|^2$ for some $a > 0$. Therefore we have

$$F(u^\star + \mathrm{d}u) - F(u^\star) = \frac{1}{2}\mathrm{d}u^\top F_{uu}(u^\star)\mathrm{d}u + o(\|\mathrm{d}u\|^2) \geq a\|\mathrm{d}u\|^2 + o(\|\mathrm{d}u\|^2).$$

For small $\|\mathrm{d}u\|$, the first term on the right dominates the second, implying that both sides are positive for small $\mathrm{d}u$.

– If F_{uu} is negative definite, the critical point is a local maxima. If F_{uu} is indefinite, the critical point is a saddle point.
– If F_{uu} is semi-definite, then higher order terms of the Taylor series expansion is needed to determine the type of the critical point (see Fig. 7).

Summary:

Proposition 2.1.1 (Second-order necessary optimality condition for (2.1.3))**.** *Consider the minimization problem (2.1.3). Let $u^\star$ be a local minimum point of twice differentiable function F. Then*

- $F_u(u^\star) = 0,$
- $F_{uu}(u^\star) \geq 0.$

Proposition 2.1.2 (Second-order sufficient optimality condition for (2.1.3))**.** *Consider the minimization problem (2.1.3). Assume that F is twice differentiable and there exists a $u^\star \in \mathbb{R}^m$ such that*

- $F_u(u^\star) = 0,$
- $F_{uu}(u^\star) > 0.$

Then, $u^\star$ is a strong (strict) local minimum point of F.

Iterative solution methods for unconstrained minimization

The first order necessary and sufficient condition for optimality, $F_u(u^\star) = 0$ gives m equations (one for each component of F_u) that can be solved algebraically to obtain the minimum point. However, such an approach may not be feasible for high dimensional or highly nonlinear systems. Typically, iterative algorithms are used to solve minimization problems. That is, starting from some initial guess we take steps in directions which successively reduce the value of the function until we hit a minimum point

$$u(k+1) = u(k) + \mathrm{d}u(k) = u(k) + \alpha(k)s(k), \quad u(0) = u_0,$$

where $s(k) \in \mathbb{R}^m$ is the search direction and $\alpha(k) \in \mathbb{R}$ is the step that we take in the search direction.

We can obtain feasible search directions using Taylor series expansion of the function around the current estimate $u(k)$.

- **Gradient decent algorithm**: Using first order Taylor series approximation we have

$$F(u(k+1)) = F(u(k) + \alpha(k)s(k)) \approx F(u(k)) + \frac{\partial F}{\partial u}\alpha(k)s(k).$$

Let $\alpha(k) > 0$, then to ensure decrease in the function value we can choose

$$s(k) = -(\frac{\partial F}{\partial u})^\top = -F_u(u(k))^\top = -g(u(k)),$$

which is the direction of *steepest decent*.

Gradient decent algorithm: $u(k+1) = u(k) - \alpha(k)\, g(u(k)), \quad u(0) = u_0$.

The desired value for $\alpha(k)$ is a value which will cause maximum reduction in function value at $x(k+1)$. Notice that the gradient decent algorithm is based on the first order approximation of the function. Larger steps can violate our first order approximation's validity and as a result, the desired reduction in the function value may not happen.

Example: Consider the cost function $F(u_1, u_2) = \frac{1}{4}u_1^4 - u_2 u_1 - u_2 + u_2^2$. Let $u(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Consider the gradient decent algorithm for the first step

$$u(1) = u(0) - \alpha(0) F_u(u(0)).$$

Notice that $F_u(u(0)) = \begin{bmatrix} u_1^3 - u_2 \\ -u_1 - 1 + 2u_2 \end{bmatrix}\Big|_{u(0)} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$. Then,

$$u(1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \alpha(0) \begin{bmatrix} 0 \\ -1 \end{bmatrix},$$

which gives $G(\alpha(0)) = F(u(1)) = 0 - 0 - \alpha(0) + \alpha(0)^2$. For minimum reduction, we pick $\alpha(0)$ as follows

$$\frac{\partial G}{\partial \alpha} = -1 + 2\alpha(0) = 0 \Rightarrow \alpha^\star(0) = \frac{1}{2}.$$

Therefore,

$$u(1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix}.$$

Minimizing $G(\alpha(k)) = F(u(k) - \alpha(k) F_u(u(k)))$ with respect to $\alpha(k)$ at each time step to obtain the optimal step size become hard for high dimensional or highly nonlinear cost functions. Because $G(\alpha(k))$ is a univariate function we can use line search methods like Golden section or bisection search methods or polynomial approximations to obtain an estimate of optimal $\alpha(k)$ at each time step. For further details, see [4].

- **Newton and Quasi-Newton algorithms**: Newton and Quasi-Newton algorithms are based on the second order Taylor series approximation. These algorithms typically provide faster termination but they are numerically expensive and complex due to need to compute F_{uu}. For details, see [4].

Matlab fminunc function

'fminunc' is a nonlinear programing solver, which finds the minimum of unconstrained multivariable function. For more details see `http://www.mathworks.com/help/optim/ug/fminunc.html`

- does not need gradients, however, can be faster and more reliable when you provide derivatives
- the choices of algorithms are 'trust-region' (default) or 'quasi-newton'.

- ‘trust-region’ (default). The ‘trust-region’ algorithm requires you to provide the gradient, or else fminunc uses the ‘quasi-newton’ algorithm.
- BFGS Quasi-Newton method with a cubic line search procedure.
- initial guess x_0 is usually very important. The recommended procedure is to try many different initial to look for global minimum.

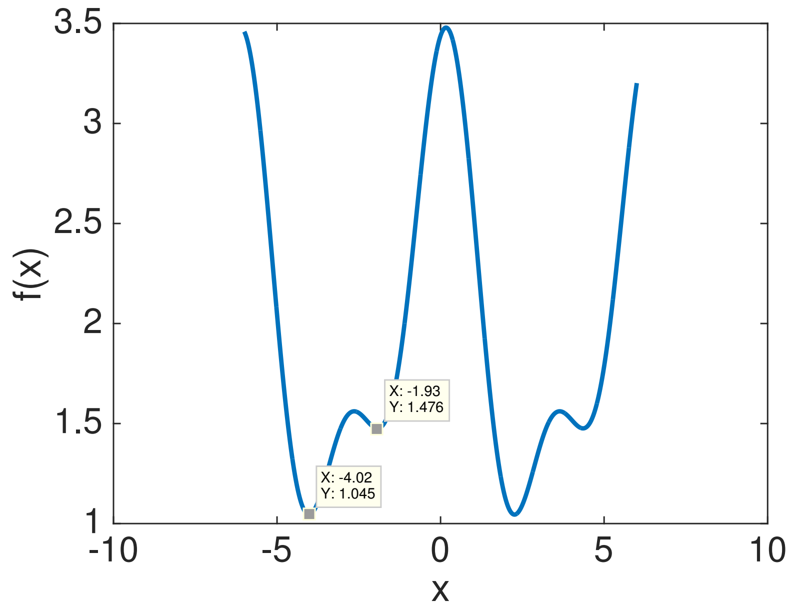


Figure 8: $f(x) = 2 + \cos(x) + 0.5\,\cos(2\,x - 0.5)$ has multiple local and global minimizer. Starting fminunc with $x_0 = -2.61$ results in converging to $x = -1.9277$, however, starting fminunc with $x_0 = -2.65$ results in converging to $x = -4.0221$.

2.1.2 A class of parameter optimization problems with equality constraints

Let the performance measure be differentiable function $F(x, u) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, which is a function of the control vector $u \in \mathbb{R}^m$ and an auxiliary (state) vector $x \in \mathbb{R}^n$. The optimization problem is to determine the control vector $u \in \mathbb{R}^m$ that minimizes $L(x, u)$ and at the same time satisfy the constraint equation $f(x, u) = 0$, where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ is a differential vector function, i.e.,

$$u^\star = \underset{u\in\mathbb{R}^m}{\text{argmin}} F(x, u), \quad \text{s.t.,} \tag{2.1.5a}$$

$$f(x, u) = 0. \tag{2.1.5b}$$

One way of solving this problem is to find x in terms of u from $f(x, u) = 0$. Then substitute this value for x in $F(x, u)$. Then one can solve the constrained optimization using unconstrained approach discussed earlier. For example, consider the optimization problem below

$$\begin{gathered}\text{minimize}\, F(x,u) = x^2+u^2, \quad s.t. \\ x+u+2=0.\end{gathered}$$

To find a minimum point, use the constraint to write $x=-2-u$. Then, the cost becomes

$$F(u) = (-u-2)^2+u^2 = 2u^2+4u+4.$$

Using first order necessary condition

$$\frac{\partial F}{\partial u} = 4u+4=0,$$

we obtain $x^\star = u^\star = 1$.

As can seen from this example, this method works the best for linear f. However, when f is nonlinear it is quite possible that finding x in terms of u can not be done in a tractable manner. Next, we will derive some necessary conditions for the candidate local minimum points by analyzing the changes in cost function in the neighborhood of the candidate points on the feasible set.

The difference between the constrained and unconstrained minimization analysis is that now our search for the minimum point is restricted to the feasible set defined by the constraint equations, i.e., the set $\mathcal{S}_{\text{feas}}$ defined by

$$\mathcal{S}_{\text{feas}} = \{(x,u)\in\mathbb{R}^m\times\mathbb{R}^n | f(x,u)=0\}.$$

For minimum point, we look for $(x^\star,u^\star)$ such that

$$\begin{cases} f(x^\star,u^\star)=0, \\ F(x^\star,u^\star)\leq F(x,u) \quad \forall (x,u)\in\mathcal{S}_{\text{feas}}.\end{cases}$$

First-order necessary condition

Definition 12 (Stationary or critical point for (2.1.5))**.** At a stationary or a critical point $\mathrm{d}F$ is equal to zero in the first-order approximation with respect to increments $\mathrm{d}u$ when $\mathrm{d}f$ is zero. □

- Let us consider a point (x,u) in the feasibility set (also can be phrased as on the constraint manifold) and investigate the variation in the cost for the neighboring points on the constraint manifold.

$$\begin{aligned} f(x+\mathrm{d}x, u+\mathrm{d}u) &\approx f(x,u) + \underbrace{(f_u)^\top \mathrm{d}u + (f_x)^\top \mathrm{d}x}, \\ F(x+\mathrm{d}x, u+\mathrm{d}u) &\approx F(x,u) + \underbrace{(F_u)^\top \mathrm{d}u + (F_x)^\top \mathrm{d}x}_{\mathrm{d}F}, \end{aligned}$$

where

$$f_x = \frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f^1}{\partial x_1} & \frac{\partial f^2}{\partial x_1} & \cdots & \frac{\partial f^n}{\partial x_1} \\ \frac{\partial f^1}{\partial x_2} & \frac{\partial f^2}{\partial x_2} & \cdots & \frac{\partial f^n}{\partial x_2} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f^1}{\partial x_n} & \frac{\partial f^2}{\partial x_n} & \cdots & \frac{\partial f^n}{\partial x_n} \end{bmatrix}, \quad f_u = \frac{\partial f}{\partial u} = \begin{bmatrix} \frac{\partial f^1}{\partial u_1} & \frac{\partial f^2}{\partial u_1} & \cdots & \frac{\partial f^n}{\partial u_1} \\ \frac{\partial f^1}{\partial u_2} & \frac{\partial f^2}{\partial u_2} & \cdots & \frac{\partial f^n}{\partial u_2} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f^1}{\partial u_m} & \frac{\partial f^2}{\partial u_m} & \cdots & \frac{\partial f^n}{\partial u_m} \end{bmatrix},$$

$$F_x = \frac{\partial F}{\partial x} = \begin{bmatrix} \frac{\partial F}{\partial x_1} \\ \frac{\partial F}{\partial x_2} \\ \vdots \\ \frac{\partial F}{\partial x_n} \end{bmatrix}, \quad F_u = \frac{\partial F}{\partial u} = \begin{bmatrix} \frac{\partial F}{\partial u_1} \\ \frac{\partial F}{\partial u_2} \\ \vdots \\ \frac{\partial F}{\partial u_m}, \end{bmatrix}.$$

Notice that $f(x + \mathrm{d}x, u + \mathrm{d}u) = 0$, therefore, for the stationary points we need

$$(f_u)^\top \mathrm{d}u + (f_x)^\top \mathrm{d}x = 0, \tag{2.1.6a}$$
$$(F_u)^\top \mathrm{d}u + (F_x)^\top \mathrm{d}x = 0. \tag{2.1.6b}$$

From (2.1.6a), we obtain

$$\underbrace{\mathrm{d}x}_{\text{depends on } \mathrm{d}u} = -(f_x)^{-\top}(f_u)^\top \underbrace{\mathrm{d}u}_{\text{free to choose in any direction}}. \tag{2.1.7}$$

Substituting for $\mathrm{d}x$ in (2.1.6b), we obtain

$$((F_u)^\top - (F_x)^\top (f_x)^{-\top}(f_u)^\top)\mathrm{d}u = 0,$$

which, because of it has to be value for any value of $\mathrm{d}u$, results in the following necessary and sufficient condition for a point (x, u) be a critical (stationary) point:

$$(F_u)^\top - (F_x)^\top (f_x)^{-\top}(f_u)^\top = 0, \quad \Leftrightarrow F_u - f_u\,(f_x)^{-1}\,F_x = 0.$$

Notice that

$$F_u - f_u\,(f_x)^{-1}\,F_x = \frac{\partial F}{\partial u}\Big|_{\mathrm{d}f}.$$

- An alternative strategy to characterize the first order optimality conditions can be obtained using Lagrange multipliers method. In this method, we adjoin the constraint $f(x, u) = 0$ to the cost using constants (Lagrange multipliers)

$$\lambda = \left[\lambda_1, \cdots, \lambda_n\right]^\top \in \mathbb{R}^n,$$

to obtain

$$H(x, u, \lambda) = F(x, u) + \lambda^\top f(x, u).$$

Notice that the values of H and F along to points on the feasible set are the same.
Given values of x and u that satisfy $f(x,u)=0$ consider differential changes to H due to differential changes in x and u:

$$dH = H_x^\top dx + H_u^\top du$$

Since u is our decision variable, we would like to keep du but we will choose λ such that $H_x=0$:

$$H_x = F_x + \lambda^\top f_x \Rightarrow \lambda^\top = (f_x)^{-1} F_x$$

We are investigating the changes in the cost function for the points on the neighboring points on the constraint manifold so to proceed, we must determine what changes are possible to the cost keeping the equality constraint satisfied.
Changes to (x,u) such that $f(x,u)=0$:

$$df = f_x^\top dx + f_u^\top d_u = 0 \Rightarrow \underbrace{dx}_{dependent} = -f_x^{-\top} f_u^\top \underbrace{du}_{arbitrary}.$$

Now lets look at the cost fucntion

$$dF = F_x^\top dx + F_u^\top du = (-F_x^\top \underbrace{f_x^{-\top} f_u^\top}_{\lambda} + F_u)du = H_u^\top du$$

You can also note that

$$H(x,u,\lambda) = F(x,u) + \lambda^\top f(x,u) \Rightarrow dH(x,u,\lambda) = dF(x,u) + \lambda^\top df(x,u)$$

therefore, for valid variations, i.e., $df(x,u)=0$ we have

$$dF = dH \underbrace{\Rightarrow}_{H_x=0} dF = H_u^\top du$$

So the gradient of cost F with respect to u while keeping the constraint $f(x,u)=0$ is just H_u. We need this gradient to be zero to have a stationary point so that $dF=0$ for $\forall du \neq 0$. Thus the necessary condition for a stationary value of F are

$$\begin{cases} H_x = 0, \\ H_u = 0, \\ H_\lambda = f(x,u) = 0 \end{cases}$$

$2n+m$ unknown $2n+m$ equations!

Constrained optimization: sufficient condition for optimality

To obtain sufficient conditions for a stationary (critical) point $(x, u) \in \mathcal{X} \times \mathcal{U}$ to be a minimum point we look at the second-order order Taylor series expansion of $F(x, u)$ and $f(x, u)$:

$$f^1(x+\mathrm{d}x, u+\mathrm{d}u) \approx f^1(x,u) + \underbrace{(f^1_u)^\top \mathrm{d}u + (f^1_x)^\top \mathrm{d}x + \frac{1}{2}\mathrm{d}x^\top f^1_{xx}\mathrm{d}x + \mathrm{d}x^\top f^1_{xu}\mathrm{d}u + \frac{1}{2}\mathrm{d}u^\top f^1_{uu}\mathrm{d}u}_{\mathrm{d}f^1},$$

$$\vdots$$

$$f^n(x+\mathrm{d}x, u+\mathrm{d}u) \approx f^n(x,u) + \underbrace{(f^n_u)^\top \mathrm{d}u + (f^n_x)^\top \mathrm{d}x + \frac{1}{2}\mathrm{d}x^\top f^n_{xx}\mathrm{d}x + \mathrm{d}x^\top f^n_{xu}\mathrm{d}u + \frac{1}{2}\mathrm{d}u^\top f^n_{uu}\mathrm{d}u}_{\mathrm{d}f^n},$$

$$F(x+\mathrm{d}x, u+\mathrm{d}u) \approx F(x,u) + \underbrace{(F_u)^\top \mathrm{d}u + (F_x)^\top \mathrm{d}x + \frac{1}{2}\mathrm{d}x^\top F_{xx}\mathrm{d}x + \mathrm{d}x^\top F_{xu}\mathrm{d}u + \frac{1}{2}\mathrm{d}u^\top F_{uu}\mathrm{d}u}_{\mathrm{d}F},$$

$$F_{uu} = \frac{\partial}{\partial u}(\frac{\partial F}{\partial u}), \quad F_{xu} = \frac{\partial}{\partial x}(\frac{\partial F}{\partial u}), \quad F_{xx} = \frac{\partial}{\partial x}(\frac{\partial F}{\partial x})$$
$$f^i_{uu} = \frac{\partial}{\partial u}(\frac{\partial f^i}{\partial u}), \quad f^i_{xu} = \frac{\partial}{\partial x}(\frac{\partial f^i}{\partial u}), \quad f^i_{xx} = \frac{\partial}{\partial x}(\frac{\partial f^i}{\partial x}), \quad i = 1, \cdots, n.$$

We are examining the increments in cost in neighboring points of a critical point on the constraint manifold, therefore, $\mathrm{d}f = 0$. Recall $H = F(x, u) + \sum_{i=1}^{n} \lambda_i f^i(x, u)$. We like to obtain the second order condition in terms of Hamiltonian. Notice that

$$\mathrm{d}H = \mathrm{d}F + \sum_{i=1}^{n} \lambda_i \mathrm{d}f^i(x,u) = \begin{bmatrix} H_x^\top & H_u^\top \end{bmatrix} \begin{bmatrix} \mathrm{d}x \\ \mathrm{d}u \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \mathrm{d}x^\top & \mathrm{d}u^\top \end{bmatrix} \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} \mathrm{d}x \\ \mathrm{d}u \end{bmatrix}.$$

We are examining the increments in cost in neighboring points of a critical point on the constraint manifold. From the first-order analysis we got that at the critical point we have $H_x = 0, \quad H_u = 0$ and $\mathrm{d}x = -(f_x)^{-\top}(f_u)^\top \mathrm{d}u$. So we can write

$$\mathrm{d}F = \frac{1}{2}\mathrm{d}u^\top \begin{bmatrix} -f_u(f_x)^{-1} & I \end{bmatrix} \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} -(f_x)^{-\top}(f_u)^\top \\ I \end{bmatrix} \mathrm{d}u$$

Thus, a sufficient condition for a local minimum are

$$H_x = 0, \quad H_u = 0 \tag{2.1.8}$$
$$\begin{bmatrix} -f_u(f_x)^{-1} & I \end{bmatrix} \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} -(f_x)^{-\top}(f_u)^\top \\ I \end{bmatrix} > 0 \tag{2.1.9}$$

Clearly, a necessary condition for local minimum is

$$H_x = 0, \quad H_u = 0 \tag{2.1.10}$$
$$\begin{bmatrix} -f_u(f_x)^{-1} & I \end{bmatrix} \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} -(f_x)^{-\top}(f_u)^\top \\ I \end{bmatrix} \geq 0 \tag{2.1.11}$$

Notice that

$$H_{xx} = F_{xx} + \sum_{i=1}^{n} \lambda_i \mathrm{d} f^i_{xx}, \quad H_{xx} = F_{xx} + \sum_{i=1}^{n} \lambda_i \mathrm{d} f^i_{xu}, \quad H_{uu} = F_{uu} + \sum_{i=1}^{n} \lambda_i \mathrm{d} f^i_{uu}.$$

Numerical example: find the point nearest the origin on the line

$$x + 2y + 3z = 10, \quad x - y + 2z = 1,$$

where x, y, z are rectangular coordinates.
The optimization problem here is

$$\begin{aligned} \text{minimize } 0.5(x^2 + y^2 + z^2), \quad &s.t., \\ x + 2y + 3z - 10 = 0,& \\ x - y + 2z - 1 = 0.& \end{aligned}$$

We first write the Hamiltonian:

$$H = 0.5(x^2 + y^2 + z^2) + \lambda_1(x + 2y + 3z - 10) + \lambda_2(x - y + 2z - 1)$$

The first order condition for gives the following equations for critical point:

$$\begin{aligned} \frac{\partial H}{\partial x} &= x + \lambda_1 + \lambda_2 = 0, \\ \frac{\partial H}{\partial y} &= y + 2\lambda_1 - \lambda_2 = 0, \\ \frac{\partial H}{\partial z} &= z + 3\lambda_1 + 2\lambda_2 = 0, \\ \frac{\partial H}{\partial \lambda_1} &= x + 2y + 3z - 10 = 0, \\ \frac{\partial H}{\partial \lambda_2} &= x - y + 2z - 1 = 0. \end{aligned}$$

You can use Matlab 'linsolve' command to solve the set of linear algebra equations above. For this problem, the solution, the critical point, is $x = 0.3220$, $y = 2.4746$, $z = 1.5763$, $\lambda_1 = -0.9322$ and $\lambda_2 = 0.6102$. To check if this critical point is minimum point, we use the sufficient conditions obtained above with treating x as decision variable u and (y, z) as states.

$$f_x = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad f_{(y,z)} = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}, \quad H_{(y,z),(y,z)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad H_{(y,z),x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad H_{xx} = 1.$$

Then,

$$\begin{bmatrix} -f_x (f_{(y,z)})^{-1} & I \end{bmatrix} \begin{bmatrix} H_{(y,z),(y,z)} & H_{(y,z),x} \\ H_{x,(y,z)} & H_{xx} \end{bmatrix} \begin{bmatrix} -(f_{(y,z)})^{-\top} (f_x)^\top \\ I \end{bmatrix} =$$

$$\begin{bmatrix} 0.1429 & -0.4286 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.1429 \\ -0.4286 \\ 1 \end{bmatrix} = 1.2041 > 0.$$

Therefore, the critical point obtained earlier is the minimizer of our cost function over the feasible set.

Constrained optimization: numerical solution by a 1^{st}-order gradient method

1. Select initial u
2. Determine numerical value of x from $f(x,u)=0$ for the given u
3. Determine $\lambda = -(\frac{\partial f}{\partial x})^{-1}\frac{\partial F}{\partial x} = -(f_x)^{-1}\,F_x$.
4. Determine $H_u = \frac{\partial F}{\partial u} + \frac{\partial f}{\partial u}\lambda = F_u + f_u\lambda$ (this ingeneral, will not be zero)
5. Interpreting H_u as a gradient vector, change the estimates of u by the amount $\Delta u = -\alpha H_u$, where $\alpha > 0$ o a positive scalar constant. The predicted change in the cost is $\Delta F = -\alpha H_u^\top H_u$.
6. Go to step 2 and use the new u to repeat the steps until $\Delta F = -\alpha H_u^\top H_u$ in the threshold you set.

Matlab numerical solver for constrained optimization: fmincon and quadprog

Sample problem:

$$\text{minimize } F(x,u) = x^4 + u^2, \ .s.t., \\ x^2 + u^2 - 2 = 0$$

Code for fmincon: (see http://www.mathworks.com/help/optim/ug/fmincon.html for further details)

- a function to list equality constraints

```
function [c,ceq] = EqualFun(x)

ceq = x(1)^2 + x(2)^2 -2;

c = [];
```

- the main code

```
fun = @(x)(x(1)^4+x(2)^2);

nonlcon = @EqualFun;

A = [];

b = [];

Aeq = [];
```

```
beq = [];

lb = [];

ub = [];

x0 = [0,0];

[x,fval,exitflag,output] = fmincon(fun,x0,A,b,Aeq,beq,lb,ub,nonlcon)
```

If your cost function is quadratic and your constraints are linear you can us Matlab 'quadprog' command to solve your constrained optimization problem. For information about this solver see http://www.mathworks.com/help/optim/ug/quadprog.html .

Note 3

The treatment corresponds to selected parts from Chapter 2 of [3] and Chapter 2 [5]. The presentation is at times informal. For rigorous treatments, students should consult the aforementioned references and the other listed texts in the class syllabus.

3.1 Optimal control of multi-stage systems over finite horizon

We are now ready to extend the methods of Note 2 to the optimization of a performance index associated with a system developing dynamically through time over finite horizon (see Fig. 5). We state objective as to choose sequence of control inputs $\{u(0), u(1), \cdots, u(N-1)\}$ that minimizes (or maximizes) a performance index of the form

$$\begin{aligned} J(u) = \phi(x_N) + \sum\nolimits_{k=0}^{N-1} L^k(x_k, u_k), \quad \text{s.t.}, \\ x_{k+1} = f^k(x_k, u_k), \quad k = 0, 1, \cdots, N-1, \\ x(0) = x_0. \end{aligned} \tag{3.1.1}$$

We start over study by using Lagrange multipliers $\lambda_1, \cdots, \lambda_N$ to adjoin the system state equations to the cost function, i.e.,

$$\bar{J}(u) = \phi(x(N)) + \sum\nolimits_{k=0}^{N-1} (L^k(x_k, u_k) + \lambda_{k+1}^\top (f^k(x_k, u_k) - x_{k+1}),$$

We define

$$H^k = L^i(x(i), u(i)) + \lambda_{k+1}^\top f^k(x_k, u_k), \quad k = 0, \cdots, N-1,$$

to obtain

$$\bar{J}(u) = \; (\phi(x_N) - \lambda_N^\top x_N) + \sum\nolimits_{k=0}^{N-1} (H^k(x_k, u_k, \lambda_{k+1}) - \lambda_k^\top x_k) + H^0(x_0, u_0, \lambda_1),$$

Note 4

Dynamic Programming

The treatment corresponds to selected parts from Chapter 3 of [1]. The presentation is at times informal. For rigorous treatments, students should consult the aforementioned references and the other listed texts in the class syllabus.

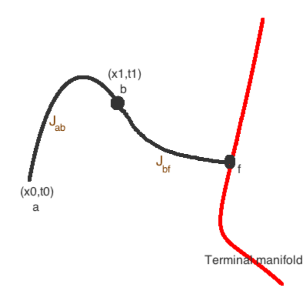
4.1 Introduction

This chapter reviews the *principle of optimality* and use of this principle in design of *dynamic programing*. Dynamic Programing (DP) is a numerical solution procedure for solving multi-stage decision making problems. Usually creativity is required before we can recognize that a particular problem can be cast effectively as a dynamic program; and often subtle insights are necessary to restructure the formulation so that it can be solved effectively. In this chapter, after introducing the dynamic programing, we will show how this optimal decision making tool can be used to solve discrete-time optimal control problems. For most problems, dynamic programming leads to sequential numerical decision making procedures, often quite challenging. However, there are a few cases that dynamic programming offers a closed-form analytical solution; one of those cases is design of optimal discrete LQR solution.

4.2 Principle of optimality

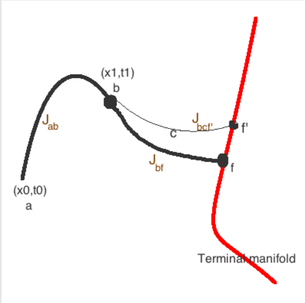
Definition: A problem is said to satisfy the Principle of Optimality if the sub-solutions of an optimal solution of the problem are themselves optimal solutions for their subproblems.

Let $a-b-f$ be the optimal path to go from point a to the terminal manifold shown in the figure below. The first decision made at a (point (x_0, t_0)) results in segment $a-b$ with cost J_{ab} and the remaining decision yields segment $b-f$ (from b, point $((x_1, t_1))$) with cost of J_{bf} to arrive at the terminal manifold. The minimum cost J_{af} from $a-f$ is



$$J^\star_{af} = J_{ab} + J_{ab}.$$

Assertion If $a-b-f$ is the optimal path from a to f then $b-f$ is the optimal path from b to f.
Proof: Proof by contradiction



Principle of optimality (due to Bellman)

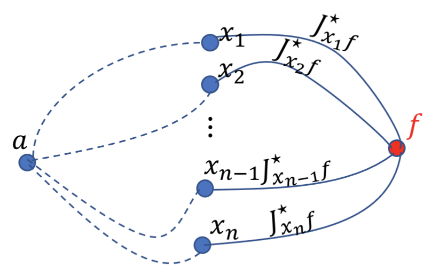
- An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.
- All points on an optimal path are possible initial points for that path
- Suppose the optimal solution for a problem passes through some intermediate point (x_1, t_1) then the optimal solution to the same problem starting at (x_1, t_1) must be the continuation of the same path.

4.3 Dynamic programming

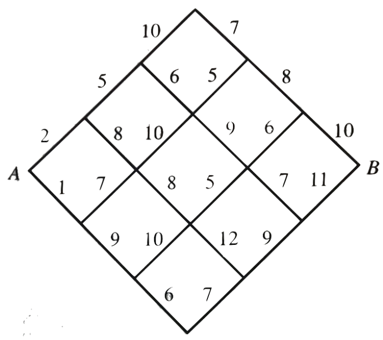
We begin by providing a general insight into the dynamic programming approach by treating a set of simple examples in some detail. These examples show the basis of dynamic programming and use of principle of optimality.

- In general, if there are numerous options at location a that next leads to locations $x_1, \cdots, x_n$ choose the action that leads to

  $$J^\star_{af} = \min_{x_i} \big\{[J_{ax_1} + J^\star_{x_1 f}], [J_{ax_2} + J^\star_{x_2 f}], \cdots, [J_{ax_n} + J^\star_{x_n f}]\big\}$$

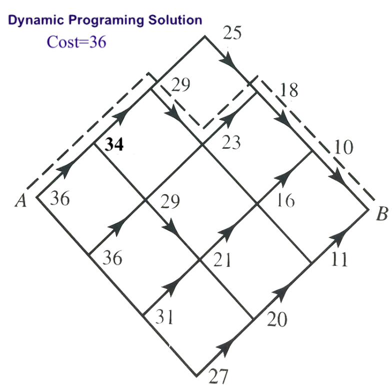


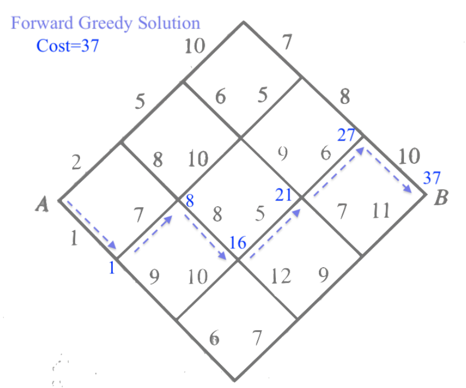
- **Problem** (Bryson): find the path from A to B traveling only to the right, such that the sum of the numbers on the segments along this path is a minimum.
  - minimum time path from A to B: if you think of numbers as time to travel
  - control decision is: up-right or down-right (only two possible value at each node
  - there are 20 possible paths from A to B (traveling only to the right)



Solution approaches

1. There are 20 possible paths: evaluate each and compute the travel time (pretty tedious approach)
2. Start at B and work backwards, invoking the principle of optimality along the way.

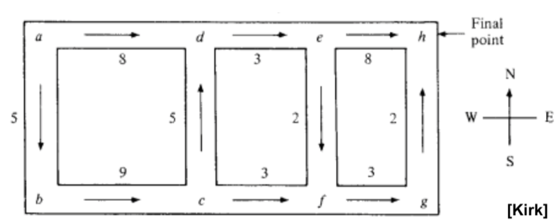




- For dynamic programing (DP) we need to find 15 numbers to solve this problem rather than evaluate the travel time for 20 paths
- Modest difference here, but scales up for larger problems.
- Let $n =$ number of segments on side (3 here) then:
  * Number of routes scales as $\sim (2n)!/(n!)^2$
  * Number DP computations scales as $\sim (n+1)^2 - 1$

Segments on a side	3	4	5	6	7	n
Number of the routes	20	70	252	724	2632	$(2n)!/(n!)^2$
DP computations	15	24	35	48	63	$(n+1)^2-1$

- **Problem**: Minimize cost to travel from c to h moving only along the direction of arrows.



[Kirk]

- g to h: goes directly to h, i.e., $J^\star_{gh} = 2$
- e to h: a possible path goes through f, we need to compute the cost of going from f to h first.
- f to h: $J^\star_{fh} = J_{fg} + J^\star_{gh} = 3 + 2 = 5$
- e to h: $J^\star_{eh} = \min\{J_{eh}, J_{efh}\} = \min\{J_{eh}, [J_{ef} + J^\star_{fh}]\} = \min\{8, 2+5\} = 7,\quad e \to f \to g \to h$
- d to h: $J^\star_{dh} = J_{de} + J^\star_{eh} = 3 + 7 = 10$
- c to h: $J^\star_{ch} = \min\{J_{cdh}, J_{cfh}\} = \min\{[J_{cd} + J^\star_{dh}], [J_{cf} + J^\star_{fh}]\} = \min\{[5+10], [3+5]\} = 8$

Optimal path: $c \to f \to g \to h$

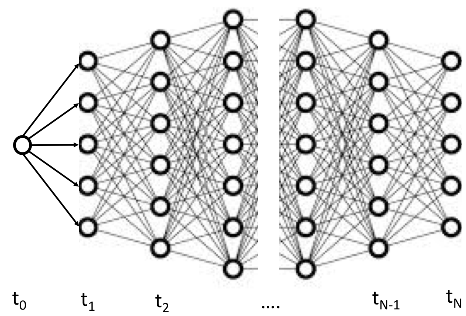
Greedy vs. Dynamic Programming:

- Both techniques are optimization techniques, and both build solutions from a collection of choices of individual elements.
- The greedy method computes its solution by making its choices in a serial forward fashion, never looking back or revising previous choices.
- Dynamic programming computes its solution bottom up by synthesizing them from smaller subsolutions, and by trying many possibilities and choices before it arrives at the optimal set of choices.
- There is no a priori litmus test by which one can tell if the Greedy method will lead to an optimal solution.
- By contrast, there is a litmus test for Dynamic Programming, called The Principle of Optimality

4.4 Dynamic Programing: optimal control

Roadmap to use DP in optimal control

- Grid the time/state and find the necessary control
- Grid the time/state and quantize control inputs
- Discrete-time problem: discrete time LQR



A discrete time/quantized space grid with the linkages showing the possible transition in state/time grid through the control commands. It is hard to evaluate all options moving forward through the grid, but we can work backwards and use the principle of optimality to reduce this load.

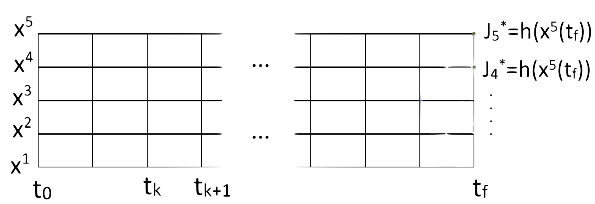
Consider

$$\text{minimize } J = h(x(t_f)) + \int_{t_0}^{t_f} g(x(t), u(t), t)\mathrm{d}t, \quad \text{s.t.}$$
$$\dot{x} = a(x, u, t),$$
$$x(t_0) = x_0 = \text{fixed},$$
$$t_f = \text{fixed}$$

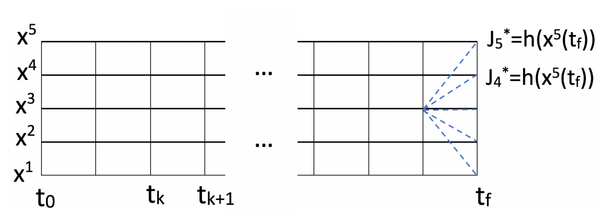
We will discuss including constraints on $x(t)$ and $u(t)$

DP solution

1. develop a grid over space/time
2. evaluate the final cost at possible final states $x^i(t_f)$: $J_i^\star = h(x^i(t_f))\ \forall i$



3. back up 1 step in time and consider all possible ways of completing the problem



To obtain the cost of a control action, we approximate the integral in the cost.

- let $u^{ij}(t_k)$ be the control action that takes the system from $x^i(t_k)$ to $x^j(t_{k+1})$ at time $t_k + \Delta t$. Then the approximate cost of going from $x^i(t_k)$ to $x^j(t_{k+1})$:

$$\int_{t_k}^{t_{k+1}} g(x(t), u(t), t)\mathrm{d}t \approx g(x^i(t_k), u^{ij}(t_k), t_k)\Delta t.$$

- $u^{ij}(t_k)$ is computed from the system dynamics:

$$\dot{x} = a(x, u, t) \Rightarrow \frac{x(t_{k+1}) - x(t_k)}{\Delta t} = a(x(t_k), u(t_k), t_k) \Rightarrow$$

$$x^j(t_{k+1}) = x^i(t_k) + a(x^i(t_k), u^{ij}(t_k), t_k)\Delta t \Rightarrow u^{ij}(t_k)$$

  Note: If the system is control affine $\dot{x} = f(x, t) + g(x, t)u$, the control $u^{ij}(t_k)$ can be computed from $u^{ij}(t_k) = g(x^i_k, t_k)^{-1}(\frac{x^j(t_{k+1}) - x^i(t_k)}{\Delta t} - f(x^i_k, t_k))$

- So far for any combination of x^i_k and x^j_{k+1} on the state/time grid we can evaluate the incremental cost $\Delta J(x^i_k, x^j_{k+1})$ of making the state transition.
- Assuming you know already the optimal path from each new terminal point x^j_{k+1}, the optimal path from x^i_k is established from

$$J^\star(x^i_k, t_k) = \min_{x^j_{k+1}} \Big[\Delta J(x^i_k, x^j_{k+1}) + J^\star(x^j_{k+1})\Big]$$

- So far for any combination of x^i_k and x^j_{k+1} on the state/time grid we can evaluate the incremental cost $\Delta J(x^i_k, x^j_{k+1})$ of making the state transition.
- Assuming you know already the optimal path from each new terminal point x^j_{k+1}, the optimal path from x^i_k is established from

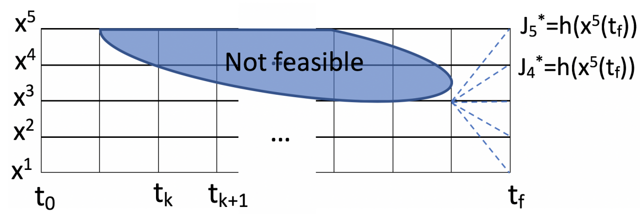
$$J^\star(x^i_k, t_k) = \min_{x^j_{k+1}} \Big[\Delta J(x^i_k, x^j_{k+1}) + J^\star(x^j_{k+1})\Big]$$

- Then for each x^i_k the output is
  - Best x^i_{k+1} to pick that gives the lowest cost
  - Control input required to achieve this best cost

4. then work backwards on time until you reach x_0, when only one value of x is allowed because of the given initial condition

Couple of points about the process that is explained above

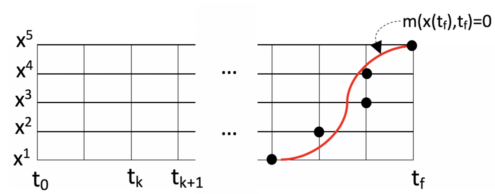
- with constraints on the state, certain values of $x(t)$ might not be allowed at certain time t.



- with bounds on the control, certain state transitions might not be allowed from one time step to another
- the process extends to higher dimensions. Just have to define a grid of points in x and t. See Kirk's book for more details.
- Extension of the method discussed earlier to the case of free end time with some additional constraint on the final state $m(x(t_f), t_f) = 0$, i.e.,

$$\begin{aligned}&\text{minimize } J = h(x(t_f)) + \int_{t_0}^{t_f} g(x(t), u(t), t)\mathrm{d}t, \quad \text{s.t.}\\ &\dot{x} = a(x, u, t),\\ &x(t_0) = x_0 = \text{fixed},\\ &m(x(t_f), t_f) = 0 \quad t_f = \text{free}\end{aligned}$$

  – find a group of points on the state/time grid that (approximately) satisfy the terminal constrain

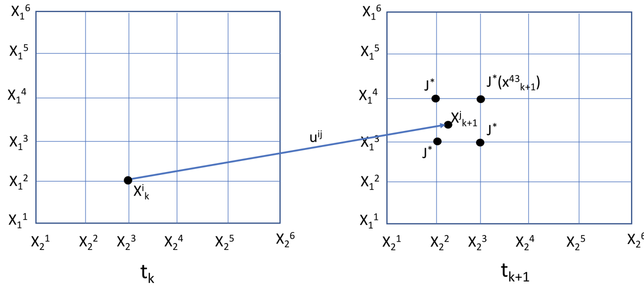


  – evaluate cost for each point and work backward from there

- The previous formulation picked x's and used the sate equation to determine the control needed to transition between the quantized states across time.
  - For more general case problems, it might be better to pick the u's and use those to determine the propagated x's

$$\begin{aligned}J^\star(x_k^i, t_k) =&\min_{u_k^{ij}}\Big[\Delta J(x_k^i, u_k^{ij}) + J^\star(x_{k+1}^j, t_{k+1})\Big] =\\ &\min_{u_k^{ij}}\Big[g(x_k^i, u_k^{ij}, t_k)\Delta t + J^\star(x_{k+1}^j, t_{k+1})\Big]\end{aligned}$$

  - To this end, the control inputs should be quantized as well.
  - Then, it is likely that terminal points from one time step to the next will not lie on the state discrete points: must interpolate the cost to go between between them



Example 1 (Kirk) Consider

$$\text{Minimize } J = x^2(2) + 2\int_0^2 u^2(t)dt \text{ subject to}$$
$$\begin{aligned}&\dot{x} = u,\\ &0 \leq x(t) \leq 1.5,\\ &-1 \leq u(t) \leq 1\end{aligned}$$

- Quantize the state within the allowable values and time within the range $t \in [0, 2]$ using $N = 2$ and $\Delta t = 1$, i.e., $k = 0, 1, 2$

Solution:

1. Use Euler integration approximation), a very common discretization process, which works well for small , to discretize the system

$$\dot{x} \approx \frac{x(t + \Delta t) - x(t)}{\Delta t} = u(t) \Rightarrow x_{k+1} = x_k + \Delta t u_k$$

2. Use approximate calculation to discretize the cost

$$J = x^2(N) + 2\sum_{k=0}^{N-1} u_k^2(t)\Delta t$$

3. Given that $0 \leq x \leq 1.5$, take x quantized into four possible values $x_k \in \{0, 0.5, 1.0, 1.5\}$

4. Compute cost associate with all possible terminal states ($k = 2$):

x_2^j	$J_2^\star = h(x_2^j) = (x_2^j)^2$
0	0
0.5	0.25
1	1
1.5	2.25

5. Back track to $k = 1$. Given any x_1^i, determine the control to go to x_2^j

$u(1)$	$x_2^j = x_1^i + u(1)$			
x_1^i	**0**	**0.5**	**1**	**1.5**
0	0	0.5	1	~~1.5~~
0.5	-0.5	0	0.5	1
1	-1	-0.5	0	0.5
1.5	~~-1.5~~	-1	-0.5	0

Compute incremental cost →

ΔJ_{12}^{ij}	x_2^j			
x_1^i	**0**	**0.5**	**1**	**1.5**
0	0	0.5	2	XX
0.5	0.5	0	0.5	2
1	2	0.5	0	0.5
1.5	XX	2	0.5	0

$\Delta J_{12}^{ij} = 2u^{ij}(1)^2\Delta t$

Compute cost at k= 1 →

J_1	x_2^j			
x_1^i	**0**	**0.5**	**1**	**1.5**
0	**0**	0.75	3	XX
0.5	0.5	**0.25**	1.5	4.25
1	2	**0.75**	1	2.75
1.5	XX	2.25	**1.5**	2.25

$J_1 = \Delta J_{12}^{ij} + J_2^\star(x_2^j)$

$x_1^i \rightarrow x_2^j$
$0 \rightarrow 0$
$0.5 \rightarrow 0.5$
$1 \rightarrow 0.5$
$1.5 \rightarrow 1$

Take min across each row to determine best action at each possible state: $x_1 \rightarrow J_1^\star(x_1^j)$

6. Compute cost associate with all possible terminal states ($k = 1$):

x_1^j	$J_1^\star$
0	0
0.5	0.25
1	0.75
1.5	1.5

7. Back track to $k = 0$. Given any x_0^i, determine the control to go to x_1^j

$u(0)$	$x_1^j = x_0^i + u(0)$			
x_0^i	**0**	**0.5**	**1**	**1.5**
0	0	0.5	1	~~1.5~~
0.5	-0.5	0	0.5	1
1	-1	-0.5	0	0.5
1.5	~~-1.5~~	-1	-0.5	0

Compute incremental cost →

ΔJ_{01}^{ij}	x_1^j			
x_0^i	**0**	**0.5**	**1**	**1.5**
0	0	0.5	2	XX
0.5	0.5	0	0.5	2
1	2	0.5	0	0.5
1.5	XX	2	0.5	0

$\Delta J_{01}^{ij} = 2u^{ij}(0)^2\Delta t$

Compute cost at $k = 0$ →

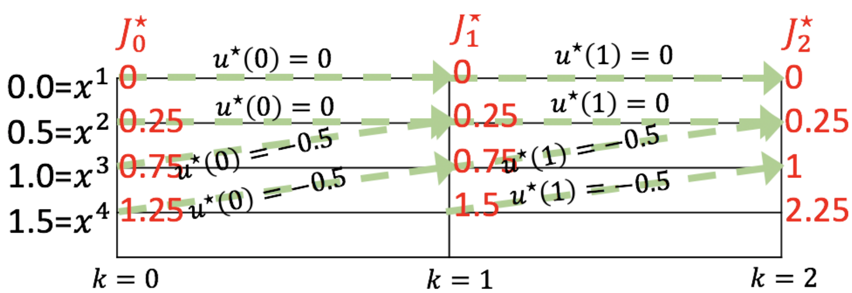
J_0	x_1^j			
x_0^i	**0**	**0.5**	**1**	**1.5**
0	**0**	0.75	2.75	XX
0.5	0.5	**0.25**	1.25	3.5
1	2	**0.75**	0.75	2
1.5	XX	2.25	**1.25**	1.5

$J_0 = \Delta J_{01}^{ij} + J_1^\star(x_1^j)$

$x_0^i \rightarrow x_1^j$
$0 \rightarrow 0$
$0.5 \rightarrow 0.5$
$1 \rightarrow 0.5$
$1.5 \rightarrow 1$

Take min across each row to determine best action at each possible state: $x_0 \rightarrow J_0^\star(x_0^j)$

8. Final DP result



9. Example: starting at $x(0) = 1$, the optimal control is $u^\star(0) = -0.5$ and $u^\star(1) = 0$

In the preceding control example all of the trial control values drive the state of the system either to a computational "grid" point or to a value outside of the allowable range. Had the numerical values not been carefully selected, this happy situation would not have been obtained and interpolation would have been required.

Example 2 (Quantized Control) Consider

$$\text{Minimize } J = (x(2) - 0.5)^2 + 2\int_0^2 u^2(t)dt \text{ subject to}$$

$$\begin{aligned}&\dot{x} = 0.5x + 0.5u,\\ &-0.5 \leq x(t) \leq 1,\\ &-1 \leq u(t) \leq 1\end{aligned}$$

Quantize the state within the allowable values and time within the range $t \in [0, 2]$ using $N = 2$ and $\Delta t = 1$, i.e., $k = 0, 1, 2$
Quantize the control input within the allowable values $u \in \{-1, 0, 1\}$
Solution:

1. Use Euler integration approximation), a very common discretization process, which works well for small $\Delta t = 1$, to discretize the system

$$\dot{x} \approx \frac{x(t + \Delta t) - x(t)}{\Delta t} = 0.5\,x(t) + 0.5\,u(t) \Rightarrow x_{k+1} = 1.5\,x_k + 0.5\,u_k$$

2. Use approximate calculation to discretize the cost

$$J = (x(N) - 0.5)^2 + 2\sum_{k=0}^{N-1} u_k^2(t)\Delta t$$

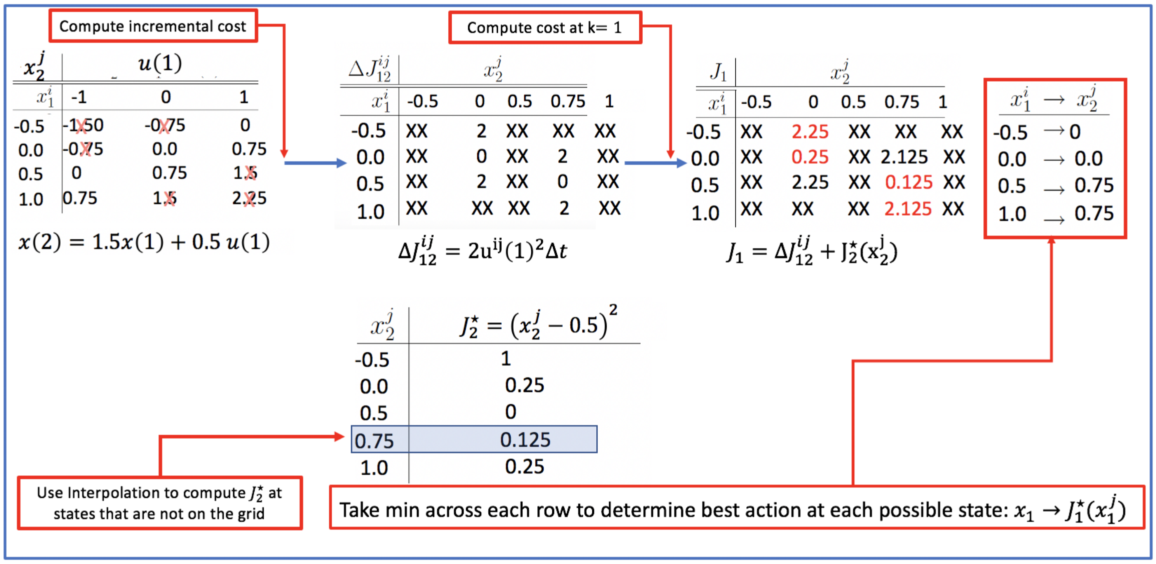
3. Given that $-0.5 \leq u \leq 1$, take x quantized into four possible values $x_k \in \{-0.5, 0.5, 1.0\}$

4. Given that $-1 \leq u \leq 1$, take u quantized into three possible values $u_k \in \{-1, 0, 1\}$

5. Compute cost associate with all possible terminal states ($k = 2$):

x_2^j	$J_2^\star = (x_2^j - 0.5)^2$
-0.5	1
0.0	0.25
0.5	0
1.0	0.25

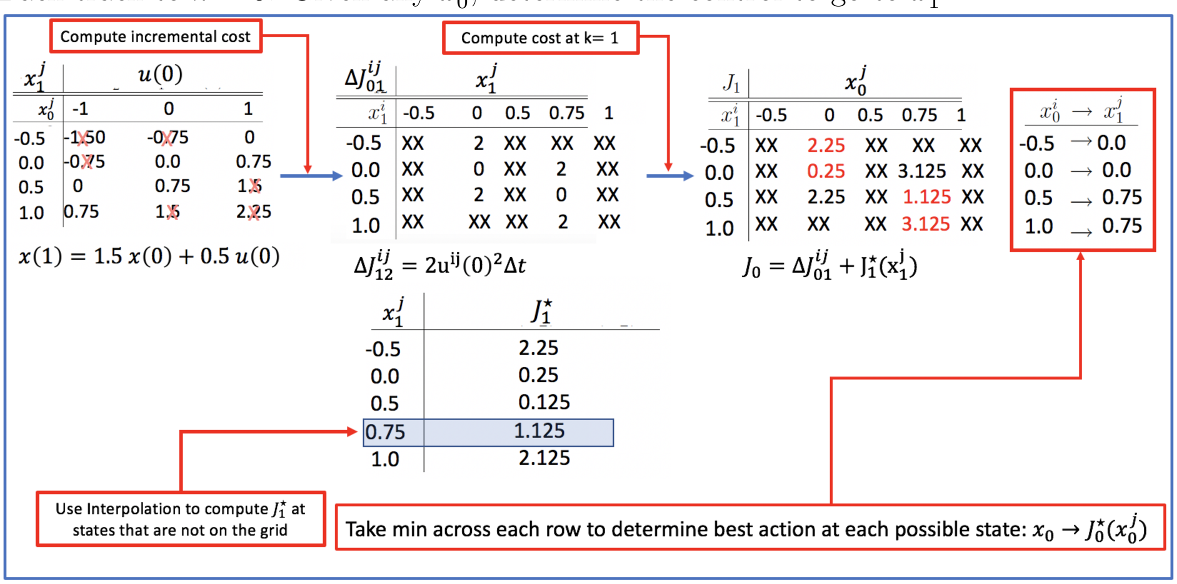
6. Back track to $k = 1$. Given any x_1^i, determine the control to go to x_2^j



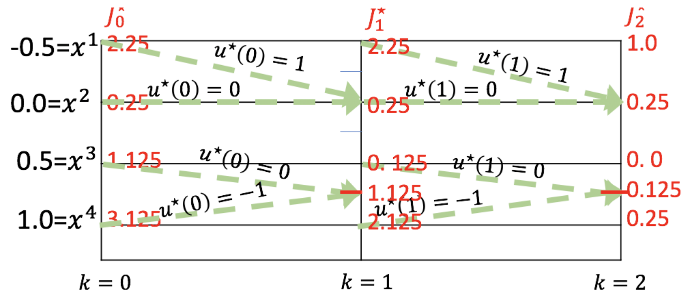
7. Compute cost associate with all possible terminal states ($k = 1$):

x_1^j	$J_1^\star$
-0.5	2.25
0.0	0.25
0.5	0.125
1.0	2.125

8. Back track to $k = 0$. Given any x_0^i, determine the control to go to x_1^j



9. Final DP result



Note: Interpolation may also be required when one is using stored data to calculate an optimal control sequence. For example, if the optimal control applied at some value of $x(0)$ drives the system to a state value $x(1)$ that is halfway between two points where the optimal controls are -1 and 0, then by linear interpolation the optimal control is -0.5.

10. Example: Given initial condition $x(0) = 0.5$, optimal control is $u^\star(0) = 0$. Because $x(1) = 0.75$ is not on the state grid, to find the optimal control $u^\star(1)$ we should use interpolation. Note $(x(1) = 0.5, u^\star(1) = 0)$ and $(x(1) = 1, u^\star(1) = -1)$. Then, after interpolation we obtain $u^\star(1) = -0.5$.

4.5 Closing Remark

Dynamic programing: curse of dimensionality

- Main concern with dynamic programming is how badly it scales
- Given m quantized states with dimension n and N points in time, the number of calculations for dynamic programing is Nm^n

"Curse of Dimensionality"

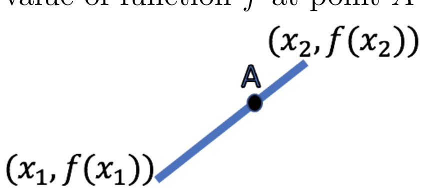
see Dynamic Programing by R. Bellman (1957),

Interpolation

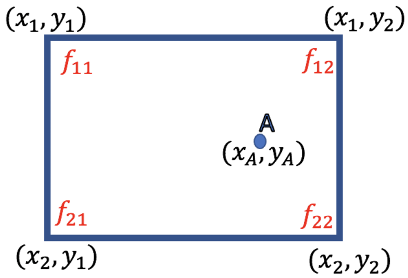
- Linear interpolation: Consider $f(x)$. Suppose you know $y_1 f_1 = f(x_1)$ and $y_2 f_2 = f(x_2)$. We want to use linear interpolation to find the value of function f at point A where $x_1 \leq x_A \leq x_2$:

$$f(x) = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) = \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2)$$

value of function f at point A where $x_1 \leq x_A \leq x_2$.



- Bilinear interpolation: Consider $f(x,y)$. Suppose you know $f_1 1 = f(x_1,y_1)$, $f_1 2 = f(x_1,y_2)$, $f_{21} = f(x_2,y_1)$, and $f_2 2 = f(x_2,y_2)$. We want to use bilinear interpolation to find the value of function f at point A where $x_1 \le x_A \le x_2$, $y_1 \le y_A \le y_2$.



We do first linear interpolation in x direction:

$$f(x,y_1) \approx \frac{x_2 - x}{x_2 - x_1} f_{11} + \frac{x - x_1}{x_2 - x_1} f_{21}$$

$$f(x,y_2) \approx \frac{x_2 - x}{x_2 - x_1} f_{12} + \frac{x - x_1}{x_2 - x_1} f_{22}$$

We proceed by interpolating in the $y-$direction to obtain the desired estimate:

$$\begin{aligned} f(x,y) \approx & \frac{y_2 - y_1}{y_2 - y_1} f(x,y_1) + \frac{y - y_1}{y_2 - y_1} f(x,y_2) + \dots \\ &= \frac{1}{(x_2 - x_1)(y_2 - y_1)} \begin{bmatrix} x_2 - x & x - x_1 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} y_2 - y \\ y - y_1 \end{bmatrix} \end{aligned}$$

Note 5

Calculus of variation

The treatment corresponds to selected parts from Chapters 4 [1]. The presentation is at times informal. For rigorous treatments, students should consult the aforementioned references and the other listed texts in the class syllabus.

5.1 Calculus of Variation and its relevance to optimal control

Another class of optimization problems we will study is the following problem

$$
\begin{aligned}
&u^\star(t)\Big|_{t\in[t_0,t_f]} = \underset{u(t)\in\mathcal{U}}{\operatorname{argmin}}(J = h(x(t_f),t_f) + \int_{t_0}^{t_f} g(x(t),u(t),t)), \;\; s.t. \\
&\dot{x}(t) = f(x(t),u(t),t), \quad t\in[t_0,t_f], \\
&x(t_0),\; t_0 \text{ is given}, \\
&m(x(t_f),t_f) = 0 \leftarrow \text{ various constraints on final state}, \\
&\\
&x(t):\mathbb{R}\to\mathbb{R}^n, \quad u(t):\mathbb{R}\to\mathbb{R}^m, \quad f:\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}\to\mathbb{R}^n.
\end{aligned}
$$

Notice her that the cost function is a function of $x(t)$ and $u(t)$ which are both themselves functions of time. Therefore $J(x(t),u(t))$ is a functional.

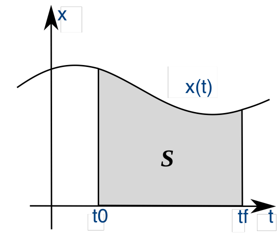
Definition 13 (function). *A function f is a rule of correspondence that assigns to each element q in a certain set $\mathcal{D}$ (domain) a unique element in a set $\mathcal{R}$ (range or image).* □

Definition 14 (functional). *A functional J is a rule of correspondence that assigns to each function x in a certain class Ω a unique real number. Ω is called the domain of functional, and the set of real numbers associated with the functions Ω is called the range of the functional.* □

Example Suppose that x is a continuous-time function of t defined in the interval $t_0, t_f]$ and

$$J(x) = \int_{t_0}^{t_f} x(t)\mathrm{d}t;$$

the real number assigned by functional J is the area under the $x(t)$ curve.



In optimal control problem in continuous-time, the objective is to determine a function that minimizes a specific functional, the performance measure. In discrete-time systems we were able to go from a problem involving variables dynamically changing through time to static parameter optimization problem through stacking the variables on top of one another to arrive at decision variables represented as vectors with finite dimensions. This approach cannot be used with the continuous-time systems, as we will end up with a decision vector of infinite dimension. The branch of mathematics that is extremely useful in solving an optimization problem of the form for continuous-time systems is the Calculus of Variation.

Calculus of Variation

- field of mathematical analysis that deals with maximization/minimization of functionals
- functionals are defined as integrals involving functions and their derivatives
- interest is in extremal functions that make the functional attain
  - maximum
  - minimum
  - or stationary functions (those where the rate of change of the functional is zero)

5.2 Preliminaries for calculus of variation

In parameter optimization, in order to investigate if a point is a local minimum/maximum/stationary point, we looked at the variation of the function for points close to that point. Calculus of variation employs the same concept to identify extremals of a functional. Before proceeding further, we need first to define the concepts of *closeness* and *increments of a functional*.

For functions, recall that we used the concept of norms to measure closeness. Norm in n-dimensional Euclidean space is defined as a rule of correspondence which assigns to each point q a real number. Any function of q represented by $\|q\|$ is a norm if it satisfies the properties below,

1. $\|q\| \geq 0$ and $\|q\| = 0$ iff $q = 0$
2. $\|\alpha q\| = |\alpha|\|q\|$ for all $\alpha \in \mathbb{R}$
3. $\|q^1 + q^2\| \leq \|q^1\| + \|q^2\|$

To points q^1 and q^2 are close together $\Leftrightarrow$ $\|q^1 - q^2\|$ is small.

For functions we also use concept of norm as a measure of size of the functions. Norm of a function is a rule of correspondence which assigns to each function $x \in \Omega$, defined for $t \in [t_0, t_f]$, a real number. Any functional of x represented by $\|qx\|$ is a norm if it satisfies the properties below,

1. $\|x\| \geq 0$ and $\|x\| = 0$ iff $x(t) = 0$ for all $t \in [t_0, t_f]$
2. $\|\alpha x\| = |\alpha| \, \|x\|$ for all $\alpha \in \mathbb{R}$
3. $\|x^1 + x^2\| \leq \|x^1\| + \|x^2\|$

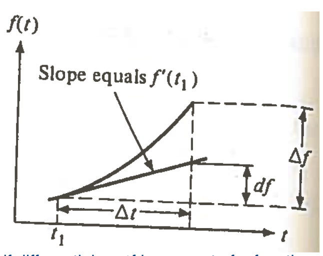
For two functions $x_1 \in \Omega$ and $x_2 \in \Omega$, intuitively speaking norm of the difference of them $\|x_1 - x_2\|$ should be

- zero if the functions are identical.
- small, if the functions are "close".
- large if the functions are "far apart".

Following are two examples of norm of functions

- $\|x\|_2 = (\int_{t_0}^{t_f} x^\top(t) x(t) \mathrm{d}t)^{1/2}$.
- $\|x\| = \max\limits_{t_0 \leq t \leq t_f} (|x(t)|)$, (scalar x)

In the following we introduce increment of a functional



df differential vs.Δf increment of a function

(photo courtesy of [1])

Increment of a function f: for two point q, $q + \Delta q$ in the domain $\mathcal{D}$ of the function, the increment of f is

$$\Delta f = f(q + \Delta q) - f(q).$$

Recall that the *differential* $\mathrm{d}f$ of a function is the first-order (in Δq) approximation of Δf:

$$\Delta f = \mathrm{d}f + O((\Delta q)^2)$$

$$\mathrm{d}f = \frac{\partial f}{\partial q} \Delta q.$$

Increment of a functional J: If x and $x + \delta x$ are functions for which the functional J is defined, then increment of J is

$$\Delta J = J(x + \delta x) - J(x).$$

δx is the *variation* of the function x. The increment of a functional can be written as

$$\Delta J(x(t), \delta x(t)) = \delta J(x(t), \delta x(t)) + g(x(t), \delta x(t)).\|\delta x(t)\|,$$

where δJ is linear in $\delta x(t)$. If $\lim_{\|\delta x(t)\| \to 0} (g(x(t), \delta x(t))) = 0$, then J is said to be *differentiable* on x and δJ is the *variation* of J evaluated for a function x. A *variation of* the functional is a linear approximation of its increment, i.e., $\delta J(x(t), \delta x(t))$ is linear in $\delta x(t)$.

$$\Delta J(x(t), \delta x(t)) = \delta J(x(t), \delta x(t)) + H.O.T.$$

The variation of $J(x(t)) = \int_{t_0}^{t_f} f(x(t))\mathrm{d}t$ (assuming f has first and second continuous derivative) can be obtained from

$$\delta J(x(t), \delta x(t)) = \int_{t_0}^{t_f} \frac{\partial f(x(t))}{\partial x(t)} \cdot \delta x \,\mathrm{d}t + f(x(t_f))\delta t_f - f(x(t_0))\delta t_0.$$

The final concept we review is the definition of a minimum of a function and a functional.

Minimizer of a function f(q): is $q^\star$ if

$$f(q^\star) \leq f(q)$$

for all admissible q in $\|q - q^\star\| \leq \epsilon$

Minimizer of functional $J(x(t))$: is $x^\star(t)$ if

$$J(x^\star(t)) \leq J(x(t))$$

for all admissible $x(t)$ in $\|x(t) - x^\star(t)\| \leq \epsilon$.

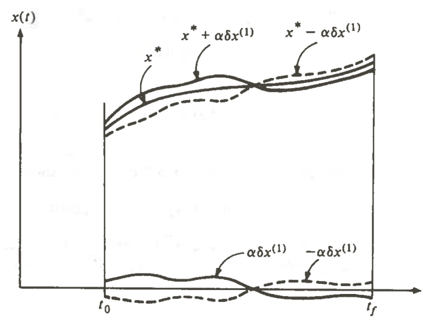
5.3 Calculus of variation

The following states the fundamental theorem of the calculus of variation:

- Let x be a vector function of t in the class Ω, and $J(x)$ be a differential functional of x.
- Assume that all $x \in \Omega$ are not constrained by any boundaries. If $x^\star$ is an extremal function, the variation of J must vanish in $x^\star$

$$\delta J(x^\star, \delta x) = 0$$

for all admissible $x \in \Omega$.



An extremal and two neighboring curves

In the following we are going to invoke the fundamental theorem of the calculus of variation to obtain equations that characterize the extremal points of different class of functional optimization problems.

5.3.1 The simplest problem in calculus of variation

Problem 1: determine scalar $x^\star(t)$ in the class of functions with continuous first derivative that is a local extremum of

$$J(x(t)) = \int_{t_0}^{t_f} g(x(t), \dot{x}, t)\mathrm{d}t$$

and respects $x(t_0) = x_0$ and $x(t_f) = x_f$ for given and fixed t_0, t_f, x_0 and x_f. □

To solve this problem we use the fundamental theorem of calculus, i.e, we look for functions $x^\star(t)$ that satisfy $\delta J(x^\star(t), \delta x(t)) = 0$.

We start with

$$\Delta J(x(t), \delta x) = \int_{t_0}^{t_f} g(x(t) + \delta x(t), \dot{x}(t) + \delta \dot{x}(t), t)\mathrm{d}t - \int_{t_0}^{t_f} g(x(t), \dot{x}, t)\mathrm{d}t.$$

We use taylor series expansion to obtain

$$\Delta J(x(t),\delta x)=\int_{t_0}^{t_f}\Big(g(x(t),\dot{x}(t),t)+\frac{\partial}{\partial x}g(x(t),\dot{x}(t),t)\delta x(t)+\frac{\partial}{\partial \dot{x}}g(x(t),\dot{x}(t),t)\delta\dot{x}(t)$$
$$+H.O.T.\Big)\mathrm{d}t-\int_{t_0}^{t_f}g(x(t),\dot{x},t)\mathrm{d}t,$$

which gives $\delta J(x^\star(t),\delta x)$, the first order approximation of $\Delta J(x(t),\delta x)$, as follows

$$\delta J(x(t),\delta x)=\int_{t_0}^{t_f}\Big(\frac{\partial}{\partial x}g(x(t),\dot{x}(t),t)\delta x(t)+\frac{\partial}{\partial \dot{x}}g(x(t),\dot{x}(t),t)\delta\dot{x}(t)\Big)\mathrm{d}t.$$

Notice that δx and $\delta\dot{x}$ are not independent variations from one anther

$$\dot{x}(t)=\frac{\mathrm{d}}{\mathrm{d}t}x(t),\quad \delta\dot{x}(t)=\frac{\mathrm{d}}{\mathrm{d}t}\delta x(t),\quad \delta x(t)=\int\delta\dot{x}\mathrm{d}t+\delta x(t_0); thatis, selecting$$

δx uniquely defines $\delta\dot{x}$. Therefore, we cannot yet draw any conclusive information from setting $\delta J(x(t),\delta x)=0$. We choose δx as a function that is varied independently. We use integration by parts (IBP) to eliminate $\delta\dot{x}$ from the expression that we have for $\delta J(x(t),\delta x)$. Recall

$$\text{IBP:}\qquad \int_1^2 u\mathrm{d}v=uv\Big|_1^2-\int_1^2 v\mathrm{d}u.$$

Let $u=g_{\dot{x}}$ and $\mathrm{d}v=\delta\dot{x}\mathrm{d}t$, then after implementing the IBP we obtain

$$\delta J(x(t),\delta x)=\int_{t_0}^{t_f}\Big\{\frac{\partial g}{\partial x}(x(t),\dot{x}(t),t)-\frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x(t),\dot{x}(t),t)]\Big\}\delta x\,\mathrm{d}t+\Big(\frac{\partial g}{\partial \dot{x}}(x(t),\dot{x}(t),t)\delta x\Big)_{t_0}^{t_f}=$$
$$\int_{t_0}^{t_f}\Big\{\frac{\partial g}{\partial x}(x(t),\dot{x}(t),t)-\frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x(t),\dot{x}(t),t)]\Big\}\delta x\,\mathrm{d}t.$$

Here, we used $\delta x(t_0)=0$ and $\delta x(t_f)=0$, because $x(t_0)=x_0$ and $x(t_f)=x_f$ are fixed and given in Problem 1. Invoking the fundamental theorem of calculus to characterize the extremal functions, for first order optimality condition we should have $\delta J(x^\star(t),\delta x)=0$ for any variation δx. Then, using the fundamental lemma of calculus of variation (see page 126 of [1]), we arrive at the following conclusion.

Both final time and final state are specified

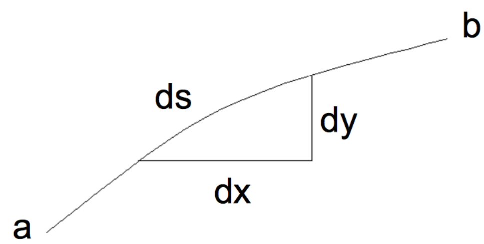
$$\frac{\partial g}{\partial x}(x^\star(t),\dot{x}^\star(t),t)-\frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x^\star(t),\dot{x}^\star(t),t)]=0,\quad\text{(Euler equation)}$$
$$x^\star(t_0)=x_0,$$
$$x^\star(t_f)=x_f$$

Problem 1: example (a)

$*$ Find the curve that gives the shortest distance between two points with known locations on a plane, $P = (x_0, y_0)$ and $Q = (x_f, y_f)$.

Cost function:

- Define the distance to be s, so, $s = \int \mathrm{d}s$



- Therefore, $s = \int \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}x)^2} = \int \sqrt{1 + (\frac{\mathrm{d}y}{\mathrm{d}x})^2}\,\mathrm{d}x$
- Take y as dependent variable, and x as independent variable, and Let $\frac{\mathrm{d}y}{\mathrm{d}x} \to \dot{y}$

Optimization problem:

$$\text{minimize } J = \int_{x_0}^{x_f} \sqrt{1 + \dot{y}^2}\mathrm{d}y \longrightarrow \text{minimize } J = \int_{x_0}^{x_f} g(\dot{y})\mathrm{d}y$$

First order necessary condition is Euler equation

$$\frac{\partial g}{\partial y}(y^\star(t), \dot{y}^\star(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{y}}(y^\star(t), \dot{y}^\star(t), t)] = 0$$

which gives the following first order necessary condition for this problem (here we have $\frac{\partial g}{\partial y} = 0$)

$$\frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{y}}(y^\star(t), \dot{y}^\star(t), t)] = 0$$

we can solve this ode problem in number of ways

- $$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\partial g}{\partial \dot{y}}\Big) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\dot{y}}{\sqrt{1 + \dot{y}^2}}\Big) = 0$$

  The first order necessary condition for optimality is then

$$\frac{\dot{y}}{\sqrt{1 + \dot{y}^2}} = \text{constant},$$
$$y(x_0) = y_0,$$
$$y(x_f) = y_f.$$

From $\frac{\dot{y}}{\sqrt{1+\dot{y}^2}}$ = constant we can deduce that $\dot{y}(x) = c_1$ = constant for $x \in [x_0, x_f]$. Therefore, we obtain

$$\dot{y} = c_1 \Rightarrow y(x) = c_1\, x + x_2$$

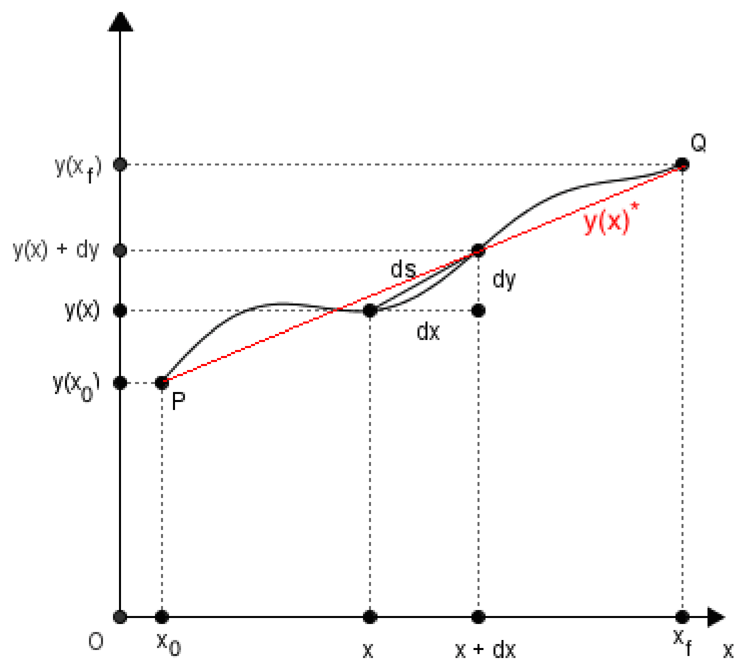
We conclude that the shortest path is a line $y = c_1\, x + c_2$. For given initial conditions, the constants are $c_1 = \frac{y_f - y_0}{x_f - x_0}$ and $c_2 = y_0 - \frac{y_f - y_0}{x_f - x_0} x_0 = \frac{y_0 x_f - y_f x_0}{x_f - x_0}$.

-
$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\partial g}{\partial \dot{y}}\Big) &= \frac{\mathrm{d}}{\mathrm{d}\dot{y}}\Big(\frac{\partial g}{\partial \dot{y}}\Big)\frac{\mathrm{d}\dot{y}}{\mathrm{d}x} \\ &= \frac{\mathrm{d}}{\mathrm{d}\dot{y}}\Big(\frac{\dot{y}}{(1+\dot{y}^2)^{1/2}}\Big)\ddot{y} = \frac{\ddot{y}}{(1+\dot{y}^2)^{3/2}} = 0\end{aligned}$$

The first order necessary condition for optimality is then

$$\begin{aligned}&\ddot{y} = 0,\\ &y(x_0) = y_0,\\ &y(x_f) = y_f.\end{aligned}$$

The most general curve with $\ddot{y} = 0$ is a line $y = c_1\, x + c_2$. For given initial conditions, the constants are $c_1 = \frac{y_f - y_0}{x_f - x_0}$ and $c_2 = y_0 - \frac{y_f - y_0}{x_f - x_0} x_0 = \frac{y_0 x_f - y_f x_0}{x_f - x_0}$.



The shortest curve connecting two points is the straight line connecting them together. □

Problem 1: example (b)-**The Brachistrochrone Problem**

Find the shortest path on which a particle in the absence of friction will slide from one point to another point in the shortest time under the action of gravity.

The most famous classical variational principle is the so-called brachistochrone problem. The compound Greek word "brachistochrone" means "minimal time". An experimenter lets a bead slide down a wire that connects two fixed points. The goal is to shape the wire in such a way that, starting from rest, the bead slides from one end to the other in minimal time. Naive guesses for the wire's optimal shape, including a straight line, a parabola, a circular arc, or even a catenary are wrong. One can do better through a careful analysis of the associated variational problem. The brachistochrone problem was originally posed by the Swiss mathematician Johann Bernoulli in 1696, and served as an inspiration for much of the subsequent development of the subject.

Solution: Let the particle slide from O along the path OP. Let at time t, the particle be at (x, y). By the principle of work and energy, we have

Kinetic energy point (x, y) ? kinetic energy at O = work done in moving the particle from O to (x, y).

$$\frac{1}{2}m(\frac{ds}{dt})^2 - 0 = mgy \Rightarrow \frac{ds}{dt} = \sqrt{2gy}.$$

Our objective is

$$\text{minimize } t_1 = \int_{t_0}^{t_1} dt, \quad s.t.\ \frac{ds}{dt} = \sqrt{2gy}.$$

We can re-write this problems as

$$t_1 = \int_0^{x_1} \frac{ds}{\sqrt{2gy}} = \frac{1}{\sqrt{2g}} \int_0^{x_1} \frac{\sqrt{1+(y')^2}}{\sqrt{y}} dx$$

Therefore, we solve the problem

$$\text{minimize } \ t_1 = \frac{1}{\sqrt{2g}} \int_0^{x_1} \frac{\sqrt{1+(y')^2}}{\sqrt{y}} dx \quad : \quad g(y, y') = \frac{\sqrt{1+(y')^2}}{\sqrt{y}}.$$

F.O.N conditions:

Since $g(y, y')$ does not explicitly depend on x, the Euler equation gives

$$g - y'\frac{\partial g}{\partial y'} = c = \text{constant}$$

$$\frac{\sqrt{1+(y')^2}}{\sqrt{y}} - y - \frac{\partial}{\partial y'}(\frac{\sqrt{1+(y')^2}}{\sqrt{y}}) = c$$

$$\frac{\sqrt{1+(y')^2}}{\sqrt{y}} - y'\frac{1}{\sqrt{y}}\frac{y'}{\sqrt{1+(y')^2}} = c$$

$$\frac{1+(y')^2-(y')^2}{\sqrt{y}\sqrt{1+(y')^2}} = c \Rightarrow \frac{1}{\sqrt{y}\sqrt{1+(y')^2}} = c \Rightarrow \sqrt{y}\sqrt{1+(y')^2} = \frac{1}{c} = \sqrt{a} \ \ (\text{we defined } \ \sqrt{a} = \frac{1}{c})$$

$$y(1+(y')^2) = a \Rightarrow y' = \sqrt{\frac{a-y}{y}} \Rightarrow \sqrt{\frac{y}{a-y}}dy = dx \Rightarrow \int_0^{x_1} dx = \int_0^{y_1} \sqrt{\frac{y}{a-y}}dy.$$

Let $y = a\sin^2\theta$. Therefore, $dy = 2\sin\theta\cos\theta d\theta$. Thus,

$$x = \int_0^\theta \sqrt{\frac{a\sin^2\theta}{a - a\sin^2\theta}} 2\sin\theta\cos\theta d\theta = \int_0^\theta \frac{\sin\theta}{\cos\theta} 2\sin\theta\cos\theta d\theta = a\int_0^\theta 2\sin^2\theta d\theta = a\int_0^\theta (1-\cos 2\theta) d\theta$$

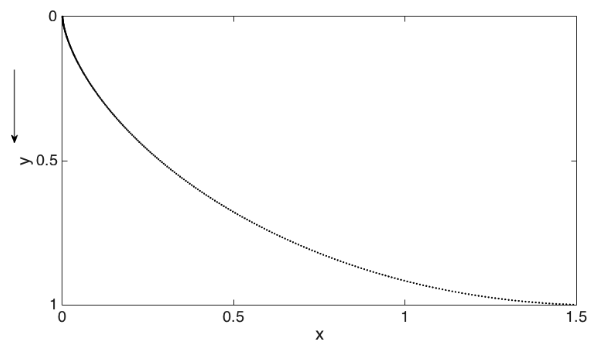
$$x = \frac{a}{2}[2\theta - \sin 2\theta].$$

Let $\frac{a}{2} = r$, $2\theta = \phi$,

$$x = r(\phi - \sin\phi); \quad y = r(1-\cos\phi),$$

which is a cycloid.

$$\begin{cases} 0 = r(\phi_0 - \sin\phi_0), \quad 0 = r(1-\cos\phi_0) \Rightarrow \phi_0 = 0 \\ x_1 = r(\phi_1 - \sin\phi_1), \quad y_1 = r(1-\cos\phi_1) \underset{x_1=1.5, y_1=1}{\Rightarrow} \phi_1 = 3.0688, r = 0.50066. \end{cases}$$




```
syms x y
[Sx,Sy] = solve(x*(y-sin(y)) == 1.5,x*(1-cos(y)) == 1)
r=Sx;
for phi=0:0.01:Sy
    x=r*(phi-sin(phi));
    y=r*(1-cos(phi));
    plot(x,-y,'k.')
    hold on
end
```

Problem 2: final time is specified but final state is free: determine scalar $x^\star(t)$ in the class of functions with continuous first derivative that is a local extremum of

$$J(x(t)) = \int_{t_0}^{t_f} g(x(t), \dot{x}, t)\mathrm{d}t$$

and respects $x(t_0) = x_0$ and $x(t_f) = x_f$ for given and fixed t_0, t_f, x_0 but free x_f. □

To solve this problem we use the fundamental theorem of calculus, i.e, we look for functions $x^\star(t)$ that satisfy $\delta J(x^\star(t), \delta x(t)) = 0$.

We start with

$$\Delta J(x(t), \delta x) = \int_{t_0}^{t_f} g(x(t) + \delta x(t), \dot{x}(t) + \delta\dot{x}(t), t)\mathrm{d}t - \int_{t_0}^{t_f} g(x(t), \dot{x}, t)\mathrm{d}t.$$

with manipulations similar to those use in Problem 1, we can show that the variation for this problem is

$$\begin{aligned}\delta J(x(t), \delta x) =& \int_{t_0}^{t_f} \Big\{\frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t)]\Big\}\delta x\,\mathrm{d}t + \Big(\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t)\delta x\Big)_{t_0}^{t_f} = \\ & \frac{\partial g}{\partial \dot{x}}(x(t_f), \dot{x}(t_f), t_f)\delta x(t_f) + \int_{t_0}^{t_f} \Big\{\frac{\partial g}{\partial x}(x(t), \dot{x}(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x(t), \dot{x}(t), t)]\Big\}\delta x\,\mathrm{d}t.\end{aligned}$$

Here, we used $\delta x(t_0) = 0$ because $x(t_0) = x_0$ is given and fixed. But because the final state is free, we have $\delta x(t_f) \neq 0$. Invoking the fundamental theorem of calculus to characterize the extremal functions, for first order optimality condition we should have $\delta J(x^\star(t), \delta x) = 0$ for any variation δx and $\delta x(t_f)$. Then, using the fundamental lemma of calculus of variation (see page 126 of [1]), we arrive at the following conclusion.

Final time is specified but final state is free

$$\frac{\partial g}{\partial x}(x^\star(t), \dot{x}^\star(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x^\star(t), \dot{x}^\star(t), t)] = 0, \quad \text{(Euler equation)}$$

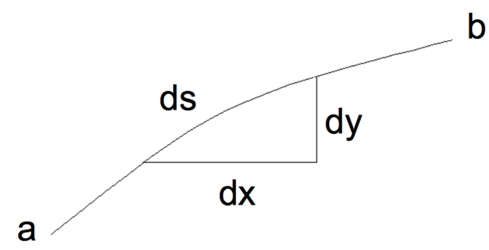
$$x^\star(t_0) = x_0,$$

$$g_{\dot{x}}(x^\star(t_f), \dot{x}^\star(t_f), t_f) = 0$$

Problem 2: example

$*$ Find the shortest length smooth curve that connects a given point $P = (x_0, y_0)$ on a plane to another point $Q = (x_f, y_f)$ where x_f is specified and given but y_f is free to attain any value. Cost function:

- Define the distance to be s, so, $s = \int \mathrm{d}s$



- Therefore, $s = \int \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}x)^2} = \int \sqrt{1 + (\frac{\mathrm{d}y}{\mathrm{d}x})^2}\,\mathrm{d}x$
- Take y as dependent variable, and x as independent variable, and Let $\frac{\mathrm{d}y}{\mathrm{d}x} \to \dot{y}$

Optimization problem:

$$\text{minimize } J = \int_{x_0}^{x_f} \sqrt{1 + \dot{y}^2}\mathrm{d}y \longrightarrow \text{minimize } J = \int_{x_0}^{x_f} g(\dot{y})\mathrm{d}y$$

First order necessary conditions are

$$\begin{aligned} &\frac{\partial g}{\partial y}(y^\star(t), \dot{y}^\star(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{y}}(y^\star(t), \dot{y}^\star(t), t)] = 0, \\ &g_{\dot{y}}(y^\star(x_f), \dot{y}^\star(x_f), x_f) = 0 \end{aligned}$$

which gives the following first order necessary condition for this problem (here we have $\frac{\partial g}{\partial y} = 0$)

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{y}}(y^\star(t), \dot{y}^\star(t), t)] = 0, \\ &\frac{\dot{y}(x_f)}{\sqrt{1 + \dot{y}(x_f)^2}} = 0 \Rightarrow \dot{y}(x_f) = 0 \end{aligned}$$

we can solve the Euler equation similar to problem 1:

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\partial g}{\partial \dot{y}}\Big) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\dot{y}}{\sqrt{1 + \dot{y}^2}}\Big) = 0$$

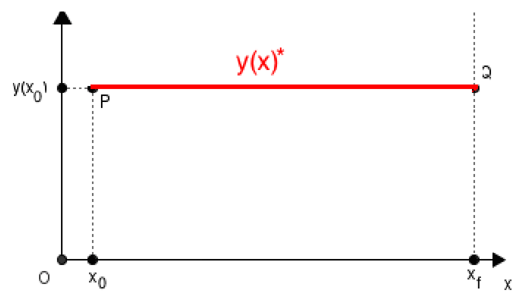
The first order necessary condition for optimality is then

$$\frac{\dot{y}}{\sqrt{1+\dot{y}^2}} = \text{constant},$$
$$y(x_0) = y_0,$$
$$\dot{y}(x_f) = 0$$

From $\frac{\dot{y}}{\sqrt{1+\dot{y}^2}} = \text{constant}$ we can deduce that $\dot{y}(x) = c_1 = \text{constant}$ for $x \in [x_0, x_f]$. From the boundary condition $\dot{y}(x_f) = 0$ we obtain that $c_1 = 0$. Therefore, we obtain

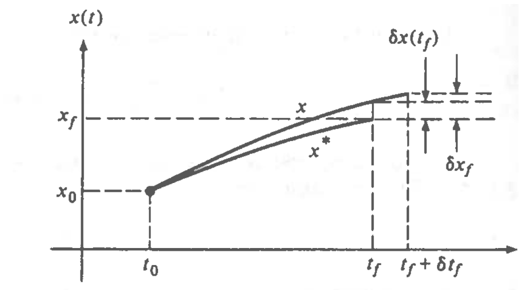
$$\dot{y} = 0 \Rightarrow y(x) = c_2$$

We conclude that the shortest path is the line $y(x) = c_2$. For given initial condition $y(x_0) = y_0$, we obtain $c_2 = y_0$.



The shortest path connecting two points P and Q where y_f is free is the straight horizontal line $y(x) = y_0$, $x \in [x_0, x_f]$. $\square$

Problem 3: both final time t_f and $x(t_f)$ are free



(photo courtesy of [1])

$$
\begin{aligned}
&\delta x(t_f) = x(t_f) - x^\star(t_f) \neq 0, \quad \delta t_f \neq 0,\\
&\delta x_f = x(t_f + \delta t_f) - x^\star(t_f) \neq 0,\\
&\qquad \delta x_f \neq \delta x(t_f),\\
&\delta x_f \approx \delta x(t_f) + \dot{x}^\star(t_f)\delta t_f
\end{aligned}
$$

The variation in this problem is

$$\delta J(x(t), \delta x) = \int_{t_0}^{t_f} \{\big(g_x - \frac{\mathrm{d}}{\mathrm{d}t} g_{\dot{x}}\big) \cdot \delta x(t)\}\mathrm{d}t + g_{\dot{x}}(x(t_f), \dot{x}(t_f), t_f) \cdot \delta x(t_f) + g(x(t_f), \dot{x}(t_f), t_f)\delta t_f$$

Next, substitute $\delta x(t_f) = \delta x_f - \dot{x}^\star(t_f)\delta t_f$, in $\delta J(x(t), \delta x)$ to obtain

$$
\begin{aligned}
\delta J(x(t), \delta x) = &\int_{t_0}^{t_f} \big(g_x - \frac{\mathrm{d} g_{\dot{x}}}{\mathrm{d}t}\big) \cdot \delta x(t)\mathrm{d}t + g_{\dot{x}}(x(t_f), \dot{x}(t_f), t_f) \cdot \delta x_f\\
&+ \big(g(x(t_f), \dot{x}(t_f), t_f) - g_{\dot{x}}(x(t_f), \dot{x}(t_f), t_f) \cdot \dot{x}^\star(t_f)\big)\delta t_f = 0
\end{aligned}
$$

Any extremum $x^\star(t)$ should satisfy the following

$$
\begin{aligned}
&\frac{\partial g}{\partial x}(x^\star(t), \dot{x}^\star(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x^\star(t), \dot{x}^\star(t), t)] = 0,\\
&x^\star(t_0) = x_0,
\end{aligned}
$$

depending on the relationship between $x(t_f)$ and t_f, different set of terminal boundary conditions are obtained

1. Unrelated
2. related by $x(t_f) = \Theta(t)$
3. constrained relationship $m(x(t_f), t_f) = 0$

Problem 3-1: both final time t_f and $x(t_f)$ are free and unrelated

In this case $\delta x_f \neq 0$ and $\delta t_f \neq 0$ are independent from each other. Therefore the variation is

$$\begin{aligned}\delta J(x(t), \delta x) = &\int_{t_0}^{t_f} \big(g_x - \frac{\mathrm{d}g_{\dot{x}}}{\mathrm{d}t}\big) \cdot \delta x(t)\mathrm{d}t + g_{\dot{x}}(x^\star(t_f), \dot{x}^\star(t_f), t_f) \cdot \delta x_f \\ &+ \big(g(x^\star(t_f), \dot{x}^\star(t_f), t_f) - g_{\dot{x}}(x^\star(t_f), \dot{x}^\star(t_f), t_f) \cdot \dot{x}^\star(t_f)\big)\delta t_f = 0\end{aligned}$$

Using the Fundamental Theorem of Calculus of Variation then the first order necessary conditions for optimality are (based on free variation for $\delta x(t)$, δx_f and δt_f)

Final time t_f and $x(t_f)$ are free and unrelated

$$\frac{\partial g}{\partial x}(x^\star(t), \dot{x}^\star(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x^\star(t), \dot{x}^\star(t), t)] = 0, \quad \text{(Euler equation)}$$
$$x^\star(t_0) = x_0,$$
$$g_{\dot{x}}(x^\star(t_f), \dot{x}^\star(t_f), t_f) = 0,$$
$$g(x^\star(t_f), \dot{x}^\star(t_f), t_f) - g_{\dot{x}}(x^\star(t_f), \dot{x}^\star(t_f), t_f) \cdot \dot{x}^\star(t_f) = 0.$$

Problem 3-2: final time t_f and $x(t_f)$ are free but related as $x(t_f) = \Theta(t_f)$

In this case $\delta x_f \neq 0$ and $\delta t_f \neq 0$ but we have

$$\delta x_f = \frac{\mathrm{d}\Theta}{\mathrm{d}t}\delta t_f$$

Substituting this relation in the variation equation below for free final time and free final state

$$\begin{aligned}\delta J(x(t), \delta x) = &\int_{t_0}^{t_f} \big(g_x - \frac{\mathrm{d}g_{\dot{x}}}{\mathrm{d}t}\big) \cdot \delta x(t)\mathrm{d}t + g_{\dot{x}}(x^\star(t_f), \dot{x}^\star(t_f), t_f) \cdot \delta x_f \\ &+ \big(g(x^\star(t_f), \dot{x}^\star(t_f), t_f) - g_{\dot{x}}(x^\star(t_f), \dot{x}^\star(t_f), t_f) \cdot \dot{x}^\star(t_f)\big)\delta t_f = 0\end{aligned}$$

we obtain

$$\begin{aligned}\delta J(x(t), \delta x) = &\int_{t_0}^{t_f} \big(g_x - \frac{\mathrm{d}g_{\dot{x}}}{\mathrm{d}t}\big) \cdot \delta x(t)\mathrm{d}t + \\ &\Big(g_{\dot{x}}(x^\star(t_f), \dot{x}^\star(t_f), t_f) \cdot \Big[\frac{\mathrm{d}\Theta}{\mathrm{d}t}\Big|_{t_f} - \dot{x}^\star(t_f)\Big] + g(x^\star(t_f), \dot{x}^\star(t_f), t_f)\Big)\delta t_f = 0\end{aligned}$$

Here, we used $\delta x(t_0) = 0$ because $x(t_0) = x_0$ is given and fixed. But because the final time is free, we have $\delta t_f \neq 0$. Invoking the fundamental theorem of calculus to characterize the extremal functions, for first order optimality condition we should have $\delta J(x^\star(t), \delta x) = 0$ for any variation δx and δt_f. Then, using the fundamental lemma of calculus of variation (see page 126 of [1]), we arrive at the following conclusion.

Final time t_f and $x(t_f)$ are free but related as $x(t_f) = \Theta(t_f)$

$$\frac{\partial g}{\partial x}(x^\star(t), \dot{x}^\star(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x^\star(t), \dot{x}^\star(t), t)] = 0, \quad \text{(Euler equation)}$$

$x^\star(t_0) = x_0,$

$x^\star(t_f) = \Theta(t_f),$

$$g_{\dot{x}}(x^\star(t_f), \dot{x}^\star(t_f), t_f) \cdot \Big[\frac{\mathrm{d}\Theta}{\mathrm{d}t}\Big|_{t_f} - \dot{x}^\star(t_f)\Big] + g(x^\star(t_f), \dot{x}^\star(t_f), t_f) = 0, \quad \text{(Transversality condition)}.$$

Problem 3-2: example

$*$ Find the shortest length smooth curve that connects a given point $P = (x_0 = 0, y_0 = 0)$ to a point $Q = (x_f, y_f)$ on the surface $\Theta(x) = -5\,x + 15$, i.e., $y_f = -5\,x_f + 15$.
The optimization problem in this case is Optimization problem:

$$\begin{aligned}&\text{minimize } J = \int_{x_0}^{x_f} \sqrt{1+\dot{y}^2}\mathrm{d}y \longrightarrow \text{minimize } J = \int_{x_0}^{x_f} g(\dot{y})\mathrm{d}y, \quad s.t.\\ &y_f = -5\,x_f + 15\end{aligned}$$

$$\Theta(x) = -5x + 15 \Rightarrow \frac{\mathrm{d}\Theta(x)}{x} = -5$$

First order necessary conditions:

- Since $g(y, \dot{y}, t) = g(\dot{y})$ is only a function of $\dot{y}$, Euler equation reduces to

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\partial g}{\partial \dot{y}}\Big) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\dot{y}}{\sqrt{1+\dot{y}^2}}\Big) = 0$$

  which after differentiating and simplifying, gives the answer as a straight line

$$y^\star(x) = c_1 x^\star + c_2$$

  but since $y(0) = 0$, then $c_2 = 0$.

- Transversality condition gives

$$\Big(\frac{\dot{y}^\star(x_f^\star)}{\sqrt{1+\dot{y}^\star(x_f^\star)^2}}\Big)(-5 - \dot{y}^\star(x_f^\star)) + (\sqrt{1+\dot{y}^\star(x_f^\star)^2}) = 0$$

  which simplifies to

$$(\dot{y}^\star(x_f^\star))(-5 - \dot{y}^\star(x_f^\star)) + (1+\dot{y}^\star(x_f^\star)^2) = -5\dot{y}^\star(x_f^\star) + 1 = 0.$$

  Therefore $\dot{y}^\star(x_f^\star) = c_1 = 1/5$. The solution is then

$$y^\star(x) = \frac{1}{5}x^\star, \quad x \in [x_0, x_f]$$

  Not a surprise, as this gives the slope of a line orthogonal to the constraint line.

- To find the final $x_f^\star$ we use

$$\begin{aligned}y^\star(x_f^\star) &= \frac{1}{5}x_f^\star,\\ y_f^\star = \Theta^\star(x_f^\star) &= -5x_f^\star + 15,\end{aligned}$$

  which gives $x_f^\star \approx 2.88$. Then, the point is $Q = (2.88, \frac{2.88}{5})$.

The shortest path connecting point $P=(0,0)$ to surface $\Theta(x)=-5x+15$ is the straight line $y(x)=\frac{1}{5}x$, $x\in[0,2.88,]$. This line is orthogonal to the surface (x). The contact point on the surface is $Q=(2.88,\frac{2.88}{5})$. $\square$

Next consider the case that the terminal surface is

$$\Theta(x)=\frac{1}{2}((x-5)^2-1).$$

In this case we have

$$\frac{\mathrm{d}\Theta(x)}{x}=x-5.$$

First order optimality conditions:

- The Euler equation and the initial condition $y(0)=0$ still gives

$$y^\star(x)=c_1x^\star,\quad x\in[0,x_f]$$

- Transversality condition here gives

$$\Big(\frac{\dot{y}^\star(x_f^\star)}{\sqrt{1+\dot{y}^\star(x_f^\star)^2}}\Big)(x_f^\star-5-\dot{y}^\star(x_f^\star))+(\sqrt{1+\dot{y}^\star(x_f^\star)^2})=0$$

  which simplifies to

$$(\dot{y}^\star(x_f^\star))(x_f^\star-5-\dot{y}^\star(x_f^\star))+(1+\dot{y}^\star(x_f^\star)^2)=c_1(x_f^\star-5)+1=0.$$

  Which gives

$$c_1=-\frac{1}{x_f^\star-5}$$

  as a result

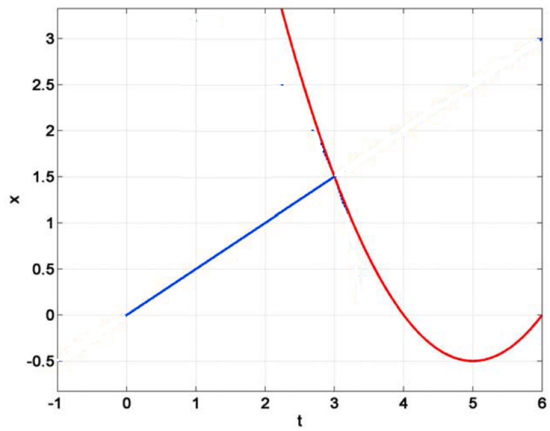
$$y^\star(x)=-\frac{1}{x_f^\star-5}x^\star,\quad x\in[0,x_f]$$

- Now consider $y^\star(x)$ and $\Theta(x)$ at $x_f^\star$

$$\begin{cases}y^\star(x_f^\star)=-\frac{1}{x_f^\star-5}x_f^\star,\\ y^\star(x_f^\star)=\Theta(x_f^\star)=\frac{1}{2}((x_f^\star-5)^2-1)\end{cases}\qquad -\frac{x_f^\star}{x_f^\star-5}=\frac{1}{2}((x_f^\star-5)^2-1),$$

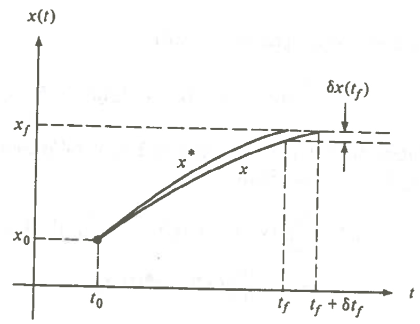
  which gives $x_f^\star=3$. As a result $c_1=\frac{1}{2}$ and

$$y^\star(x)=\frac{1}{2}x^\star,\quad x\in[0,x_f]$$



The shortest path connecting point $P = (0, 0)$ to surface $\Theta(x) = \frac{1}{2}((x-5)^2 - 1)$ is the straight line $y(x) = \frac{1}{2}x$, $x \in [0, 3]$. This line is orthogonal to the surface (x) (how would you check this). The contact point on the surface is $Q = (3, \frac{3}{2})$. $\square$

Problem 4: free final time but fixed and pre-specified final state



(photo courtesy of [1])

t_f is unknown.

$\delta x(t_f)$ is neither zero or free, it depends on $\delta t_f \neq 0$ However, in this case $\delta x_f = x(t_f + \delta t_f) - x^\star(t_f) = 0$

$$\delta x(t_f) + \dot{x}(t_f)\delta t_f = 0 \quad \text{(see the fig)}$$

$$\delta x(t_f) = -\dot{x}(t_f)\delta t_f, \quad \delta t_f \neq 0$$

The variation in this case is

$$\delta J(x(t), \delta x) = \int_{t_0}^{t_f} \{\big(g_x - \frac{\mathrm{d}}{\mathrm{d}t} g_{\dot{x}}\big) \cdot \delta x(t)\}\, \mathrm{d}t + \big(g(x(t_f), \dot{x}(t_f), t_f) - g_{\dot{x}}(x(t_f), \dot{x}(t_f), t_f) \cdot \dot{x}(t_f)\big)\delta t_f = 0$$

Here, we used $\delta x(t_0) = 0$ because $x(t_0) = x_0$ is given and fixed. But because the final time is free, we have $\delta t_f \neq 0$. Invoking the fundamental theorem of calculus to characterize the extremal functions, for first order optimality condition we should have $\delta J(x^\star(t), \delta x) = 0$ for any variation δx and δt_f. Then, using the fundamental lemma of calculus of variation (see page 126 of [1]), we arrive at the following conclusion.

Free final time t_f but fixed and pre-specified final state $x(t_f)$

$$\frac{\partial g}{\partial x}(x^\star(t), \dot{x}^\star(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x^\star(t), \dot{x}^\star(t), t)] = 0, \quad \text{(Euler equation)}$$
$$x^\star(t_0) = x_0,$$
$$x^\star(t_f) = x_f,$$
$$g(x^\star(t_f), \dot{x}^\star(t_f), t_f) - g_{\dot{x}}(x^\star(t_f), \dot{x}^\star(t_f), t_f) \cdot \dot{x}^\star(t_f) = 0.$$

5.3.2 Piecewise-smooth extremals

So far we focused on admissible $x(t)$ that are continuous with continuous first first derivatives. We want to expand to class if piecewise-smooth admissible functions. Some reasons for expanding to this class of admissible curves are

- control input is no smooth (e.g., subject to nonlinearities)

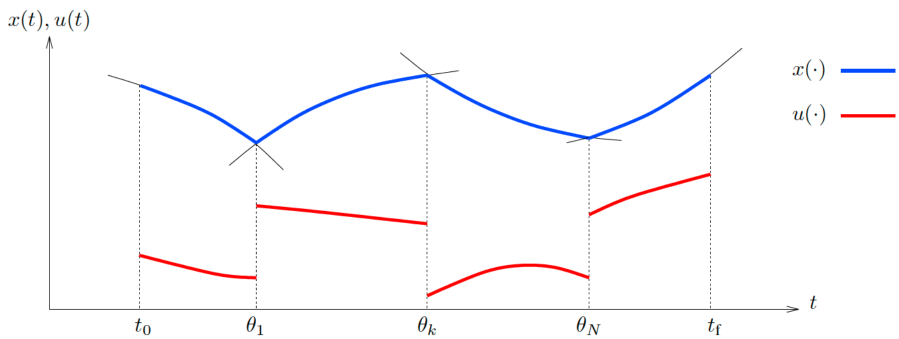


Illustration of a piecewise continuous control $u \in \hat{\mathcal{C}}[t_0, t_f]$ (red line), and the corresponding piecewise continuously differentiable response $x \in \hat{\mathcal{C}}^1[t_0, t_f]$ (blue line).

- intermediate state constraints are imposed (e.g., we want to go from a point at t_0 to another point in t_f while minimizing our cost but we want to touch a point on a constraint surface in between $[t_0, t_f]$, i.e., $x(t_1) = \theta(t_1)$, $t_1 \in [t_0, t_f]$).

Problem: determine vector function $x^\star(t)$ in the class of functions with *piecewise-continuous* first derivative that is a local extremum of

$$J(x(t)) = \int_{t_0}^{t_f} g(x(t), \dot{x}, t)\mathrm{d}t$$

and respects $x(t_0) = x_0 \in \mathbb{R}^n$ for given and fixed t_0, x_0, t_f and $x(t_f)$. $\square$

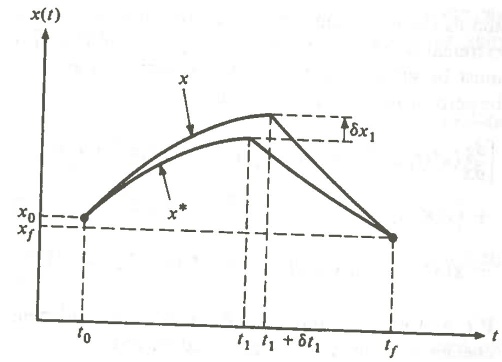
In the following developments we assume that there is only one corner between $[t_0, t_f]$. Expansion to include more corners is explained later.

- Assume $\dot{x}$ has a discontinuity at $t_1 \in (t_0, t_f)$, where t_1 is not fixed (or known)

$$J(x(t)) = \int_{t_0}^{t_f} g(x(t), \dot{x}, t)\mathrm{d}t = \underbrace{\int_{t_0}^{t_1} g(x(t), \dot{x}, t)\mathrm{d}t}_{J_1} + \underbrace{\int_{t_1}^{t_f} g(x(t), \dot{x}, t)\mathrm{d}t}_{J_2}.$$

- Do as before

$$
\begin{aligned}
\delta J =&\delta J_1 + \delta J_2 = \\
&\int_{t_0}^{t_1} (\frac{\partial g}{\partial x}\delta x + \frac{\partial g}{\partial \dot{x}}\delta \dot{x})\mathrm{d}t + g(t_1^-)\delta t_1 + \\
&\int_{t_1}^{t_f} (\frac{\partial g}{\partial x}\delta x + \frac{\partial g}{\partial \dot{x}}\delta \dot{x})\mathrm{d}t - g(t_1^+)\delta t_1
\end{aligned}
$$



(photo courtesy of [1])

IBP:

$$
\begin{aligned}
&\delta J = \delta J_1 + \delta J_2 = \\
&\int_{t_0}^{t_1} (\frac{\partial g}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial g}{\partial \dot{x}})\delta x \mathrm{d}t + \big(g(t_1^-) - g_{\dot{x}}(t_1^-)\dot{x}(t_1^-)\big)\delta t_1 + g_{\dot{x}}(t_1^-)\delta x_1 + \\
&\int_{t_0}^{t_1} (\frac{\partial g}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial g}{\partial \dot{x}})\delta x \mathrm{d}t + \big(g(t_1^+) - g_{\dot{x}}(t_1^+)\dot{x}(t_1^+)\big)\delta t_1 + g_{\dot{x}}(t_1^+)\delta x_1
\end{aligned}
$$

$$
\begin{aligned}
\dot{x}(t_1^-) &\neq \dot{x}(t_1^+), \\
\text{from left side:} \quad \delta x_1 &\approx \delta x(t_1^-) + \dot{x}(t_1^-)\delta t_1, \Rightarrow \\
\delta x(t_1^-) &= \dot{x}(t_1^-)\delta t_1 - \delta x_1, \\
\text{from right side:} \quad \delta x_1 &\approx \delta x(t_1^+) + \dot{x}(t_1^+)\delta t_1, \Rightarrow \\
\delta x(t_1^+) &= \dot{x}(t_1^+)\delta t_1 - \delta x_1
\end{aligned}
$$

− − − − − − − − − − − − − − −

For problem with non-smooth class of admissible curves defined above (only one corner), the following equations characterize its extremals:

$$
\begin{aligned}
&\frac{\partial g}{\partial x}(x^\star(t), \dot{x}^\star(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x^\star(t), \dot{x}^\star(t), t)] = 0, \quad \text{(Euler equation)} \\
&x^\star(t_0) = x_0, \\
&x^\star(t_f) = x_f, \\
&\left.\begin{array}{l} g(t_1^-) - g_{\dot{x}}(t_1^-)\dot{x}(t_1^-) = g(t_1^+) - g_{\dot{x}}(t_1^+)\dot{x}(t_1^+), \\ g_{\dot{x}}(t_1^-) = g_{\dot{x}}(t_1^+), \end{array}\right\} \quad \text{Weierstrass-Erdmann conditions.}
\end{aligned}
$$

- For several corners, there are a set of Weierstrass-Erdmann conditions for each corner.

− − − − − − − − − − − − − − −

When the corner point is not free, instead it is required to satisfy an intermediate time constraint of the form $x(t_1) = \theta(t_1)$, δx_1 and δt_1 are not independent. In this case we use

$$
\delta x_1 = \frac{\mathrm{d}\theta}{\mathrm{d}t}\delta t_1 = \dot{\theta}\delta t_1,
$$

to write the variation δJ in terms of independent variations δx and δt_1 as follows

$$
\begin{aligned}
&\delta J = \delta J_1 + \delta J_2 = \\
&\int_{t_0}^{t_1} (\frac{\partial g}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial g}{\partial \dot{x}})\delta x \mathrm{d}t + \big(g(t_1^-) - g_{\dot{x}}(t_1^-)\dot{x}(t_1^-) + g_{\dot{x}}(t_1^-)\dot{\theta}(t_1^-)\big)\delta t_1 + \\
&\int_{t_0}^{t_1} (\frac{\partial g}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial g}{\partial \dot{x}})\delta x \mathrm{d}t + \big(g(t_1^+) - g_{\dot{x}}(t_1^+)\dot{x}(t_1^+) + g_{\dot{x}}(t_1^+)\dot{\theta}(t_1^+)\big)\delta t_1
\end{aligned}
$$

Then the necessary conditions for the extremal is

$- - - - - - - - - - - - - - -$

$$
\begin{aligned}
&\frac{\partial g}{\partial x}(x^\star(t), \dot{x}^\star(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x^\star(t), \dot{x}^\star(t), t)] = 0, \quad \text{(Euler equation)}\\
&x^\star(t_0) = x_0,\\
&x^\star(t_f) = x_f,\\
&g(t_1^-) + g_{\dot{x}}(t_1^-)(\dot{\theta})(t_1^-) - \dot{x}(t_1^-)) = g(t_1^+) + g_{\dot{x}}(t_1^+)(\dot{\theta}(t_1^+) - \dot{x}(t_1^+)),\\
&x(t_1) = \theta(t_1)
\end{aligned}
$$

- For several corners, there are a set of the last two conditions for each corner.

$- - - - - - - - - - - - - - -$

Example: find the shortest path that connects $x(-2) = 0$ to $x(1) = 0$ and touches the curve $x(t) = t^2 + 3$.

The shortest path cost function is

$$J = \int_{-2}^{1} \sqrt{1 + \dot{x}^2}\mathrm{d}t.$$

Because of the constraint that requires our curve to touch a constraint surface in an intermediate time, we allow our admissible curves $x(t)$ to have one corner at time t_1. The conditions for the extremal are

$$\frac{\partial g}{\partial x}(x^\star(t), \dot{x}^\star(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}[\frac{\partial g}{\partial \dot{x}}(x^\star(t), \dot{x}^\star(t), t)] = 0, \rightarrow \begin{cases} \text{left segment: } x^\star(t) = c_{11} + c_{12}\, t, & t \in [-2, t_1^-] \\ \text{right segment: } x^\star(t) = c_{21} + c_{22}\, t, & t \in [t_1^+, 1] \end{cases}$$

$$x^\star(-2) = 0, \rightarrow c_{11} + c_{12}(-2) = 0,$$

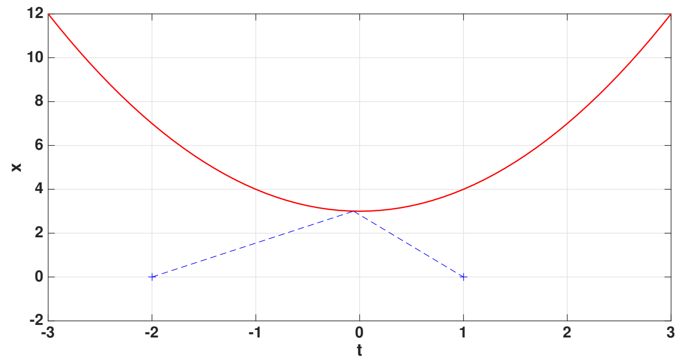
$$x^\star(t_f) = x_f, \rightarrow c_{21} + c_{22}(1) = 0,$$

$$g(t_1^-) + g_{\dot{x}}(t_1^-)(\dot{\theta})(t_1^-) - \dot{x}(t_1^-)) = g(t_1^+) + g_{\dot{x}}(t_1^+)(\dot{\theta}(t_1^+) - \dot{x}(t_1^+)) \rightarrow (1 + \dot{x}^2(t_1^-))^{1/2} + \frac{\dot{x}^\star(t_1^-)}{(1 + \dot{x}^2(t_1^-))^{1/2}}(2t_1^- - \dot{x}(t_1^-)) = (1 + \dot{x}^2(t_1^+))^{1/2} + \frac{\dot{x}^\star(t_1^+)}{(1 + \dot{x}^2(t_1^+))^{1/2}}(2t_1^+ - \dot{x}(t_1^+)),$$

$$x(t_1) = \theta(t_1) \rightarrow \begin{cases} c_{11} + c_{12}t_1 = t_1^2 + 3, & t \in [-2, t_1^-], \\ c_{21} + c_{22}t_1 = t_1^2 + 3, & t \in [t_1^+, 1]. \end{cases}$$

The equations to solve are

$$c_{11} - 2\,c_{12} = 0,$$
$$c_{21} + c_{22} = 0,$$
$$c_{11} + c_{12}t_1 = t_1^2 + 3,$$
$$c_{11} + c_{12}t_1 = t_1^2 + 3,$$
$$\frac{1 + 2c_{12}t_1}{(1 + c_{12}^2)^{1/2}} = \frac{1 + 2c_{22}t_1}{(1 + c_{22}^2)^{1/2}}.$$



Using Matlab 'fsolve' command you can solve these equations to obtain

$c_{11} = 3.0947,\ c_{12} = 1.5474,\ c_{21} = 2.8362,\ c_{22} = -2.8362,\ t_1 = -0.0590.$ $\square$

Matlab code

```
function F=myfunc(x); %
% x=[c11 c12  c21 c22 t1]; %
F=[x(1)-2*x(2);
x(3)+x(4);
x(1)+x(2)*x(5) -(x(5)^2+3);
x(3)+x(4)*x(5) -(x(5)^2+3);
(1+2*x(2)*x(5))/(1+x(2)^2)^(1/2) -(1+2*x(4)*x(5))/(1+x(4)^2)^(1/2)];
return %
x = fsolve('myfunc',[2 1 2 -2 0]')
```

Note 6

Optimal Control of Continuous-time systems

The treatment corresponds to selected parts from Chapter 5 [1]. The presentation is at times informal. For rigorous treatments, students should consult the aforementioned references and the other listed texts in the class syllabus.

6.1 Optimal control for problems with no inequality constraint

Problem definition

$$u^\star(t)\Big|_{t\in[t_0,t_f]} = \underset{u(t)\in\mathcal{U}}{\operatorname{argmin}}(J = h(x(t_f),t_f) + \int_{t_0}^{t_f} g(x(t),u(t),t))\mathrm{d}t, \;\; s.t.$$

$$\dot{x}(t) = a(x(t),u(t),t),$$

$$x(t_0),\; t_0 \text{ is given},$$

$$m(x(t_f),t_f) = 0 \leftarrow \text{ when final state is constrained},$$

$$x(t):\mathbb{R}\to\mathbb{R}^n, \quad u(t):\mathbb{R}\to\mathbb{R}^m, \quad f:\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}\to\mathbb{R}^n.$$

Hamiltonian $H(x,u,p,t) = g(x(t),u(t),t) + p(t)^\top a(x(t),u(t),t)$,

- First order conditions for extremal solution

$$\dot{p} = -H_x, \qquad (n \text{ dimensional})$$

$$0 = H_u, \qquad (m \text{ dimensional})$$

$$\dot{x} = H_p : \quad \dot{x} = a(x,u,t), \qquad (n \text{ dimensional})$$

- Boundary conditions

- $x(t_0) = x_0$
- if t_f free: $\frac{\partial h}{\partial t}\Big|_{t_f} + H(t_f) = 0$
- if $x_i(t_f)$ is fixed: $x_i(t_f) = x_{i_f}$
- if $x_i(t_f)$ is free: $p_i(t_f) = \frac{\partial h}{\partial x_i}(t_f)$

$$m(x(t_f), t_f) = 0$$

Let $w(x(t_f), v, t_f) = h(x(t_f), t_f) + v^\top m(x(t_f), t_f)$

- $x(t_0) = x_0$
- since $\mathbf{x}(t_f)$ is not directly given we need $p(t_f) = \frac{\partial w}{\partial x}(t_f)$
- if t_f free: $\frac{\partial w}{\partial t}\Big|_{t_f} + H(t_f) = 0$ (disappears if t_f known)

6.1.1 BVP4C Matlab package to solve optimal control problems

BVP4C package of Matlab solves problems of the form

$$\begin{aligned}&\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, t, \mathbf{p}), \quad a \le t \le b, \text{ subject}\\ &\mathbf{g}(\mathbf{y}(a), \mathbf{y}(b)) = \mathbf{0},\end{aligned} \tag{6.1.1}$$

where $\mathbf{y}$ are the variables of interest, and $\mathbf{p}$ are extra variables in the problem that can also be optimized. For technical details regarding the solution approach used by the BVP4C package see https://www.mathworks.com/help/matlab/ref/bvp4c.html.

BVP4C package can be solved optimal control problems with the exception of free end time problems, because the time period $t \in [a, b]$ should be specified in BVP4C. To use BVP4C, for problems with unspecified time t_f, we need to convert our optimal control problem of interest into the standard form (6.1.1) using some intermediate conversions described in
U. Ascher and R. D. Russell, "Reformulation of Boundary Value Problems into Standard Form", SIAM Review, Vol. 23, No. 2, 238-254, 1981.
The Key step is to rescale time so that

$$\tau = \frac{t}{t_f},$$

so that $\tau \in [0, 1]$; the assumption is that $t_0 = 0$. The, since

$$\mathrm{d}\tau = \frac{\mathrm{d}t}{t_f},$$

we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\tau} = t_f \frac{\mathrm{d}}{\mathrm{d}t}.$$

Next, we need to introduce a dummy state r that corresponds to t_f with the trivial dynamics $\dot{s} = 0$, and to replace all instances of t_f in the boundary conditions with state s. Then the solver will peak an appropriate constant for $s = t_f$. The followings demonstrates the process.

$$\begin{aligned}&t \in [0, t_f] \to \tau \in [0, 1],\\ &\dot{\mathbf{x}} = \mathbf{a}(\mathbf{x}, \mathbf{u}, t), \to \mathbf{x}' = t_f \mathbf{a}(\mathbf{x}, \mathbf{u}, t_f \tau) \to \mathbf{x}' = s\, \mathbf{a}(\mathbf{x}, \mathbf{u}, s, \tau)\\ &\dot{\mathbf{p}} = -H_{\mathbf{x}}, \to \mathbf{p}' = -t_f H_{\mathbf{x}}, \to \mathbf{p}' = -s\, H_{\mathbf{x}},\\ &H_{\mathbf{u}} = 0.\end{aligned}$$

For boundary conditions, any explicit instance of t_f that appears at the boundary conditions should be replaced by $s(1)$. The terminal conditions should be evaluated at $b = 1$.

6.1.2 Sample problems

Sample Problem 1: Consider the optimal control problem

$$\begin{aligned} J =& \text{minimize} \frac{1}{2}\int_0^1 \big(\frac{1}{2}u^2(t) - 2x(t)\big)\mathrm{d}t, \quad s.t. \\ &\dot{x} = 2 - u(t), \\ &x(0) = 2. \end{aligned}$$

- To find the optimal control we start by forming the Hamiltonian

$$H = \frac{1}{2}u^2(t) - 2x(t) + p(t)(2 - u(t)).$$

- Candidate optimal controls should satisfy the F.O.N.Cs below

$$\begin{aligned} &\dot{x} = H_p = 2 - u(t), \quad x(0) = 2, \\ &\dot{p} = -H_x = 2, \qquad\quad p(1) = 0, \text{(free terminal state; there is no terminal cost)}, \\ &0 = H_u = u(t) - p(t). \end{aligned}$$

- The co-state equation trivially yields,

$$p^\star(t) = 2\,t - 2, \quad t \in [0, 1]$$

- Then, the optimal control from $H_u = u(t) - p(t) = 0$ is obtained as

$$u^\star(t) = 2\,t - 2, \quad t \in [0, 1].$$

- Note that $u^\star$ is indeed the candidate minimizer because $H_{uu} = 1 \geq 0$.
- Using the optimal control yields

$$\dot{x} = 2 - (2\,t - 2) = -2\,t + 4,$$

  which given the initial condition $x(0) = 2$ yields

$$x^\star(t) = -t^2 + 4\,t + 2, \quad t \in [0, 1].$$

- This is a problem that the Hamiltonian has no explicit dependency on time t, therefore, we expect that the Hamiltonian is a fixed value over $t \in [0, 1]$ (see Section 6.2),

$$\begin{aligned} H =& \frac{1}{2}u^2(t) - 2x(t) + p(t)(2 - u(t)) \\ =& \frac{1}{2}(2t - 2)^2 + 2\,t^2 - 8t - 4 - (2\,t - 2)^2 + 2(2\,t - 2) \\ =& -6, \qquad t \in [0, 1]. \end{aligned}$$

Sample Problem 2: Consider the second integrator system $\ddot{y} = u$ starting from rest at $y(0) = 10$. Find the optimal control that drives this system to origin to stop while minimizing the cost

$$J = \frac{1}{2}\alpha t_f^2 + \frac{1}{2}\int_0^{t_f} bu^2(t)\mathrm{d}t.$$

-Solution

- State space representation and boundary conditions: $x_1 = y$ and $x_2 = \dot{y}$ so that

$$\begin{aligned}&\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + Bu(t), \quad A = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}, \quad B = \begin{bmatrix}0\\1\end{bmatrix}.\\ &x_1(0) = 10, \quad x_2(0) = 0 \ \ \text{(start from rest)}\\ &x_1(t_f) = 0, \quad x_2(t_f) = 0 \ \ \text{(stops at the end)}\\ &\text{free final time}\end{aligned}$$

- Write $\mathbf{p}(t) = [p_1(t) \ \ p_2(t)]^\top$, to define the Hamiltonian:

$$H = g + \mathbf{p}^\top\mathbf{a} = \frac{1}{2}bu^2 + \mathbf{p}(t)^\top(A\mathbf{x}(t) + Bu(t))$$

- The necessary conditions for optimality

$$\begin{aligned}&\dot{\mathbf{p}} = -H_x^\top \rightarrow \begin{cases}\dot{p}_1 = -\frac{\partial H}{\partial x_1} = 0 \rightarrow p_1(t) = c_1 & t \in [0, t_f],\\ \dot{p}_2 = -\frac{\partial H}{\partial x_2} = 0 \rightarrow p_2(t) = c_1 t + c_2 & t \in [0, t_f],\end{cases}\\ &H_u = bu + p_2 = 0 \rightarrow u = -\frac{p_2}{b} = -\frac{c_2}{b} + \frac{c_1}{b}t\end{aligned}$$

- Free final time boundary condition:

$$\begin{aligned}\frac{\partial h}{\partial t}|_{t_f} + H(t_f) &= \alpha t_f + \frac{1}{2}bu(t_f)^2 + p_1(t_f)x_2(t_f) + p_2(t_f)u(t_f)\\ &= \alpha t_f + \frac{1}{2}bu(t_f)^2 + (-bu(t_f))u(t_f)\\ &= -\frac{1}{2}bu(t_f)^2 + \alpha t_f = 0 \rightarrow t_f = \frac{1}{2b\alpha}(-c_2 + c_1 t_f)^2.\end{aligned}$$

Here, we used $H_u = bu + p_2 = 0$ to write $p_2(t_f) = -bu(t_f)$, and also $x_2(t_f) = 0$.

- Substituting the optimal control $u = -\frac{c_2}{b} + \frac{c_1}{b}t$ in the state equations we get

$$\dot{x}_2(t) = -\frac{c_2}{b} + \frac{c_1}{b}t \rightarrow \quad x_2(t) = c_3 - \frac{c_2}{b}t + \frac{c_1}{2b}t^2 \underset{x_2(0)=0}{\rightarrow} c_3 = 0$$

$$\dot{x}_1(t) = -\frac{c_2}{b}t + \frac{c_1}{2b}t^2 \rightarrow \quad x_1(t) = c_4 - \frac{c_2}{2b}t^2 + \frac{c_1}{6b}t^3 \underset{x_1(0)=10}{\rightarrow} c_4 = 10$$

- Using the final boundary conditions we obtain

$$\begin{cases} x_2(t_f) = -\frac{c_2}{b}t_f + \frac{c_1}{2b}t_f^2 = 0 \\ x_1(t) = 10 - \frac{c_2}{2b}t_f^2 + \frac{c_1}{6b}t_f^3 = 0 \end{cases} \Rightarrow c_2 = \frac{60b}{t_f^2}, \quad c_1 = \frac{120b}{t_f^3}$$

- Substituting for c_1 and c_2 in the free final time boundary condition we obtain

$$t_f = \frac{1}{2b\alpha}(-c_2 + c_1 t_f)^2 = \frac{1}{2b\alpha}(-\frac{60b}{t_f^2} + \frac{120b}{t_f^3}t_f)^2 = \frac{(60b)^2}{2b\alpha t_f^4} \Rightarrow t_f = (1800\frac{b}{\alpha})^{\frac{1}{5}} \approx 4.48(\frac{b}{\alpha})^{\frac{1}{5}}.$$

- $t_f = 4.48(\frac{b}{\alpha})^{\frac{1}{5}}$ makes sense: α going up t_f goes down.
- $c_2 = 2.99b^{3/5}\alpha^{2/5}$ and $c_1 = 1.33b^{2/5}\alpha^{3/5}$: $u = -\frac{2.99b^{3/5}\alpha^{2/5}}{b} + \frac{1.33b^{2/5}\alpha^{3/5}}{b}t$

Next we solve this problem numerically via Matlab's

- The necessary conditions can be written as

$$\begin{cases} \dot{\mathbf{x}} = A\mathbf{x} + Bu, \\ \dot{\mathbf{p}} = -A'\mathbf{p}, \\ 0 = bu + [0 \ \ 1]\mathbf{p} \end{cases} \Rightarrow \begin{cases} \dot{\mathbf{x}} = A\mathbf{x} - \frac{1}{b}B[0 \ \ 1]\mathbf{p}, \\ \dot{\mathbf{p}} = -A'\mathbf{p}, \end{cases}$$

- Boundary conditions

$$\begin{aligned} &\mathbf{x}_1(0) = 10 \\ &\mathbf{x}_2(0) = 0 \\ &\mathbf{x}_1(t_f) = 0 \\ &\mathbf{x}_2(t_f) = 0 \\ &-0.5\, b\, u(t_f)^2 + \alpha\, t_f = 0 \end{aligned}$$

- Define the state of interest for the BVP4C as $\mathbf{z} = [\mathbf{x}^\top \ \mathbf{p}^\top \ r]^\top$ and note that

$$\frac{d\mathbf{z}}{d\tau} = \begin{bmatrix} \frac{d z_1}{d\tau} \\ \frac{d z_2}{d\tau} \\ \frac{d z_3}{d\tau} \\ \frac{d z_4}{d\tau} \\ \frac{d z_5}{d\tau} \end{bmatrix} = z_5 \left(\begin{bmatrix} A & -\frac{1}{b}B[0 \ \ 1] & 0 \\ 0 & -A^\top & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{z} \right) \Rightarrow \mathbf{z}' = f(\mathbf{z}) \ \ \text{(nonlinear function)}$$

with boundary conditions

$$\begin{aligned} &z_1(0) = 10, \\ &z_2(0) = 0, \\ &z_1(1) = 0, \\ &z_2(1) = 0, \\ &\frac{-0.5}{b} z_4(1)^2 + \alpha\, z_5(1) = 0. \end{aligned}$$

Notice at that in the boundary condition $-0.5\, b\, u(t_f)^2 + \alpha\, t_f = 0$, the explicit incidence of t_f is replaced by $z_5(1)$.

- The code to solve this problem is given next
  - the code gives the numerical solution and compares it to the closed-form analytic solutions

- Main code (courtesy of [6])

```
clc
clear all
b=0.1;
%alp=[.05 .1 1 10 20];
alp=logspace(-2,2,10);
t=[];
for alpha=alp
m=TPBVP(b,alpha);
t=[t;m(1,end)];
end

figure(1); clf
semilogx(alp,(1800*b./alp).^0.2,'-','Linewidth',2)
hold on; semilogx(alp,t,'rs'); hold off
xlabel('\alpha','FontSize',12); ylabel('t_f','FontSize',12)
legend('Analytic','Numerical')
title('Comparison with b=0.1')

% code from opt1.m on the analytic solution
b=0.1; alpha=0.1;
m=TPBVP(b,alpha);
tf=(1800*b/alpha)^0.2;
c1=120*b/tf^3;
c2=60*b/tf^2;
u=(-c2+c1*m(1,:))/b;
A=[0 1;0 0];B=[0 1]';C=eye(2);D=zeros(2,1);G=ss(A,B,C,D);X0=[10 0]';
[y3,t3]=lsim(G,u,m(1,:),X0);
figure(2); clf
subplot(211)
plot(m(1,:),u,'g-','LineWidth',2);
xlabel('Time','FontSize',12); ylabel('u(t)','FontSize',12)
hold on; plot(m(1,:),m(6,:),'--'); hold off
subplot(212)
plot(m(1,:),abs(u-m(6,:)),'-')
xlabel('Time','FontSize',12)
ylabel('u_{Analytic}(t)-U_{Numerical}','FontSize',12)
legend('Analytic','Numerical')
figure(3); clf
subplot(221)
plot(m(1,:),y3(:,1),'c-','LineWidth',2);
xlabel('Time','FontSize',12); ylabel('X(t)','FontSize',12)
hold on; plot(m(1,:),m([2],:),'k--'); hold off
legend('Analytic','Numerical')
subplot(222)
plot(m(1,:),y3(:,2),'c-','LineWidth',2);
xlabel('Time','FontSize',12); ylabel('dX(t)/dt','FontSize',12)
hold on; plot(m(1,:),m([3],:),'k--'); hold off
legend('Analytic','Numerical')
```

```
subplot(223)
plot(m(1,:),abs(y3(:,1)-m(2,:)'),'k-')
xlabel('Time','FontSize',12);ylabel('Error','FontSize',12)
subplot(224)
plot(m(1,:),abs(y3(:,2)-m(3,:)'),'k-')
xlabel('Time','FontSize',12);ylabel('Error','FontSize',12)
```

- Functions

  - TPBVPbc

```
function res=TPBVPbc(ya,yb)

global A B x0 b alp
res=[ya(1)-x0(1);ya(2)-x0(2);yb(1);yb(2);-0.5*yb(4)^2/b+alp*yb(5)];
```

  - TPBVPode

```
function dydt=TPBVPode(t,y)

global A B x0 b alp

dydt=y(5)*[ A -B*[0 1]/b zeros(2,1); zeros(2,2) -A' zeros(2,1);zeros(1,5)]*y;
```

  - TPBVPinit

```
function v=TPBVPinit(t)
global A B x0 b alp
v=[x0;1;0;1];
return
```

  - TPBVP

```
function m = TPBVP(p1,p2)

global A B x0 b alp;

A=[0 1;0 0];
B=[0 1]';
x0=[10 0]';
b=p1;
alp=p2;
solinit = bvpinit(linspace(0,1),@TPBVPinit);

sol = bvp4c(@TPBVPode,@TPBVPbc,solinit);

time = sol.y(5)*sol.x;

state = sol.y([1 2],:);
adjoint = sol.y([3 4],:);
control = -(1/b)*sol.y(4,:);
m(1,:) = time;
m([2 3],:) = state;
m([4 5],:) = adjoint;
m(6,:) = control;
```

6.2 Properties of the Hamiltonian

The following properties hold for optimal control problems with and without inequality constraints. Recall that the Hamiltonian is defined as

$$H = g(x(t), u(t), t)) + p(t)^\top a(x(t), u(t), t)$$

- If $g(x, u)$ and $a(x, u)$ do not explicitly depend on time t, then the Hamiltonian H is at least piece-wise contract.

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + (\frac{\partial H}{\partial x})\frac{dx}{dt} + (\frac{\partial H}{\partial p})\frac{dp}{dt} + (\frac{\partial H}{\partial u})\frac{du}{dt}$$
$$\frac{dH}{dt} = H_x\dot{x} + H_p\dot{p} + H_u\dot{u}$$

From F.O.N condition : $\dot{x} = H_p$ and $\dot{p} = -H_x$, then

$$\frac{dH}{dt} = H_u\dot{u}$$

- The third necessary condition is $H_u = 0$ so

$$\frac{dH}{dt} = (0)\dot{u} = 0$$

  which suggest H is constant

  - It might be possible for the value of this constant to change at a discontinuity of u, since then $\dot{u}$ would be infinite, and $0.\infty$ is not defined.
  - This H is at least piece-wise constant.

- For free final time problems, the transversality condition gives

$$h_t + H(t_f) = 0$$

- If h is not function of time then $h_t = 0$ and as a result $H(t_f) = 0$:

  - with no jumps in u, H is constant: $H(t) = 0$ for all $t \in [t_0.t_f]$.

6.3 Optimal control with inequality constraints

First consider an optimal control where the control inputs are constrained

$$\begin{aligned}
&u^{\star}(t) = \arg\min\ J = h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x(t), u(t), t)\mathrm{d}t, \quad s.t. \\
&\dot{x} = a(x(t), u(t), t), \quad x(t_0) = x_0, \\
&C(u, t) \leq 0.
\end{aligned}$$

Suppose initial condition and final time are specified but final state is free.

- If we define $H = g + p.a$, then according to what we have learned so far, we have

$$\delta J = \int_{t_0}^{t_f} (H_u \delta u)\mathrm{d}t, \quad \dot{p} = -H_x, \quad p(t_f) = \frac{\partial h}{\partial x}|_{t=t_f}.$$

- For $u(t)$ to be minimizing, we must have $\delta J \geq 0$ for all t and all admissible $\delta u(t)$.
- Hence, at all points on $C(u, t) = 0$, the optimal u has the property that

$$H_u \delta u \geq 0, \quad \delta C = C_u \delta u \leq 0.$$

- Another way of stating the condition above is that $\delta H = H_u\, \delta u$ must be non-improving over the set of possible $\delta u(t)$.
- A much stronger statement, " must be minimizing over the set of all possible ", is true; this compact statement is due to McShane (1939) and Pontryagin (1962) and is known as the "Minimum Principle."

Another approach to solve optimal control problems subject to inequality constraints (this time assume the more general case of $C(x, u, t) \leq 0$) is

- by studying minimizer of the augmented cost

$$J_a = h(x(t_f), t_f) + \int_{t_0}^{t_f} \big(g(x(t), u(t), t) + p(t).(a(x(t), u(t), t) - \dot{x}) + v(t).C(x(t), u(t), t)\big)\mathrm{d}t,$$

where

$$v_i(t) = \begin{cases} 0 & \text{if } C(x, u, t) < 0, \\ \geq 0 & \text{if } C(x, u, t) = 0, \end{cases}$$

so that $v_i C_i = 0$ for all $i \in \{1, \cdots, m\}$, where m is the number of inequality constraints. We note that if v is allowed to be negative, the J_a will decrease if the inequality constraint is violated. Non-negative v_i penalizes violation of inequality $C_i \leq 0$. When at any time t any of the inequality constraint C_i is strictly satisfied, i.e., $C_i < 0$, the inequality is inactive and we do not need to worry about whether $\delta u(t)$ is admissible or not. Therefore, for such cases we consider $v_i = 0$.

- Next, we find the variation of the augmented cost

$$J_a = h_x\delta x(t_f) + h_{t_f}\delta t_f + \int_{t_0}^{t_f}\Big((H_x + \dot p + v.C_x).\delta x + (H_u + v.C_u).\delta u + (H_p - \dot x).\delta + C.\delta v\Big)\mathrm{d}t,$$

- Define

$$H_a(x,u,p,v,t) = \underbrace{g + p(t).a}_{H(x,u,p,t)} + v(t).C,\ \text{where } v_i(t) = \begin{cases} 0 & \text{if } C(x,u,t) < 0,\\ \geq 0 & \text{if } C(x,u,t) = 0,\end{cases}$$

- Therefore, the necessary conditions for $\delta J_a = 0$ for $t \in [t_0, t_f]$ are

$$\begin{aligned}\dot x =& a(x,u,t),\\ \dot p =& -(H_a)_x,\\ 0 =& (H_a)_u,\end{aligned}$$

subject to appropriate boundary conditions and $v_i(t) = \begin{cases} 0 & \text{if } C(x,u,t) < 0,\\ \geq 0 & \text{if } C(x,u,t) = 0,\end{cases}$

$$u^\star = argmin\, J = -(x(4) - 1) + \frac{1}{2}\int_0^4 u^2(t)dt$$
$$\dot x = x + u, \qquad x(0) = 0$$
$$u(t) \leq 4$$

$$\boxed{\begin{aligned}&\dot x = x + u, \quad x(0) = 0\\ &\dot p = -p, \quad p(4) = \frac{\partial h}{\partial x}\Big|_{t=4} = -1 \ , \quad where\ H_a = \frac{1}{2}u^2 + p(t)(x+u) + v(t)(u-4), \quad v(t) = \begin{cases} 0 & u(t) - 4 < 0\\ \geq 0 & u(t) - 4 = 0\end{cases}\\ &0 = u(t) + p(t) + v(t)\end{aligned}}$$

$$\dot p = -p, \qquad p(4) = -1 \to p(t) = -e^{4-t} \qquad \boxed{u^\star = -(p + v)}$$

- Let us first assume if the inequality is inactive initially:
  $v = 0$, $u^\star = -p = e^{4-t}$, then $u^\star(0) = -p(0) = e^4 > 4$. Thus, the assumption **is not correct**, and we should proceed with inequality being active initially.

- The inequality is active initially: $u^\star(t) = 4$, $t \in [0, t_1]$
  $v(t) \geq 0$, $v(t) = -(u(t) + p(t)) = -4 + e^{4-t} \geq 0 \to t \leq 4 - \ln(4) \approx 2.61 = t_1$

Now for $t > t_1$, the inequality should be inactive (because for $t > t_1$, we have $v(t) < 0$)

$$v(t) = 0, \qquad u^\star = p = e^{4-t}, \qquad t \in [t_1, 4].$$

The optimal control is:

$$u^\star(t) = \begin{cases} 4 & t \in [0, t_1]\\ e^{4-t} & t \in (t_1, 4]\end{cases} \to x(t) = \begin{cases} 4e^t - 4 & t \in [0, t_1]\\ -\frac{1}{2}e^{4-t} + (4 - 8e^{-4})e^t & t \in (t_1, 4]\end{cases}$$

6.3.1 Sample examples

Example 1: Use first order necessary conditions to characterize a candidate optimal solution for the problem below

$$\begin{aligned}\text{minimize } J = -\int_0^2 (2x-3u)dt, \quad s.t.,\\ \dot{x} = x+u,\\ x(0)=4, \quad x(2)=\text{free},\\ 0 \le u \le 2.\end{aligned}$$

Solution:

- Hamiltonian for this problem is

$$H = -(2x-3u) + p(x+u)$$

  Then, we obtain

$$\begin{cases}\dot{p} = -H_x = 2-p,\\ p(t_f) = p(2) = \frac{\partial h^{\text{terminal cost}}}{\partial x}|_{t_f} = 0\end{cases} \Rightarrow p(t) = -2\,e^{2-t}+2, \quad t\in[0,2].$$

- To obtain optimal control, we use PMP

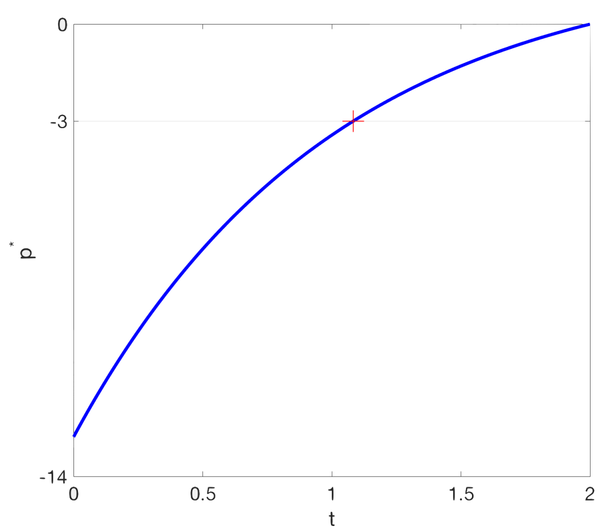
$$\begin{cases}u^\star = \text{argmin } H(x^\star,u,p^\star,t) = \text{argmin } (p^\star+3)\,u,\\ 0\le u\le 2.\end{cases}$$

  which gives us

$$u^\star = \begin{cases}2 & p^\star+3<0,\\ ? & p^\star+3=0,\\ 0 & p^\star+3>0\end{cases}$$

- Is singular arc possible, i.e., $p^\star+3=0$ for some interval of time in $[0,2]$?

  To answer, we start by checking the trajectory of $p(t)$, $t\in[0,2]$



Notice that $p(0) = -12.778$ and $p(2) = 0$, therefore, the trajectory will go through $p(t_1) = -3$ at some time $t_1 \in (0, 2)$. Therefore, singular arc is not possible in this problem, and we can write

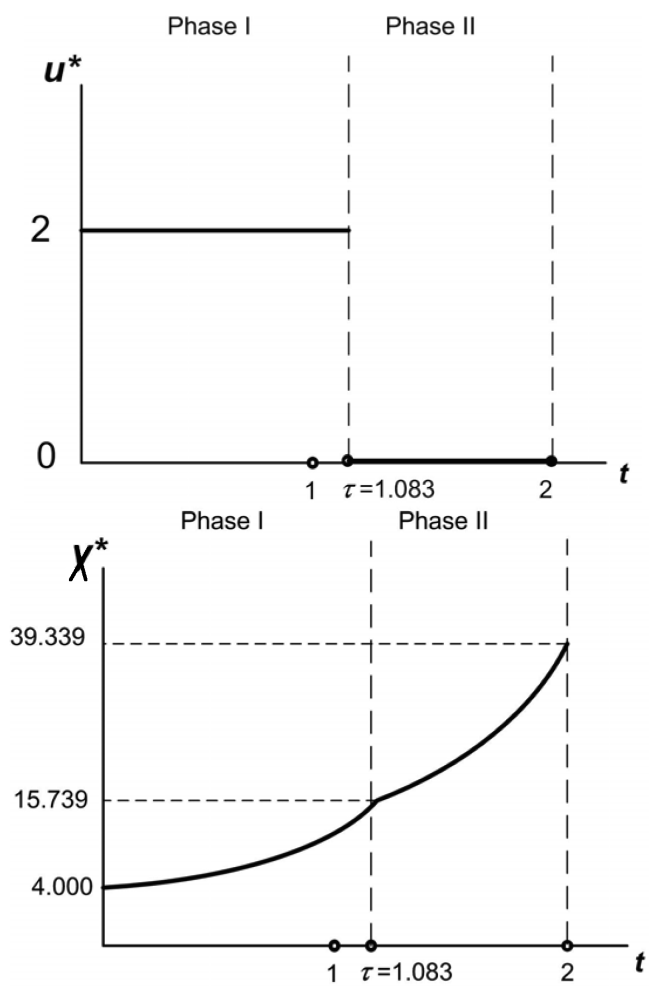
$$u^\star = \begin{cases} 2 & p^\star < -3, \\ ? & p^\star = -3, \\ 0 & p^\star > -3 \end{cases} \quad \Rightarrow \quad u^\star = \begin{cases} 2 & t \in [0, t_1], \\ 0 & t \in (t_1, 2]. \end{cases}$$

In finalizing the optimal control we use $p(0) = -12.7778 < -3$ to decide that $u^\star(t) = 2$ for $t \in [0, t_1]$. The terminal condition $p(2) = 0 > -3$ also confirms that $u^\star(t) = 0$ for $t \in (t_1, 2]$.

- Therefore, the state trajectory is obtained from

$$\begin{cases} \dot{x} = x + 2, & t \in [0, t_1], \\ \dot{x} = x + 0, & t \in (t_1, 2]. \end{cases}$$
$$x(0) = 4$$

- Results: switching time $t_1 = 2 - \ln(5/2) = 1.0837$, the control history and state trajectory are shown in the plots below



Example 2: bang-bang optimal control: Consider the second order integrator dynamics

$$\ddot{y} = u, \quad |u(t)| < 1.$$

Our objective is to drive this system from a given initial condition $y(0), \dot{y}(0) \in \mathbb{R}$ to $y(t_f) = \dot{y}(t_f) = 0$, in minimum time, i.e, find a control that minimizes $J = \int_0^{t_f} 1\mathrm{d}t$.
Solution: System dynamics: $x_1 = y$ and $x_2 = \dot{y}$: $\dot{x}_1 = x_2, \quad \dot{x}_2 = u$.
Drive this system from a given initial condition $x_1(0), x_2(0) \in \mathbb{R}$ to $x_1(t_f) = x_2(t_f) = 0$, in minimum time

observation:

- if $u = 1$ for all $t \in \mathbb{R}_{\geq 0}$, the system response is

$$x_2(t) = t + c_1 \rightarrow x_2(t) = (t + x_2(0)),$$
$$x_1(t) = \frac{1}{2}(t + x_2(0))^2 + c_2 \rightarrow x_1(t) = \frac{1}{2}(t + x_2(0))^2 + x_1(0) - \frac{1}{2}x_2(0)^2,$$

therefore: $x_1(t) = \frac{1}{2}x_2(t)^2 + (x_1(0) - \frac{1}{2}x_2(0)^2)$

- if $u = -1$ for all $t \in \mathbb{R}_{\geq 0}$, the system response is

$$x_2(t) = -t + c_1 \rightarrow x_2(t) = -(t - x_2(0)),$$
$$x_1(t) = -\frac{1}{2}(t - x_2(0))^2 + c_2 \rightarrow x_1(t) = -\frac{1}{2}(t - x_2(0))^2 + x_1(0) + \frac{1}{2}x_2(0)^2,$$

therefore: $x_1(t) = -\frac{1}{2}x_2(t)^2 + (x_1(0) + \frac{1}{2}x_2(0)^2)$

Figure 9 shows the possible trajectories of the system based on given initial conditions and the control that is used.

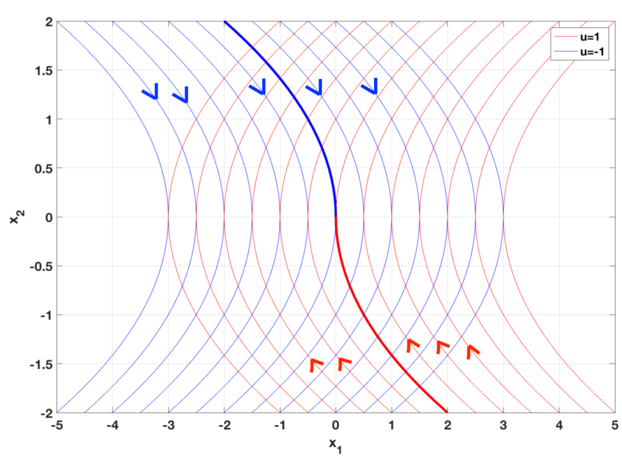


Figure 9: Possible response curves. Arrows show direction of the motion on the curves.

- Hamiltonian: $H = 1 + p_1 x_2 + p_2 u$
- Co-state equations:

$$\begin{cases} \dot{p}_1 = -H_{x_1} = 0 & \rightarrow \quad p_1^\star(t) = c_1, \\ \dot{p}_2 = -H_{x_2} = -p_1 & \rightarrow \quad p_2^\star(t) = -c_1 t + c_2. \end{cases} \tag{6.3.1}$$

- Optimal control (PMP): $u^\star = \underset{-1 \le u \le 1}{\operatorname{argmin}}(1 + p_1^\star x_2^\star + p_2^\star u) = \underset{-1 \le u \le 1}{\operatorname{argmin}}(p_2^\star u)$, which gives

$$u^\star(t) = \begin{cases} -1 & p_2^\star > 0 \\ ? & p_2^\star = 0, \\ 1 & p_2^\star < 0 \end{cases}$$

So the control depends on $p_2^\star$, but because $p_2(t) = -c_1 t + c_2$ is a linear function of time, $p_2(t)$ can only be zero at $t = t_1$ and switching can happen only at one point in time $t = t_1$: bang-bang control. We show below that $p_2(t)$ cannot be zero over any finite interval of time $[\bar{t}_1, \bar{t}_2] \subset [0, t_f]$ (singular arc is not possible in this problem. Therefore, $u^\star(t) = \begin{cases} -1 & p_2^\star > 0 \\ 0 & p_2^\star = 0, \\ 1 & p_2^\star < 0. \end{cases}$

  - Since $\dot{x}_2 = u$, and since we must stop at t_f, then we must have $u = \pm 1$ at t_f.

- To complete the solution, next we impose the boundary conditions: transversality condition along with $x_2(t_f) = 0$ gives

$$H(t_f) + h_t(t_f) = 0 \rightarrow 1 + p_1(t_f)x_2(t_f) + p_2(t_f)u(t_f) = 0 \rightarrow 1 + p_2^\star(t_f)u^\star(t_f) = 0 \rightarrow$$

$$p_2^\star(t_f) = -\frac{1}{u^\star(t_f)}.$$

  - If $u^\star(t_f) = 1$ then we obtain $p_2^\star(t_f) = -1 < 0$, which is consistent with the selection rule for optimal control.
  - If $u^\star(t_f) = -1$ then we obtain $p_2^\star(t_f) = 1 > 0$, which is consistent with the selection rule for optimal control.

  So, the terminal conditions do not help us determine if $u = 1$ or $u = -1$ at t_f because both are possible.

- To solve the problem, we are going to consider a scenario and see if the solution validates our scenario
  - First, let us look at the case where $u(t_f) = 1$, which implies that $p_2(t_f) = -1$. Let us also assume that $c_1 > 0$, which gives us a possible switching case similar to that shown in Fig. 10.

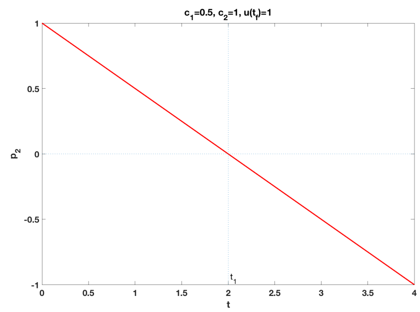


Figure 10: Possible switching case, but t_f and c_1 are unknown at this point.

- Then setting $p_2^\star(t_1) = c_2 - c_1 t_1 = 0$ and $p_2^\star(t_f) = c_2 - c_1 t_f = -1$ gives $t_1 = t_f - \frac{1}{c_1}$.
- Now look at the state response
  * During $(t_1, t_f]$ we have $u(t) = 1$, therefore using the boundary condition $x_1(t_f) = x_2(t_f) = 0$, we obtain

$$\dot{x}_2 = 1 \to x_2(t) = t + c_3 \to x_2(t) = t - t_f$$

$$\dot{x}_1 = x_2 = t - t_f \to x_1(t) = \frac{1}{2}(t - t_f)^2 + c_4 \to x_1(t) = \frac{1}{2}(t - t_f)^2.$$

    Therefore, during $(t_1, t_f]$, the states are on $\boxed{\mathrm{x}_1(t) = \frac{x_2(t)^2}{2}}$ in the phase plane. For the scenario that we have considered $(u(t_f) = 1)$, the response is in lower quadrant of $x_1 - x_2$ plane (note here that from $x_1(t) = \frac{x_2(t)^2}{2}$ you know $x_1 \geq 0$ and from $x_2(t) = t - t_f$ you know that $x_2(t) \leq 0$ during $(t_1, t_f]$).
  * During $[0, t_1]$ we have $u(t) = -1$, therefore using the given boundary conditions $x_1(0), x_2(t_0)$, we obtain

$$\dot{x}_2 = -1 \to x_2(t) = -t + c_5 \to x_2(t) = -t + x_2(0)$$

$$\dot{x}_1 = x_2 = -(t + x_2(0)) \to x_1(t) = -\frac{1}{2}(t - x_2(0))^2 + c_6 \to x_1(t) = -\frac{1}{2}(t - x_2(0))^2 + x_1(0) + \frac{1}{2}x_2(0)^2.$$

    Therefore, during $[0, t_1]$, the states are on $\boxed{\mathrm{x}_1(t) = -\frac{x_2(t)^2}{2} + x_1(0) + \frac{1}{2}x_2(0)^2}$ in the phase plane (Do you see how this curve correlates with the trajectories on Fig. 9?)
- We can validate our scenario by calculating t_1, t_f and c_1 and checking if our expectations of $0 < t_1 < t_f$ and $c_1 > 0$ are correct. If the scenario is valid, we can show that $t_1 = x_2(0) + \sqrt{x_1(0) + 0.5x_2(0)^2}$ and $t_f = x_2(0) + 2\sqrt{x_1(0) + 0.5x_2(0)^2}$.

  Lets consider two examples:

∗ First case is $x_1(0) = 4$ and $x_2(0) = 0$. The state trajectory at t_1 is continuous therefore,

$$\begin{cases} x_2(t_1) = t_1 - t_f, \\ x_2(t_1) = -t_1 + x_2(0) = -t_1, \end{cases} \quad \rightarrow t_1 = \frac{1}{2}t_f$$

$$\begin{cases} x_1(t_1) = \frac{1}{2}(t_1 - t_f)^2, \\ x_1(t_1) = -\frac{1}{2}(t_1 - x_2(0))^2 + x_1(0) + \frac{1}{2}x_2(0)^2 = -\frac{1}{2}t_1^2 + 4, \end{cases} \quad \rightarrow \frac{1}{2}(t_1 - t_f)^2 = -\frac{1}{2}t_1^2 + 4.$$

Therefore, $t_1 = 2$ and $t_f = 4$. Next, we compute c_1 from $t_1 = t_f - \frac{1}{c_1}$, which gives $c_1 = 0.5$. Therefore, this scenario is valid scenario. That is

$$u^\star = \begin{cases} -1 & t \in [0, 2], \\ 1 & t \in (2, 4]. \end{cases}$$

On the phase plane, the system starts on the curve $x_1(t) = -\frac{x_2(t)^2}{2} + 4$ and moves on it until it hits $x_1(t) = \frac{x_2(t)^2}{2}$ at t_1. Then switches to move on curve $x_1(t) = \frac{x_2(t)^2}{2}$ until it arrives at the origin. You can think of $x_1(t) = \frac{x_2(t)^2}{2}$ as the switching curve, see Fig. 11.

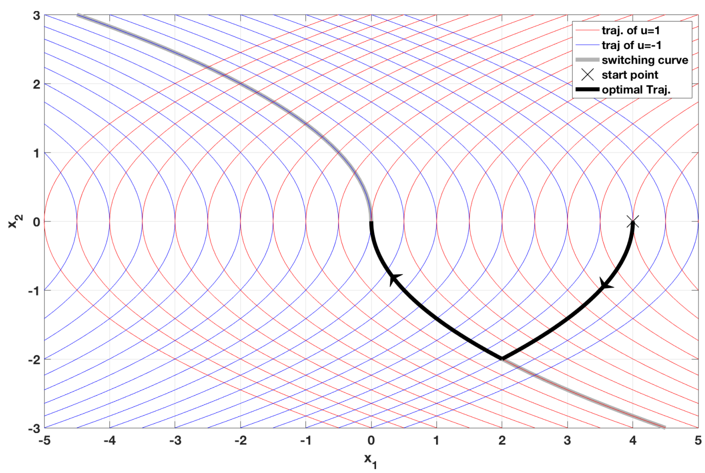


Figure 11: Optimal trajectory of the case of $x_1(0) = 4$ and $x_2(0) = 0$.

∗ Next case is defined by $x_1(0) = -4$ and $x_2(0) = 0$. The state trajectory at t_1 is continuous therefore,

$$\begin{cases} x_2(t_1) = t_1 - t_f, \\ x_2(t_1) = -t_1 + x_2(0) = -t_1, \end{cases} \quad \rightarrow t_1 = \frac{1}{2}t_f$$

$$\begin{cases} x_1(t_1) = \frac{1}{2}(t_1 - t_f)^2, \\ x_1(t_1) = -\frac{1}{2}(t_1 - x_2(0))^2 + x_1(0) + \frac{1}{2}x_2(0)^2 = -\frac{1}{2}t_1^2 - 4, \end{cases} \quad \rightarrow \frac{1}{2}(t_1 - t_f)^2 = -\frac{1}{2}t_1^2 - 4.$$

Here, we get $\frac{t_f}{2}^2 = -4$, which shows that this scenario is not valid. Therefore, for this case, we need to consider $u(t_f) = -1$ and repeat all the processes above to find the valid optimal solution.

- Next, let us look at the case where $u(t_f) = -1$, which implies that $p_2(t_f) = 1$, and also $c_1 < 0$, which gives us a possible switching case similar to that shown in Fig. 12.

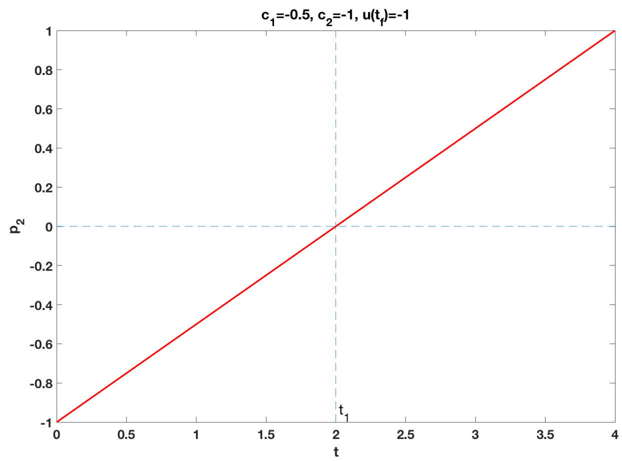


Figure 12: Possible switching case, but t_f and c_1 are unknown at this point.

- Then setting $p_2(t_1) = c_2 - c_1 t_1 = 0$ and $p_2(t_f) = c_2 - c_1 t_f = 1$ gives $t_1 = t_f + \frac{1}{c_1}$.
- Now look at the state response
  - During $(t_1, t_f]$ we have $u(t) = -1$, therefore using the boundary condition $x_1(t_f) = x_2(t_f) = 0$, we obtain

$$\dot{x}_2 = -1 \rightarrow x_2(t) = -t + c_3 \rightarrow x_2(t) = -(t - t_f)$$

$$\dot{x}_1 = x_2 = -(t - t_f) \rightarrow x_1(t) = -\frac{1}{2}(t - t_f)^2 + c_4 \rightarrow x_1(t) = -\frac{1}{2}(t - t_f)^2.$$

    Therefore, during $(t_1, t_f]$, the states are on $\boxed{x_1(t) = -\frac{x_2(t)^2}{2}}$ in the phase plane. For the scenario that we have considered ($u(t_f) = -1$), the response is in upper left quadrant of $x_1 - x_2$ plane (note here that from $x_1(t) = -\frac{x_2(t)^2}{2}$ you know $x_1 \leq 0$ and from $x_2(t) = -(t - t_f)$ you know that $x_2(t) \geq 0$ during $(t_1, t_f]$).

  - During $[0, t_1]$ we have $u(t) = 1$, therefore using the given boundary conditions $x_1(0), x_2(t_0)$, we obtain

$$\dot{x}_2 = 1 \rightarrow x_2(t) = t + c_5 \rightarrow x_2(t) = t + x_2(0)$$

$$\dot{x}_1 = x_2 = (t + x_2(0)) \rightarrow x_1(t) = \frac{1}{2}(t + x_2(0))^2 + c_6 \rightarrow x_1(t) = -\frac{1}{2}(t + x_2(0))^2 + x_1(0) - \frac{1}{2}x_2(0)^2.$$

    Therefore, during $[t_0, t_1]$, the states are on $\boxed{x_1(t) = -\frac{x_2(t)^2}{2} + x_1(0) - \frac{1}{2}x_2(0)^2}$ in the phase plane (Do you see how this curve correlates with the trajectories on Fig. 9?)

- We can validate our scenario by calculating t_1, t_2, t_f and c_1 and checking if our expectations of $0 \leq t_1 \leq t_2 \leq t_f$ and $c_1 > 0$ are correct. If the scenario is valid, we can show that $t_1 = x_2(0) + \sqrt{x_1(0) + 0.5x_2(0)^2}$ and $t_f = x_2(0) + 2\sqrt{x_1(0) + 0.5x_2(0)^2}$.

  Lets consider two examples:

  - First case is $x_1(0) = 4$ and $x_2(0) = 1$. The state trajectory at t_1 is continuous therefore,

$$\begin{cases} x_2(t_1) = -(t_1 - t_f), \\ x_2(t_1) = t_1 + x_2(0) = t_1, \end{cases} \rightarrow t_1 = \frac{1}{2}t_f$$
$$\begin{cases} x_1(t_1) = -\frac{1}{2}(t_1 - t_f)^2, \\ x_1(t_1) = \frac{1}{2}(t_1 + x_2(0))^2 + x_1(0) - \frac{1}{2}x_2(0)^2 = \frac{1}{2}t_1^2 + 4, \end{cases} \rightarrow -\frac{1}{2}(t_1 - t_f)^2 = \frac{1}{2}t_1^2 + 4.$$

    Then, we obtain $-\frac{1}{2}(\frac{t_f}{2})^2 = \frac{1}{2}(\frac{t_f}{2})^2 + 4$, which shows that this case is not a valid scenario. Therefore, for this case, we need to consider $u(t_f) = 1$ and repeat all the processes above to find the valid optimal solution (as we saw earlier).

  - Next case is defined by $x_1(0) = -4$ and $x_2(0) = 0$. The state trajectory at t_1 is continuous therefore,

$$\begin{cases} x_2(t_1) = -(t_1 - t_f), \\ x_2(t_1) = t_1 + x_2(0) = t_1, \end{cases} \rightarrow t_1 = \frac{1}{2}t_f$$
$$\begin{cases} x_1(t_1) = -\frac{1}{2}(t_1 - t_f)^2, \\ x_1(t_1) = \frac{1}{2}(t_1 + x_2(0))^2 + x_1(0) - \frac{1}{2}x_2(0)^2 = \frac{1}{2}t_1^2 - 4, \end{cases} \rightarrow -\frac{1}{2}(t_1 - t_f)^2 = \frac{1}{2}t_1^2 - 4.$$

    Therefore, $t_1 = 2$ and $t_f = 4$. Next, we compute c_1 from $t_1 = t_f + \frac{1}{c_1}$, which gives $c_1 = -0.5$. Therefore, this scenario is a valid scenario. That is

$$u^\star = \begin{cases} 1 & t \in [0, 2], \\ -1 & t \in (2, 4]. \end{cases}$$

    On the phase plane, the system starts on the curve $x_1(t) = \frac{x_2(t)^2}{2} - 4$ and moves on it until it hits $x_1(t) = -\frac{x_2(t)^2}{2}$ at t_1. Then it switches to move on curve $x_1(t) = -\frac{x_2(t)^2}{2}$ until it arrives at the origin. You can think of $x_1(t) = -\frac{x_2(t)^2}{2}$ as the switching curve, see Fig. 13.

- **State-based switching law**: based on all the discussions above, we can conclude that here, the optimal control law may be written as

$$\boxed{u^\star(t) = \begin{cases} +1 & \text{if } x_2^2\,\mathrm{sgn}(x_2) < -2x_1 \text{ or } \big(x_2^2\,\mathrm{sgn}(x_2) = -2x_1, \quad x_1 > 0\big), \\ -1 & \text{if } x_2^2\,\mathrm{sgn}(x_2) > -2x_1 \text{ or } \big(x_2^2\,\mathrm{sgn}(x_2) = -2x_1, \quad x_1 < 0\big), \end{cases}}$$

Figure 13: Optimal trajectory of the case of $x_1(0) = -4$ and $x_2(0) = 0$.

- **Ruling out singular intervals**: recall that from the PMP we concluded that $u^\star = \underset{-1\leq u\leq 1}{\text{argmin}}(1+p_1x_2 + p_2u) = \underset{-1\leq u\leq 1}{\text{argmin}}(p_2u)$, which gives

$$u^\star(t) = \begin{cases} -1 & p_2 > 0 \\ ? & p_2 = 0. \\ 1 & p_2 < 0 \end{cases}$$

We show next that $P_2(t)$ cannot be zero over any finite interval of time $[\bar{t}_1, \bar{t}_2] \subset [0, t_f]$. Recall the co-state equations in (6.3.1). We note that for $p_2^\star(t)$ to be zero over $[\bar{t}_1, \bar{t}_2]$, it is necessary that

$$c_1 = 0, \quad c_2 = 0,$$

which means that $p_1^\star(t) = 0$ and also $p_2^\star(t) = 0$ for $t \in [0, t_f]$. Substituting in the Hamiltonian, we obtain

$$H(x^\star(t), p^\star(t), u^\star(t)) = 1, \quad \forall t \in [0, t_f]. \tag{6.3.2}$$

But since the final time is free and H is explicitly independent of time, (6.3.2) violates the necessary condition that (review the properties of the Hamiltonian in Section 6.2)

$$H(x^\star(t), p^\star(t), u^\star(t)) = 0, \quad \forall t \in [0, t_f]. \tag{6.3.3}$$

We conclude that $p_2^\star(t)$ cannot be zero over any finite interval of time and, thus, the singular interval cannot exists.

Example 3: bang-off-bang optimal control: Consider the second order integrator dynamics

$$\ddot{y} = u, \quad |u(t)| < 1.$$

Our objective is to drive this system from a given initial condition $y(0), \dot{y}(0) \in \mathbb{R}$ to $y(t_f) = \dot{y}(t_f) = 0$, in minimum time, i.e, find a control that minimizes

$$J = \int_0^{t_f} (1 + b|u|)\mathrm{d}t, \quad b > 0.$$

Solution: System dynamics: $x_1 = y$ and $x_2 = \dot{y}$: $\dot{x}_1 = x_2, \quad \dot{x}_2 = u$.
Drive this system from a given initial condition $x_1(0), x_2(0) \in \mathbb{R}$ to $x_1(t_f) = x_2(t_f) = 0$, in minimum time.

- Hamiltonian: $H = 1 + b\,|u| + p_1\,x_2 + p_2\,u$
- Co-state equations:

$$\begin{cases} \dot{p}_1 = -H_{x_1} = 0 & \rightarrow \quad p_1^\star(t) = c_1, \\ \dot{p}_2 = -H_{x_2} = -p_1 & \rightarrow \quad p_2^\star(t) = -c_1 t + c_2. \end{cases} \tag{6.3.4}$$

- Optimal control (PMP): $u^\star = \underset{-1\leq u\leq 1}{\operatorname{argmin}}(1 + b\,|u| + p_1^\star\,x_2^\star + p_2^\star\,u) = \underset{-1\leq u\leq 1}{\operatorname{argmin}} \begin{cases} bu + p_2^\star\,u & u \geq 0, \\ -bu + p_2^\star\,u & u \leq 0, \end{cases}$, which gives

$$u^\star(t) = \begin{cases} -1 & p_2^\star > b \\ 0 & -b < p_2^\star < b, \\ 1 & p_2^\star < -b, \\ \text{undetermined, but } \geq 0 & p_2^\star = -b, \\ \text{undetermined, but } \leq 0 & p_2^\star = b. \end{cases}$$

- **Ruling out singular intervals**: we show that $p_2(t)$ cannot be $\pm b$ over any finite interval of time $[\bar{t}_1, \bar{t}_2] \subset [0, t_f]$. Recall the co-state equations in (6.3.4). We note that for $p_2^\star(t)$ to be $\pm b$ over $[\bar{t}_1, \bar{t}_2]$, it is necessary that

$$c_1 = 0,$$

and

$$c_2 = \pm b,$$

which means that $p_1^\star(t) = 0$ and also $p_2^\star(t) = \pm b$ for $t \in [0, t_f]$. Substituting in the Hamiltonian, we obtain

$$H(x^\star(t), p^\star(t), u^\star(t)) = 1, \quad \forall t \in [0, t_f]. \tag{6.3.5}$$

But since the final time is free and H is explicitly independent of time, (6.3.5) violates the necessary condition that (review the properties of the Hamiltonian in Section 6.2)

$$H(x^\star(t), p^\star(t), u^\star(t)) = 0, \quad \forall t \in [0, t_f]. \tag{6.3.6}$$

We conclude that $p_2^\star(t)$ cannot be $\pm b$ over any finite interval of time and, thus, the singular interval cannot exists.

- Next, observe that the optimal control cannot include $u^\star(t) = 0$ for $[\bar{t}, t_f]$, for $\bar{t} \geq 0$. That is optimal control cannot end with an interval of time of $u = 0$ because the system does not move to the origin with no control applied.

- Give the observations above, and since the control depends on $p_2^\star = -c_1 t + c_2$, which is a linear function of time, The possibilities for optimal control depending on the initial condition of the problem are

  – **Case 1**: $u^\star(t) = \begin{cases} -1 & t \in [0, t_1), \\ 0 & t \in [t_1, t_2), \\ 1 & t \in [t_2, t_f], \end{cases}$ for $0 < t_1 < t_2 < t_f$. Here, $u^\star(t_f) = 1$, which implies that $p_2^\star(t_f) = -(1 + b)$ and $c_1 > 0$, with a possible switching case similar to that shown in Fig. 14.

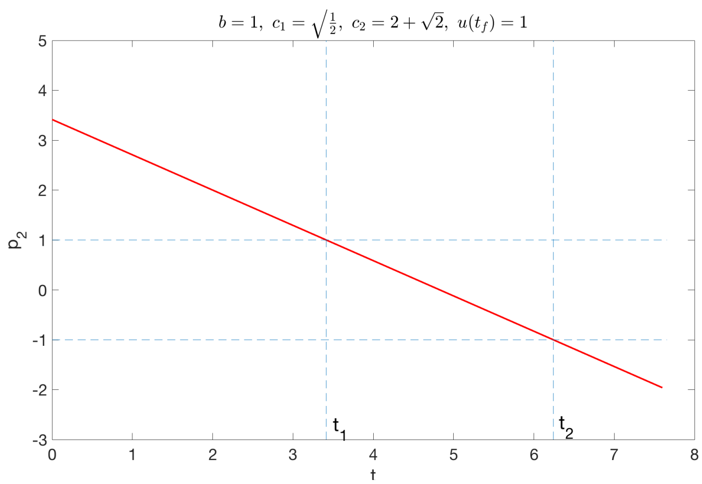


Figure 14: Possible switching case, but t_f and c_1 are unknown at this point (Case 1).

  – **Case 2**: $u^\star(t) = \begin{cases} 0 & t \in [0, t_1), \\ 1 & t \in [t_1, t_f], \end{cases}$ for $0 < t_1 < t_f$. Here, $u^\star(t_f) = 1$, which implies that $p_2^\star(t_f) = -(1 + b)$ and $c_1 > 0$.

  – **Case 3**: $u^\star(t) = \begin{cases} 1 & t \in [0, t_1), \\ 0 & t \in [t_1, t_2), \\ -1 & t \in [t_2, t_f], \end{cases}$ for $0 < t_1 < t_2 < t_f$. Here, $u^\star(t_f) = -1$, which implies that $p_2^\star(t_f) = (1 + b)$ and $c_1 < 0$ with a possible switching case similar to that shown in Fig. 15.

  – **Case 4**:$u^\star(t) = \begin{cases} 0 & t \in [0, t_1), \\ -1 & t \in [t_1, t_f], \end{cases}$ for $0 < t_1 < t_f$. Here, $u^\star(t_f) = -1$, which implies that $p_2^\star(t_f) = b + 1$ and $c_1 < 0$, with a possible switching case similar to that shown in Fig. 16.

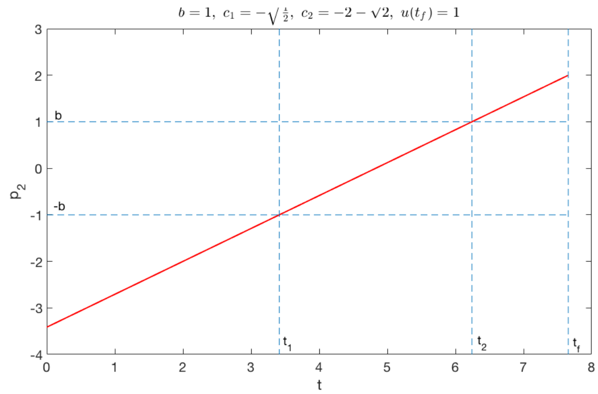


Figure 15: Possible switching case, but t_f and c_1 are unknown at this point (Case 3).

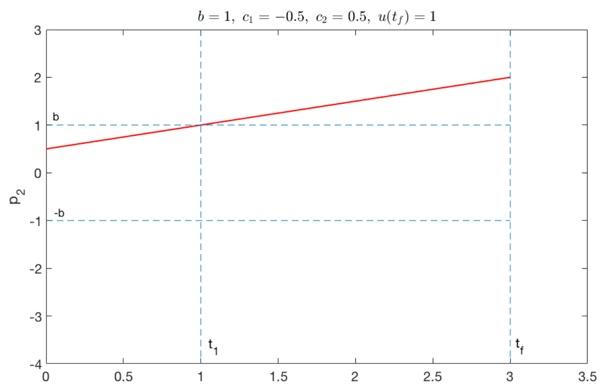


Figure 16: Possible switching case, but t_f and c_1 are unknown at this point (Case 4).

- To complete the solution, next we impose the boundary conditions: transversality condition along with $x_2(t_f)=0$ gives

$$H(t_f)+h_t(t_f)=0\rightarrow 1+b|u^\star(t_f)|+p_1^\star(t_f)x_2^\star(t_f)+p_2^\star(t_f)u^\star(t_f)=0\rightarrow$$
$$1+b|u^\star(t_f)|+p_2^\star(t_f)u^\star(t_f)=0\rightarrow p_2^\star(t_f)=-(b\,\mathrm{sgn}(u^\star(t_f)+\frac{1}{u^\star(t_f)}).$$

  - If $u^\star(t_f)=1$ then we obtain $p_2^\star(t_f)=-(1+b)<-b$, which is consistent with the selection rule for optimal control.
  - If $u^\star(t_f)=-1$ then we obtain $p_2^\star(t_f)=-(-1-b)>b$, which is consistent with the selection rule for optimal control.

So, the terminal conditions do not help us determine if $u^\star=1$ or $u^\star=-1$ at t_f because both are possible.

- To solve the problem, we are going to consider the 4 cases listed above.

$$- - - - -$$

  - Case 1: The optimal co-state in this case follows the form shown in Fig. 14 and

$$t_0 < t_1 < t_2 < t_f.$$

  - Setting $p_2^\star(t_2) = c_2 - c_1\, t_2 = -b$ and $p_2^\star(t_f) = c_2 - c_1\, t_f = -b-1$ gives $t_2 = t_f - \frac{1}{c_1}$. And, setting $p_2^\star(t_1) = c_2 - c_1 t_1 = b$ gives $t_1 = t_f - \frac{(2b+1)}{c_1}$.
  - Now look at the state response
    * During $(t_2, t_f]$ we have $u(t) = 1$, therefore using the boundary condition $x_1(t_f) = x_2(t_f) = 0$, we obtain

$$\dot{x}_2 = 1 \rightarrow x_2(t) = t + c_3 \rightarrow x_2(t) = t - t_f$$

$$\dot{x}_1 = x_2 = t - t_f \rightarrow x_1(t) = \frac{1}{2}(t-t_f)^2 + c_4 \rightarrow x_1(t) = \frac{1}{2}(t-t_f)^2.$$

    Therefore, during $(t_2, t_f]$, the states are on $\boxed{\mathrm{x}_1(t) = \frac{x_2(t)^2}{2}}$ in the phase plane. For the scenario that we have considered ($u(t_f) = 1$), the response is in lower quadrant of $x_1 - x_2$ plane (note here that from $x_1(t) = \frac{x_2(t)^2}{2}$ you know $x_1 \geq 0$ and from $x_2(t) = t - t_f$ you know that $x_2(t) \leq 0$ during $(t_2, t_f]$).
    * Between times $t_1 - t_2$, control input is zero: **coasting phase**.
      · Because of the continuity of the state, terminal condition for coast phase can be obtained from

$$x_1(t_2) = \frac{1}{2}(t_2 - t_f)^2 = \frac{1}{2}((t_f - \frac{1}{c_1}) - t_f)^2 = \frac{1}{2c_1^2},$$

$$x_2(t_2) = t_2 - t_f = (t_f - \frac{1}{c_1}) - t_f = -\frac{1}{c_1}.$$

      · On a coasting arc $x_2 = \dot{y}$ is a constant (so $x_2(t) = -\frac{1}{c_1}$ for $t \in [t_1, t_2]$), and thus

$$\text{for } t \in (t_1, t_2]:\ \dot{x}_1 = x_2 = -\frac{1}{c_1} \rightarrow x_1(t) = -\frac{1}{c_1}t + c_5,$$

then using $x_1(t_2) = \frac{1}{2c_1^2}$, we obtain $c_5 = \frac{1}{c_1}t_2 + \frac{1}{2c_1^2}$, which gives,

$$x_1(t) = -\frac{1}{c_1}t + \frac{1}{c_1}t_2 + \frac{1}{2c_1^2} = \frac{1}{c_1}(t_2 - t) + \frac{1}{2c_1^2}, \text{ and as a result}$$

$$x_1(t_1) = \frac{1}{c_1}(t_2 - t_1) + \frac{1}{2c_1^2}.$$

      Then,

$$\begin{aligned} x_1(t_1) =& \frac{1}{c_1}(t_2 - t_1) + \frac{1}{2c_1^2} = \frac{1}{c_1}(t_f - \frac{1}{c_1} - t_f + \frac{(2b+1)}{c_1}) + \frac{1}{2c_1^2} = (2b + \frac{1}{2})(\frac{1}{c_1^2}) \\ =& (2b + \frac{1}{2})x_2(t_1)^2. \end{aligned}$$

* So the first transition from coasting phase to $u = -1$ at t_1 occurs on the curve $\boxed{x_1(t) = (2b + \frac{1}{2})x_2^2.}$
* During $[0, t_1]$ we have $u(t) = -1$, therefore using the given boundary conditions $x_1(0), x_2(t_0)$, we obtain

$$\dot{x}_2 = -1 \rightarrow x_2(t) = -t + c_6 \rightarrow x_2(t) = -t + x_2(0)$$

$$\dot{x}_1 = x_2 = -(t + x_2(0)) \rightarrow x_1(t) = -\frac{1}{2}(t - x_2(0))^2 + c_7 \rightarrow x_1(t) = -\frac{1}{2}(t - x_2(0))^2 + x_1(0) + \frac{1}{2}x_2(0)^2.$$

Therefore, during $[t_0, t_1]$, the states are on $\boxed{\mathrm{x}_1(t) = -\frac{x_2(t)^2}{2} + x_1(0) + \frac{1}{2}x_2(0)^2}$ in the phase plane (Do you see how this curve correlates with the trajectories on Fig. 9?)

- We can validate our scenario by calculating t_1, t_2, t_f and c_1 and checking if our expectations of $0 < t_1 < t_2 < t_f$ and $c_1 > 0$ are correct.

Lets consider two examples with $b = 1$:

* First case is $x_1(0) = 4$ and $x_2(0) = 2$. The state trajectory at t_1 is continuous therefore,

$$\begin{cases} x_2(t_1) = -\frac{1}{c_1}, \\ x_2(t_1) = -t_1 + x_2(0) = -t_1 + 2, \end{cases} \rightarrow t_1 = 2 + \frac{1}{c_1}$$
$$\begin{cases} x_1(t_1) = (2b + \frac{1}{2})(\frac{1}{c_1^2}) = (2 + \frac{1}{2})(\frac{1}{c_1^2}), \\ x_1(t_1) = -\frac{1}{2}(t_1 - x_2(0))^2 + x_1(0) + \frac{1}{2}x_2(0)^2 = -\frac{1}{2}(t_1 - 2)^2 + 6, \end{cases}$$
$$\rightarrow (2 + \frac{1}{2})(\frac{1}{c_1^2}) = -\frac{1}{2}(t_1 - 2)^2 + 6.$$

The state trajectory at t_2 is also continuous therefore,

$$\begin{cases} x_2(t_2) = -\frac{1}{c_1}, \\ x_2(t_2) = t_2 - t_f, \end{cases} \rightarrow t_2 = t_f - \frac{1}{c_1}$$
$$\begin{cases} x_1(t_2) = \frac{1}{2c_1^2} \\ x_1(t_2) = \frac{1}{2}(t_2 - t_f)^2, \end{cases} \rightarrow \frac{1}{c_1^2} = (t_2 - t_f)^2.$$

Also recall that from the co-state equations we obtained

$$t_2 = t_f - \frac{1}{c_1}, \quad t_1 = t_f - \frac{(2b+1)}{c_1}.$$

Therefore, $c_1 = \sqrt{\frac{1}{2}}$, $t_1 = 2 + \sqrt{2}$, $t_2 = 2 + 3\sqrt{2}$ and $t_f = 2 + 4\sqrt{2}$. Therefore, this scenario is a valid scenario. That is

$$u^\star = \begin{cases} -1 & t \in \left[0, 2 + \sqrt{2}\right), \\ 0 & t \in \left[2 + \sqrt{2}, 2 + 3\sqrt{2}\right), \\ 1 & t \in \left[2 + 3\sqrt{2}, 2 + 4\sqrt{\frac{5}{3}}\right]. \end{cases}$$

See Fig. 14 for $p_2^\star$ vs. time. On the phase plane, the system starts on the curve $x_1(t) = -\frac{x_2(t)^2}{2} + 6$ and moves on it until it hits $x_1(t) = (2b + \frac{1}{2})x_2^2$ at t_2. Then, it coast on constant x_2 until it hits $x_1(t) = \frac{x_2(t)^2}{2}$ and switches to it and moves along it until it arrives at the origin, see Fig. 17.

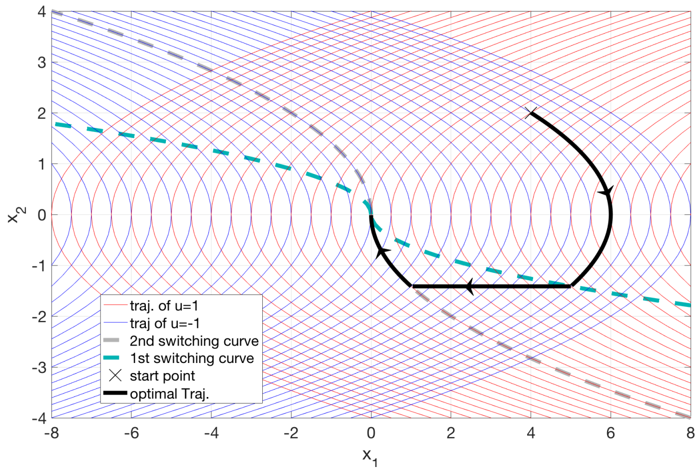


Figure 17: Optimal trajectory when $x_1(0) = 4$ and $x_2(0) = 2$ ($b = 1$). optimal control is follows the form given in Case 1.

$*$ Next case is defined by $x_1(0) = -4$ and $x_2(0) = 2$. The state trajectory at t_1 is continuous therefore,

$$\begin{cases} x_2(t_1) = -\frac{1}{c_1}, \\ x_2(t_1) = -t_1 + x_2(0) = -t_1 + 2, \end{cases} \quad \rightarrow t_1 = 2 + \frac{1}{c_1}$$

$$\begin{cases} x_1(t_1) = (2b + \frac{1}{2})(\frac{1}{c_1^2}) = (2 + \frac{1}{2})(\frac{1}{c_1^2}), \\ x_1(t_1) = -\frac{1}{2}(t_1 - x_2(0))^2 + x_1(0) + \frac{1}{2}x_2(0)^2 = -\frac{1}{2}(t_1 - 2)^2 - 2, \end{cases}$$

$$\rightarrow (2 + \frac{1}{2})(\frac{1}{c_1^2}) = -\frac{1}{2}(t_1 - 2)^2 - 2.$$

Here, we get $c_1^2 = -\frac{3}{2}$, which shows that this scenario is not valid. Therefore, for this case, we need to consider the other remaining 3 cases for the optimal control.

– – – – –

- Case 3: The optimal co-state in this case follows the form shown in Fig. 15 and

$$t_0 < t_1 < t_2 < t_f.$$

- Setting $p_2^\star(t_2) = c_2 - c_1\, t_2 = b$ and $p_2^\star(t_f) = c_2 - c_1\, t_f = b + 1$ gives $t_2 = t_f + \frac{1}{c_1}$. And, setting $p_2^\star(t_1) = c_2 - c_1 t_1 = -b$ gives $t_1 = t_f + \frac{(2b+1)}{c_1}$.

– Now look at the state response

* During $(t_2, t_f]$ we have $u(t) = -1$, therefore using the boundary condition $x_1(t_f) = x_2(t_f) = 0$, we obtain

$$\dot{x}_2 = 1 \rightarrow x_2(t) = -t + c_3 \rightarrow x_2(t) = -(t - t_f)$$

$$\dot{x}_1 = x_2 = -(t - t_{f \rightarrow} x_1(t) = -\frac{1}{2}(t - t_f)^2 + c_4 \rightarrow x_1(t) = -\frac{1}{2}(t - t_f)^2.$$

Therefore, during $(t_2, t_f]$, the states are on $\boxed{\mathrm{x}_1(t) = -\frac{x_2(t)^2}{2}}$ in the phase plane. For the scenario that we have considered ($u(t_f) = -1$), the response is in the upper-left quadrant of $x_1 - x_2$ plane (note here that from $x_1(t) = -\frac{x_2(t)^2}{2}$ you know $x_1 \leq 0$ and from $x_2(t) = -(t - t_f)$ you know that $x_2(t) \geq 0$ during $(t_2, t_f]$).

* Between times $t_1 - t_2$, control input is zero: **coasting phase**.

  · Because of the continuity of the state, terminal condition for coast phase can be obtained from

$$x_1(t_2) = -\frac{1}{2}(t_2 - t_f)^2 = -\frac{1}{2}((t_f + \frac{1}{c_1}) - t_f)^2 = -\frac{1}{2c_1^2},$$
$$x_2(t_2) = -(t_2 - t_f) = -(t_f + \frac{1}{c_1}) - t_f = -\frac{1}{c_1}.$$

  · On a coasting arc $x_2 = \dot{y}$ is a constant (so $x_2(t) = -\frac{1}{c_1}$ for $t \in [t_1, t_2]$), and thus

$$\text{for } t \in (t_1, t_2]: \ \dot{x}_1 = x_2 = -\frac{1}{c_1} \rightarrow x_1(t) = -\frac{1}{c_1}t + c_5,$$

then using $x_1(t_2) = -\frac{1}{2c_1^2}$, we obtain $c_5 = \frac{1}{c_1}t_2 - \frac{1}{2c_1^2}$, which gives,

$$x_1(t) = -\frac{1}{c_1}t + \frac{1}{c_1}t_2 - \frac{1}{2c_1^2} = \frac{1}{c_1}(t_2 - t) - \frac{1}{2c_1^2}, \text{ and as a result}$$
$$x_1(t_1) = \frac{1}{c_1}(t_2 - t_1) - \frac{1}{2c_1^2}.$$

Then,

$$\begin{aligned} x_1(t_1) =& \frac{1}{c_1}(t_2 - t_1) - \frac{1}{2c_1^2} = \frac{1}{c_1}(t_f + \frac{1}{c_1} - t_f - \frac{(2b+1)}{c_1}) - \frac{1}{2c_1^2} = -(2b + \frac{1}{2})(\frac{1}{c_1^2}) \\ =& -(2b + \frac{1}{2})x_2(t_1)^2. \end{aligned}$$

* So the first transition from coasting phase to $u = 1$ at t_1 occurs on the curve $\boxed{x_1(t) = -(2b + \frac{1}{2})x_2^2.}$

* During $[0, t_1]$ we have $u(t) = 1$, therefore using the given boundary conditions $x_1(0), x_2(0)$, we obtain

$$\dot{x}_2 = 1 \rightarrow x_2(t) = t + c_6 \rightarrow x_2(t) = t + x_2(0)$$

$$\dot{x}_1 = x_2 = (t + x_2(0)) \rightarrow x_1(t) = \frac{1}{2}(t + x_2(0))^2 + c_7 \rightarrow x_1(t) = \frac{1}{2}(t + x_2(0))^2 + x_1(0) - \frac{1}{2}x_2(0)^2.$$

Therefore, during $[t_0, t_1]$, the states are on $\boxed{x_1(t) = \frac{x_2(t)^2}{2} + x_1(0) - \frac{1}{2}x_2(0)^2}$ in the phase plane (Do you see how this curve correlates with the trajectories on Fig. 9?)

– We can validate our scenario by calculating t_1, t_2, t_f and c_1 and checking if our expectations of $0 < t_1 < t_2 < t_f$ and $c_1 < 0$ are correct.

Lets consider two examples with $b = 1$:

∗ First case is $x_1(0) = -4$ and $x_2(0) = -2$. The state trajectory at t_1 is continuous therefore,

$$\begin{cases} x_2(t_1) = -\frac{1}{c_1}, \\ x_2(t_1) = t_1 + x_2(0) = t_1 - 2, \end{cases} \rightarrow t_1 = 2 - \frac{1}{c_1}$$

$$\begin{cases} x_1(t_1) = -(2b + \frac{1}{2})(\frac{1}{c_1^2}) = -(2 + \frac{1}{2})(\frac{1}{c_1^2}), \\ x_1(t_1) = \frac{1}{2}(t_1 + x_2(0))^2 + x_1(0) - \frac{1}{2}x_2(0)^2 = \frac{1}{2}(t_1 - 2)^2 - 6, \end{cases}$$

$$\rightarrow -(2 + \frac{1}{2})(\frac{1}{c_1^2}) = \frac{1}{2}(t_1 - 2)^2 - 6.$$

The state trajectory at t_2 is also continuous therefore,

$$\begin{cases} x_2(t_2) = -\frac{1}{c_1}, \\ x_2(t_2) = -(t_2 - t_f), \end{cases} \rightarrow t_2 = t_f + \frac{1}{c_1}$$

$$\begin{cases} x_1(t_2) = -\frac{1}{2c_1^2} \\ x_1(t_2) = -\frac{1}{2}(t_2 - t_f)^2, \end{cases} \rightarrow \frac{1}{c_1^2} = -(t_2 - t_f)^2.$$

Also recall that from the co-state equations we obtained

$$t_2 = t_f + \frac{1}{c_1}, \quad t_1 = t_f + \frac{(2b + 1)}{c_1}.$$

Therefore, $c_1 = -\sqrt{\frac{1}{2}}$, $t_1 = 2 + \sqrt{2}$, $t_2 = 2 + 3\sqrt{2}$ and $t_f = 2 + 4\sqrt{2}$. Therefore, this scenario is a valid scenario. That is

$$u^\star = \begin{cases} 1 & t \in \left[0, 2 + \sqrt{2}\right), \\ 0 & t \in \left[2 + \sqrt{2}, 2 + 3\sqrt{2}\right), \\ -1 & t \in \left[2 + 3\sqrt{2}, 2 + 4\sqrt{\frac{5}{3}}\right]. \end{cases}$$

See Fig. 15 for $p_2^\star$ vs. time. On the phase plane, the system starts on the curve $x_1(t) = \frac{x_2(t)^2}{2} - 6$ and moves on it until it hits $x_1(t) = -(2b + \frac{1}{2})x_2^2$ at t_2. Then, it coast on constant x_2 until it hits $x_1(t) = -\frac{x_2(t)^2}{2}$ and switches to it and moves along it until it arrives at the origin, see Fig. 18.

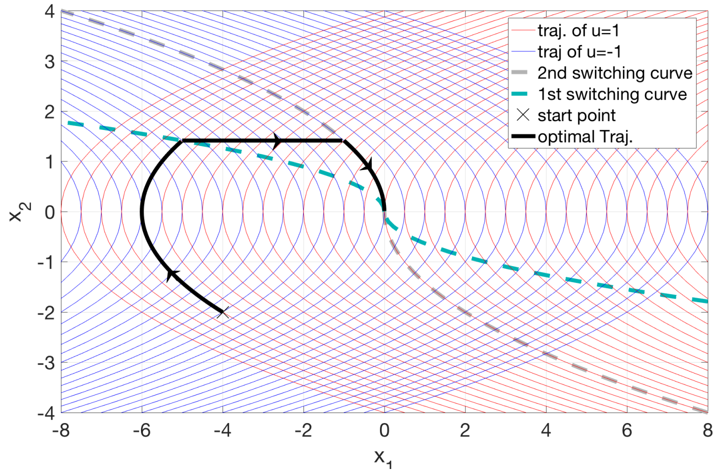


Figure 18: Optimal trajectory of the case of $x_1(0) = 4$ and $x_2(0) = 2$ ($b = 1$). Optimal control has the form given in Case 3.

* Next case is defined by $x_1(0) = -4$ and $x_2(0) = 2$. The state trajectory at t_1 is continuous therefore,

$$\begin{cases} x_2(t_1) = -\frac{1}{c_1}, \\ x_2(t_1) = t_1 + x_2(0) = t_1 + 2, \end{cases} \quad \to t_1 = -2 - \frac{1}{c_1}$$

$$\begin{cases} x_1(t_1) = -(2b + \frac{1}{2})(\frac{1}{c_1^2}) = -(2 + \frac{1}{2})(\frac{1}{c_1^2}), \\ x_1(t_1) = \frac{1}{2}(t_1 + x_2(0))^2 + x_1(0) - \frac{1}{2}x_2(0)^2 = \frac{1}{2}(t_1 + 2)^2 - 6, \end{cases}$$

$$\to -(2 + \frac{1}{2})(\frac{1}{c_1^2}) = \frac{1}{2}(t_1 + 2)^2 - 6.$$

Therefore, $c_1^2 = \frac{1}{2}$, which results in $t_1 < 0$. Therefore, the optimal control is not in the form given in Case 3. We have already seen that the optimal control is not in the form of Case 1 either. Therefore, we have to investigate Case 2 and Case 4.

− − − − −

- Case 4: The optimal co-state in this case follows the form shown in Fig. 16 and

$$t_0 < t_1 < t_f.$$

- Setting $p_2^\star(t_1) = c_2 - c_1\, t_1 = b$ and $p_2^\star(t_f) = c_2 - c_1\, t_f = b + 1$ gives $t_1 = t_f + \frac{1}{c_1}$.
- Now look at the state response
  * During $[t_1, t_f]$ we have $u(t) = -1$, therefore using the boundary condition $x_1(t_f) = x_2(t_f) = 0$, we obtain

$$\dot{x}_2 = 1 \to x_2(t) = -t + c_3 \to x_2(t) = -(t - t_f)$$

$$\dot{x}_1 = x_2 = -(t - t_{f\rightarrow} x_1(t) = -\frac{1}{2}(t-t_f)^2 + c_4 \rightarrow x_1(t) = -\frac{1}{2}(t-t_f)^2.$$

Therefore, during $(t_1, t_f]$, the states are on $\boxed{\mathrm{x}_1(t) = -\frac{x_2(t)^2}{2}}$ in the phase plane. For the scenario that we have considered ($u(t_f) = -1$), the response is in the upper-left quadrant of $x_1 - x_2$ plane (note here that from $x_1(t) = -\frac{x_2(t)^2}{2}$ you know $x_1 \leq 0$ and from $x_2(t) = -(t - t_f)$ you know that $x_2(t) \geq 0$ during $(t_1, t_f]$).

* Between times $0 - t_1$, control input is zero: **coasting phase**.
  - Because of the continuity of the state, terminal condition for coast phase can be obtained from

$$x_1(t_1) = -\frac{1}{2}(t_1 - t_f)^2 = -\frac{1}{2}((t_f + \frac{1}{c_1}) - t_f)^2 = -\frac{1}{2c_1^2},$$
$$x_2(t_1) = -(t_1 - t_f) = -(t_f + \frac{1}{c_1}) - t_f = -\frac{1}{c_1}.$$

  - On a coasting arc $x_2 = \dot{y}$ is a constant (so $x_2(t) = -\frac{1}{c_1}$ for $t \in [0, t_1]$), and thus

$$\text{for } t \in [0, t_1): \ \dot{x}_1 = x_2 = -\frac{1}{c_1} \rightarrow x_1(t) = -\frac{1}{c_1}t + c_5,$$

then using $x_1(t_1) = -\frac{1}{2c_1^2}$, we obtain $c_5 = \frac{1}{c_1}t_1 - \frac{1}{2c_1^2}$, which gives,

$$x_1(t) = -\frac{1}{c_1}t + \frac{1}{c_1}t_1 - \frac{1}{2c_1^2} = \frac{1}{c_1}(t_1 - t) - \frac{1}{2c_1^2}.$$

Then, using the boundary conditions $x_1(0)$ and $x_2(0)$, during $[0, t_1)$ we have

$$x_2(0) = -\frac{1}{c_1} \rightarrow c_1 = -\frac{1}{x_2(0)}.$$
$$x_1(0) = \frac{1}{c_1}(t_1) - \frac{1}{2c_1^2} = -x_2(0)\, t_1 - 0.5x_2(0)^2 \rightarrow t_1 = -\frac{1}{x_2(0)}(x_1(0) + 0.5x_2(0)^2).$$

Then, from $t_1 = t_f + \frac{1}{c_1}$, we obtain

$$t_f = -\frac{1}{x_2(0)}(x_1(0) + 0.5x_2(0)^2) + x_2(0)$$

. Finally using $p_2^\star(t_f) = c_2 - c_1\, t_f = b + 1$, we obtain $c_2 = b + \frac{1}{x_2(0)^2}(x_1(0) + 0.5x_2(0)^2)$.

* We can validate our scenario by calculating t_1, t_f and c_1 and checking if our expectations of $0 < t_1 < t_f$ and $c_1 < 0$ are correct. Obviously, here we expect

$$c_1 = -\frac{1}{x_2(0)} \rightarrow x_2(0) > 0,$$
$$t_1 > 0 \rightarrow x_1(0) < -0.5x_2(0)^2.$$

Lets consider an example with $b = 1$:

· Let $x_1(0) = -4$ and $x_2(0) = 2$. Recall that Case 1 and Case 3 were not valid for the given initial condition. Let us assume that the optimal control follows Case 4. then:

$$
\begin{aligned}
c_1 &= -\frac{1}{x_2(0)} = -0.5 < 0, \\
c_2 &= b + \frac{1}{x_2(0)^2}(x_1(0) + 0.5x_2(0)^2) = 1 + 0.25(-4+2) = 0.5, \\
t_1 &= -\frac{1}{x_2(0)}(x_1(0) + 0.5x_2(0)^2) = -0.5(-2) = 1 \\
t_f &= -\frac{1}{x_2(0)}(x_1(0) + 0.5x_2(0)^2) + x_2(0) = -0.5(-2) + 2 = 3.
\end{aligned}
$$

All the parameters are consistent, therefore, Case 4 is the valid scenario for this example. See Fig. 16 for $p_2^\star$ vs. time. On the phase plane, the system starts in the coasting phase and moves until it hits $x_1(t) = -\frac{x_2(t)^2}{2}$ and switches to it and moves along it until it arrives at the origin, see Fig. 19.

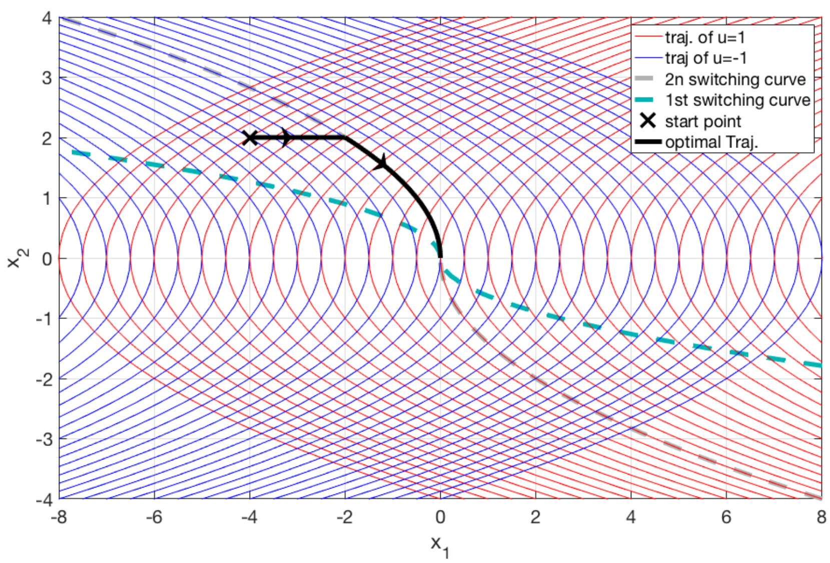


Figure 19: Optimal trajectory of the case of $x_1(0) = -4$ and $x_2(0) = 2$ ($b = 1$). Optimal control has the form given in Case 4.

∗ Can you investigate Case 2 yourself?

– **State-based switching law**: $x_1(t) = \begin{cases} (2b+\frac{1}{2})x_2^2 & x_2 \leq 0 \\ -(2b+\frac{1}{2})x_2^2 & x_2 0 \end{cases}$ is the first switching curve, and $x_1(t) = \begin{cases} \frac{x_2(t)^2}{2} & x_2 \leq 0 \\ -\frac{x_2(t)^2}{2} & x_2 \geq 0 \end{cases}$ is the second switching curve. Can you write the state dependent switching law?

Note 7

Appendix

7.1 Math Lab

- $\dot{x} = a\,x + b, \quad \Rightarrow \begin{cases} x(t) = x(t_1)\,e^{a(t-t_1)} - \frac{b}{a}(1 - e^{a(t-t_1)}), & t \in [t_1, t] \\ x(t) = x(t_2)\,e^{a(t_2-t)} + \frac{b}{a}(e^{a(t_2-t_1)} - 1), & t \in [t, t_2] \end{cases}$

  For this, note: $e^{-at}\dot{x} - a\,e^{-at}x = b\,e^{-at} \Rightarrow \frac{d}{dt}(e^{-at}\,x) = b\,e^{-at} \Rightarrow (e^{-at}\,x)_{t_1}^{t_2} = -\frac{b}{a}\,(e^{-at})_{t_1}^{t_2} \Rightarrow$ $e^{-a\,t_2}\,x(t_2) - e^{-a\,t_1}\,x(t_1) = -\frac{b}{a}\,(e^{-a\,t_2} - e^{-a\,t_1})$

Bibliography

[1] D. E. Kirk, "Optimal control theory: an introduction", Dover Publications, Inc., Mineola, New York, 2004, ISBN: 9780486434841 (originally published in 1970 by Prentice-Hall, Inc.)

[2] D. Liberzon, "Calculus of variations and optimal control theory: a concise introduction," Princeton University Press, Princeton, NJ, 2012.

[3] F. L. Lewis, D. l. Arabie, and V. L. Syrmos, "Optimal Control", 3rd Ed.,Wiley, 2012.

[4] D. G. Luenberger, Y. Ye, "Linear and nonlinear programming," 3rd Ed., Springer, New York, NY, 2008.

[5] A. E. Bryson Jr., and Y. C. Ho, "Applied Optimal Control," Hemisphere, New York, NY, 1975.

[6] https://ocw.mit.edu/courses/aeronautics-and-astronautics/16-323-principles-of-optimal-control-spring-2008/lecture-notes/lec7.pdf