Optimal control SF2852, KTH

Optimal Control

Lecture notes by

Ulf Jönsson

in collaboration with

Claes Trygger and Petter Ögren

2010

Optimization and Systems Theory

ROYAL INSTITUTE OF TECHNOLOGY

SE-100 44 STOCKHOLM, SWEDEN

Contents

Preface

The first version of these lecture notes were written in the winter quarter of 2001. They have been continuously extended and improved since then. Some material in Chapter 3 and Chapter 2 are adopted from previous lecture notes in Swedish by C. Trygger and P. Ögren. We are grateful for many comments and suggestions from A. Blomqvist, A. Hansson, R. Nagamune and V. Nieto.

Chapter 1

Introduction

Optimal control is an important topic for several reasons. For example,

- Optimal control is one of the most useful systematic methods for controller design. It has several advantages
  - Most control problems are extremely hard to solve using ad-hoc techniques and engineering intuition. Optimal control gives a systematic approach to the solution of control problems.
  - There are normally many possible solutions to a control problem. Some are good, others are poor. Optimal control reduces this redundancy by selecting a controller that is best according to some cost criterium.
  - Nature behaves optimally and it is natural to pose many engineering tasks as optimal control problems.
- Applications for optimal control abound in engineering, science, and economics. We will in the course consider examples from
  - Economics and logistics
  - Aeronautical systems
  - Autonomous systems and robotics
  - Biomathematics
- Optimal control is an important branch of mathematics
  - The field of optimal control has its roots in the calculus of variations developed by such giants as Bernoulli, Euler, Lagrange, Weierstrass and others.
  - Optimal control as an independent field emerged in the 1950s during the space race. We will in the course learn the main results which were developed by some of the most well known systems theorists.

    * Dynamic programming (Richard Bellman)
    * Pontryagin minimum principle
    * Linear quadratic control (Rudolph Kalman)
  - Optimal control is still a vital research field with many directions.

1.1 Examples of Optimal Control Problems

We will here introduce three examples that will be solved during the course. The examples are taken from rather different application areas: Economics, robotics, and data interpolation. They illustrate our claim that optimal control is a versatile systematic tool for engineering design which has wide applicability. Other examples discussed during the course are from areas such as aeronautics and biology.

Example 1 (Optimal storage strategy). A producer of a commodity wants to find an optimal storage strategy. Let us use the notation

x stock size

u production rate which is constrained to the interval $0 \leq u \leq M$

and assume that the storage is empty at time $t = 0$ and that all produced commodity will be stored. It follows that the stock grows according to the differential equation

$$\dot{x}(t) = u(t), \quad x(0) = 0.$$

The goal is to have at least A units of commodity in storage at some prescribed time t_f. This goal should be obtained at minimum cost. If we let

r be production cost growth rate

c storage cost per time unit

then the total cost becomes

$$cost = \int_0^{t_f} (u(t)e^{rt} + cx(t))dt$$

It is not optimal to have $x(t_f) > A$. Indeed, once $x(t) = A$, we can only increase the cost for larger times since $x(t)$ is positive, increasing, and appears with positive weight in the cost integral and u is also positive and appears with positive weight in the cost integral. In other words, we can replace the inequality $x(t_f) \geq A$ by the equality $x(t_f) = A$. This implies that our optimal control problem is

$$\text{minimize} \int_0^{t_f} (u(t)e^{rt} + cx(t))dt \quad \text{subj. to} \quad \begin{cases} \dot{x}(t) = u(t), & 0 \leq u(t) \leq M \\ x(0) = 0, \; x(t_f) = A. \end{cases}$$

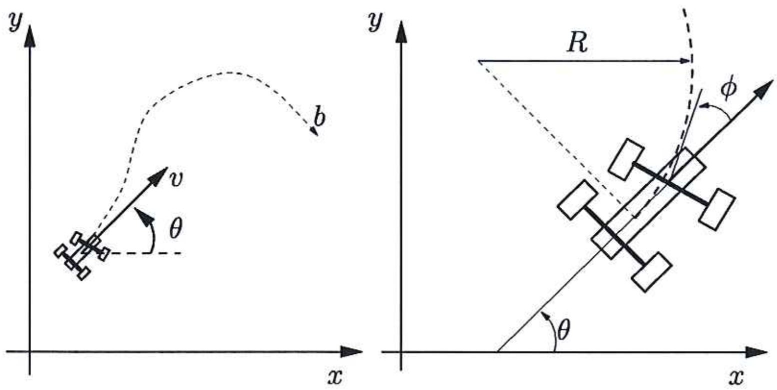

Figure 1.1: A mobile robot with constant speed v and a heading θ. The problem is to find a control law for the heading θ that corresponds to the shortest path to point b. The turning radius is bounded below by R, as is illustrated in the right hand side of the figure.

Example 2 (Dubins' car)**.** We consider the problem of steering a mobile robot from a point a to a point b as illustrated in the left hand side of Figure 1.1.

The kinematics of the robot is described by the differential equations

$$\begin{aligned} \dot{x} &= v\cos(\theta) \\ \dot{y} &= v\sin(\theta) \\ \dot{\theta} &= \omega \end{aligned} \tag{1.1}$$

Here (x, y) are the coordinates of the reference point of the car, θ is the direction (heading) of the car (i.e., the angle between the robots main axis and the positive x-axis), and v is the constant speed. The turning radius has the lower bound R, which means that the angular velocity is bounded as

$$|\omega| \leq v/R.$$

To understand why there must be a bound on the turning radius we notice that even if we can turn the steering wheel (i.e., change ϕ) instantaneously, this does not mean that θ changes instantaneously since the back wheels cannot slip on the surface.

Our aim is to find the shortest possible path from the initial point to a point b. In mathematical language this means that we wish to minimize

$$\int_0^T (\dot{x}^2 + \dot{y}^2)^{1/2} dt = vT$$

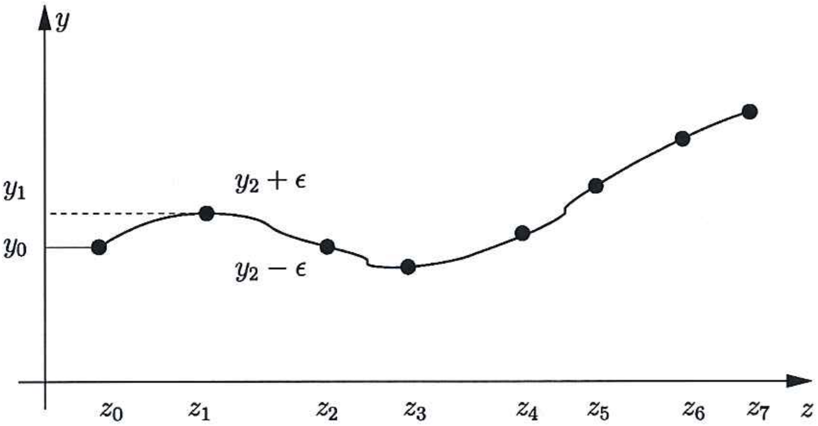

Figure 1.2: The problem is to find a smooth curve that starts at (z_0, y_0) and then passes through the the points (z_k, y_k) before it ends at (z_N, y_N).

subject to the kinematic equations (1.1) and the initial and terminal constraints

$$\begin{bmatrix} x(0) \\ y(0) \\ \theta(0) \end{bmatrix} = \begin{bmatrix} a_x \\ a_y \\ a_\theta \end{bmatrix}, \qquad \begin{bmatrix} x(T) \\ y(T) \\ \theta(T) \end{bmatrix} = \begin{bmatrix} b_x \\ b_y \\ b_\theta \end{bmatrix}$$

Note that the assumption of constant speed turns our shortest path problem into a time optimal control problem.

Example 3 (Data interpolation). We are given a set of points in the plane and the problem is to find a smooth curve that interpolates these points. Let us assume that the horizontal coordinates are ordered as $z_0 < z_1 < z_2 < \ldots < z_N$ and that the corresponding vertical coordinates must interpolate given points y_k, $k = 0, \ldots, N$. This is illustrated in Figure 1.2.

One way to obtain resonable smoothness properties is to aim at making the curvature small. The square of the curvature is given by

$$\kappa^2 = \frac{y''(z)^2}{(1 + y'(z)^2)^3} \leq y''(z)^2$$

It appears reasonable to try to minimize the second derivate, which motivates us to formulate the following optimization problem: Find a twice differentiable function $y(z)$ such that $y(z_k) = y_k$, $k = 0, \ldots, N$, for $k = 1, \ldots N$, and such that $\int_{z_0}^{z_N} y''(z)^2 dz$ is minimized.

The first step in addressing this problem is to formulate it in state space form. Let the "time axis" correspond to z and in particular let $t_k := z_k$. Furthermore,

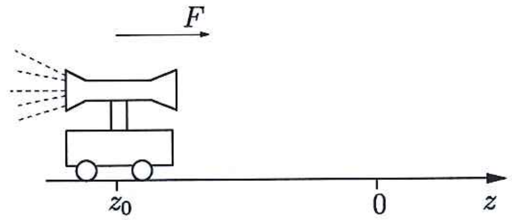

Figure 1.3: Control problem: Move the rocket car to rest at $z = 0$.

define the state vector to be $x = (x_1, x_2) = (y(z), y'(z))$, and finally define the control as $u = y''(z)$. Then the state space representation becomes

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t), \quad \begin{bmatrix} x_1(t_0) \\ x_2(t_0) \end{bmatrix} = \begin{bmatrix} y_0 \\ x_{20} \end{bmatrix}$$
$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

where x_{20} is a parameter we need to choose. If we define

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix},$$

then our optimal control problem can be formulated as

$$\text{minimize} \int_{t_0}^{t_N} u(t)^2 dt \quad \text{subj to} \quad \begin{cases} \dot{x} = Ax + Bu, \\ Cx(t_k) = y_k, \; k = 1, \dots, N \end{cases}$$

This problem and many generalizations have been solved in [19, 26, 7]. The resulting optimal curve will in our example be a polynomial spline. The problem differs from most problems discussed in the lecture notes since there are state constraints between the two end times. It is still possible to solve it using the methods discussed in the notes see e.g. [9].

1.2 Reduction of Redundancy

We will here by an example show that in general there may be many solutions to a control problem. Optimal control theory can then single out solutions that are best according to some cost criterium.

Consider control of the rocket car pictured in Figure 1.3. The control problem is to move the rocket car from rest (i.e., the speed is zero) at a point z_0 to rest at the origin. It is assumed that the car moves along a frictionless track. Newtons law gives the dynamical equation $m\ddot{z} = F$, where m is the mass of the car. If we introduce the coordinates $x_1 = z$ and $x_2 = \dot{z}$ and the control function $u = F/m$, then we get the equivalent state space realization

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix}u, \quad \begin{bmatrix}x_1(0)\\ x_2(0)\end{bmatrix} = \begin{bmatrix}z_0\\ 0\end{bmatrix}.$$

If we define

$$A = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}, \quad B = \begin{bmatrix}0\\ 1\end{bmatrix}, \quad x_0 = \begin{bmatrix}z_0\\ 0\end{bmatrix}$$

then our control problem can be stated as follows: *Find a control function* $u : [0, t_f] \to \boldsymbol{R}$ *such that the solution to* $\dot{x} = Ax + Bu$, $x(0) = x_0$ *satisfies* $x(t_f) = 0$.

It turns out that there are (infinitely) many solutions to this problem. In fact, if we define (see [14] or Chapter 4 to learn about these topics)

1. the transition matrix $\Phi(t, s)$ as the solution of the differential equation

$$\frac{\partial \Phi(t,s)}{\partial t} = A(t)\Phi(t,s), \quad \Phi(s,s) = I$$

which for the case of a constant A matrix becomes $\Phi(t,s) = e^{A(t-s)}$

2. the controllability Gramian

$$W(t_f, 0) = \int_0^{t_f} \Phi(t_f, s)BB^T\Phi(t_f, s)^T ds \tag{1.2}$$

then it follows that

$$u(t) = -B^T\Phi(t_f, t)^T W(t_f, 0)^{-1}\Phi(t_f, 0)x_0 + u_0(t) \tag{1.3}$$

is a solution to our control problem for every u_0 that satisfies

$$\int_0^{t_f} \Phi(t_f, s)Bu_0(s)ds = 0.$$

Optimal control theory can be used to reduce the number of choices by introducing a cost criterium that should be minimized. In this way we pick a solution that is in some sense better than the others. The following optimal control formulations are treated in, for example, [22].

1. Minimization of a quadratic cost $\int_0^{t_f} u(t)^2 dt$ gives an optimal control problem on the form

$$\text{minimize} \int_0^{t_f} u(t)^2 dt \quad \text{subj to} \quad \begin{cases} \dot{x} = Ax + Bu, \\ x(0) = x_0, \ x(t_f) = 0 \end{cases} \tag{1.4}$$

It is shown in [14] that the optimal solution is obtained by choosing $u_0 = 0$ in (1.3). The corresponding optimal cost is

$$J(t_f) = x_0^T \Phi(t_f, 0)^T W(t_f, 0)^{-1} \Phi(t_f, 0) x_0.$$

We will in this course learn how to derive this result. It should be noted that different choices of transition time t_f gives different solutions to the optimal control problem in (1.4).

2. The quadratic cost function has the effect that the control function, i.e., the force function F, becomes small in the sense that its energy is minimized. It is often the case that there exists an upper bound on the admissible acceleration force obtained from the jet-engine, i.e., $|F| \le K$ and hence $|u(t)| \le K/m$. Assume that the variables are scaled such that $K/m = 1$. The book [22] solves the following two optimization problems

   (*a*) Minimization of the transition time (Example 4.2)

   $$\text{minimize}\, t_f \quad \text{subj to} \quad \begin{cases} \dot{x} = Ax + Bu, \ |u| \le 1 \\ x(0) = x_0, \ x(t_f) = 0, \ t_f \ge 0 \end{cases}$$

   This is a so called *time optimal control problem.*

   (*b*) Minimization of the "fuel" for a fixed transition time (Example 5.11)

   $$\text{minimize} \int_0^{t_f} |u(t)| dt \quad \text{subj to} \quad \begin{cases} \dot{x} = Ax + Bu, \ |u| \le 1 \\ x(0) = x_0, \ x(t_f) = 0 \end{cases}$$

   This is the so called *fuel optimal control problem.*

   We treat problem (*a*) in the exercises.

1.3 Formal Problem Statement

The optimal control problems considered in this course are defined in terms of the system dynamics, the boundary conditions, the control constraints, and the cost criterium.

System dynamics: The system dynamics will be defined in terms of state space equations $\dot{x}_k = f_k(t, x_1, \ldots, x_n, u_1, \ldots, u_m)$, for $k = 1, \ldots, n$, or equivalently in vector form

$$\dot{x} = f(t, x, u), \quad \text{where} \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad f = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$$

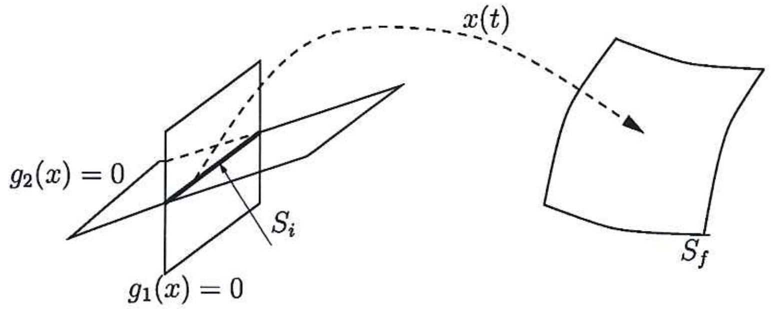

Figure 1.4: Optimal control from a line in $S_i = \{x : g_k(x) = 0,\ k = 1, 2\}$ in three dimensional space to a surface $S_f = \{x : g_3(x) = 0\}$.

The function f is called the *vector field.* It will generally be assumed that f is continuous with continuous partial derivatives with respect to t and x. Sometimes we also assume continuous differentiability with respect to u (for example in Chapter 4).

Boundary conditions: The initial and final (terminal) times are denoted by t_i and t_f, respectively, and the state vector is constrained to satisfy the boundary conditions $x(t_i) \in S_i$ and $x(t_f) \in S_f$. The sets S_i and S_f are in general smooth manifolds of the form

$$\begin{aligned} S &= \{x \in \mathbf{R}^n : g_k(x) = 0, k = 1, \ldots, p\}, \qquad (p < n) \\ &= \{x \in \mathbf{R}^n : G(x) = 0\}, \quad \text{where} \quad G(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_p(x), \end{bmatrix} \end{aligned}$$

where the gradients ∇g_k are assumed to be linearly independent[1]; this is equivalent to require that the functional matrix

$$G_x(x) = \begin{bmatrix} \frac{\partial g_1(x)}{\partial x_1} & \cdots & \frac{\partial g_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_p(x)}{\partial x_1} & \cdots & \frac{\partial g_p(x)}{\partial x_n} \end{bmatrix}$$

has rank p for all $x \in S$.

An example is illustrated in Figure 1.4 where the initial set is a line in $\mathbf{R}^3$ and the end point set is a surface. We will in the first few chapters focus on the important special case when we have

- fixed initial state, i.e., $S_i = \{x_i\}$.

[1]It is enough that they are linearly independent in a neighborhood of the optimal point.

- free terminal state, i.e., $S_f = \mathbf{R}^n$.

Control variable: The control variable will be restricted to be a piecewise continuous function that takes values in a set $U \subset \mathbf{R}^m$. The most common choices in the course will be $U = \mathbf{R}^m$ and $U = \{u \in \mathbf{R}^m : -1 \leq u_k \leq 1,\ k = 1, \ldots, m\}$.

Cost function: The cost function will in general consist of two terms

$$\phi(x(t_f)) + \int_{t_i}^{t_f} f_0(t, x(t), u(t))dt$$

where

- the terminal cost $\phi(x(t_f))$ penalizes deviation from some desired final state or manifold,
- the integral part of the cost is a cost associated with the state and control trajectories,
- ϕ and f_0 are generally assumed to be C^1, i.e., continuously differentiable with respect to the arguments.

The complete optimal control problem can now be formulated as

$$\text{minimize}\, \phi(x(t_f)) + \int_{t_i}^{t_f} f_0(t, x(t), u(t))dt \quad \text{subj to} \quad \begin{cases} \dot{x} = f(t, x, u), u \in U, \\ x(t_i) \in S_i,\ x(t_f) \in S_f, \\ u(\cdot) \text{ is piecewise cont} \end{cases} \tag{1.5}$$

The optimal control problem is thus to find an admissible control $u(t) \in U$ which transfers the state from some point in S_i to some point in S_f in such a way that the value of the cost functional becomes as small as possible. The initial time is always assumed fixed while the final time sometimes is a variable. In that case we add the constraint that the transition time is positive, $t_f - t_i \geq 0$, to the optimal control problem (1.5).

Remark 1. Maximization problems can be treated in exactly the same way as minimization problems since

$$\begin{aligned} &\text{maximize}\, \phi(x(t_f)) + \int_{t_i}^{t_f} f_0(t, x(t), u(t))dt \\ &\qquad = -\text{minimize} \left\{ -\phi(x(t_f)) - \int_{t_i}^{t_f} f_0(t, x(t), u(t))dt \right\} \end{aligned}$$

Admissible controls and optimal controls

A control signal is called *admissible* if it is piecewise continuous and such that all the constraints of (1.5) hold, i.e., $u(t) \in U$ for all $t \in [t_i, t_f]$ and such that an initial point $x_i \in S_i$ is transfered to a final point $x_f \in S_f$ by the dynamical equation $\dot{x} = f(t, x, u)$. The corresponding state function is also called admissible. We often use the notation $u(\cdot)$, and $x(\cdot)$ to denote a control function $u(t)$, $t \in [t_i, t_f]$ and the corresponding state function $x(\cdot)$. Sometimes we call $u(\cdot)$ and $x(\cdot)$ the *control and state trajectories.*

For every admissible control function $u(\cdot)$ and the corresponding state function $x(\cdot)$, we define the cost function

$$J(x(\cdot), u(\cdot)) = \phi(x(t_f)) + \int_{t_i}^{t_f} f_0(t, x(t), u(t))dt.$$

Then an admissible pair $(x^*(\cdot), u^*(\cdot))$ is *optimal* if

$$J(x^*(\cdot), u^*(\cdot)) \leq J(x(\cdot), u(\cdot))$$

for all other admissible pairs $(x(\cdot), u(\cdot))$.

Well-posedness of the optimal control problem

It is not always the case that there exists an optimal control and state function $(x^*(\cdot), u^*(\cdot))$. Several problems can appear

- There may not exist a control that transfers a point in S_i to the set S_f. In this case the value of the optimal control problem (1.5) is defined to be infinity.

  It is in general complicated to establish if the system is controllable in the sense that there exists a control that transfers a point in S_i to the set S_f. The situation is, however, simplified when the dynamics is linear

  $$\dot{x} = A(t)x + B(t)u.$$

  In particular, if the control constraint is $U = \mathbf{R}^m$, then there exist (many) admissible controls if the controllability Gramian defined in (1.2) is positive definite.

- Another potential difficulty is to establish that the solution of the differential equation $\dot{x} = f(t, x, u)$ is well defined on the interval $[t_i, t_f]$ for any piecewise continuous function $u(t) \in U$. If this is not the case then the constraint set in (1.5) is complicated to deal with. We discuss existence and uniqueness of solutions to ordinary differential equations in Chapter 3.

- It may happen that there does not exist an optimal control for (1.5). As an example, consider the optimization problem (1.4) but with the difference that the transition time t_f is a free variable to be chosen optimally. We have already seen that for fixed t_f the optimal solution becomes

$$J(t_f) = x_0^T \Phi(t_f, 0)^T W(t_f, 0)^{-1} \Phi(t_f, 0) x_0 = \frac{12 x_0^2}{t_f^3}$$

where the last equality is obtained after some simple calculations (using the particular A and B matrices in that example). It follows that $J(t_f) \to 0$ as $t_f \to \infty$. The corresponding optimal control (for given t_f) is

$$u(t) = -B^T \Phi(t_f, t)^T W(t_f, 0)^{-1} \Phi(t_f, 0) x_0 = \frac{12}{t_f^3}(t - t_f/2).$$

This shows that, as $t_f \to \infty$, the cost function decreases to the lower bound $J(\infty) = 0$. This means that if we add $t_f \geq 0$ as a free variable to be chosen optimally in (1.4) then there exists no optimal solution. The optimal value can only be achieved in the limit as $t_f \to \infty$ (and then the control tends to zero). The situation is completely analogous to the optimization problem

$$\text{minimize}\, x \quad \text{subj. to} \quad x > 0$$

where the optimal value 0 cannot be obtained because $x = 0$ does not belong to the constraint set $x > 0$.

- The above situation when the optimal control only exist in the limit gives an indication that the optimal control problem is not well posed. A slight change in the formulation of the optimization problem will often give a solution that is close to optimal. For example, the previous problem will have a solution if we introduce an upper bound for t_f ($t_f \leq T$) in the constraint set.

We will for the most part ignore these potential difficulties. This is not an essential loss since the main results of optimal control can be established and used without detailed knowledge and understanding of the topics we just pointed out.

Main approaches for optimal control

We will in the course focus on the following approaches for optimal control

Dynamic programming: Sufficient conditions for optimality are obtained by studying the optimal cost function (it is also a necessary condition under some additional assumptions). Some characteristics of this approach are:

+ It gives a sufficient condition for optimality. Hence, with this approach we immediately resolve the question of existence of an optimal control.
+ We obtain the optimal control in feedback form, $u(t) = \mu(t, x)$, for some function μ of time and the current state of the system.
- The optimal control is obtained by solving a partial differential equation.
- It requires the optimal cost function to be sufficiently smooth which is not always the case.

Pontryagin Minimum Principle: Necessary conditions for optimality are obtained by investigating local properties around the optimal control and state trajectories. The characteristics of the approach are:

+ It can be used in cases when dynamic programming fails due to lack of smoothness of the optimal cost function
+ It gives optimality conditions that are in general substantially easier to verify than solving the partial differential equation which arises in the dynamic programming approach.
- It only gives a necessary condition for optimality. This means that it can only be used to obtain candidates for optimality. Then these candidates must be investigated further to establish the optimal solution. This situation is analogous to scalar optimization, where the derivative condition $f'(x) = 0$ gives a necessary but not sufficient condition for optimality.

Computational Algorithms: It is in most cases impossible to find analytical solutions to optimal control problems. Instead we have to use numerical methods to obtain a solution. Several algorithms will be presented.

1.4 Notation

Let $J : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}$ be twice continuously differentiable with respect to x (C^2 w.r.t x). Then

$$J_x(t,x) = \frac{\partial J}{\partial x}(t,x) = \begin{bmatrix} \frac{\partial J(t,x)}{\partial x_1} \\ \vdots \\ \frac{\partial J(t,x)}{\partial x_n} \end{bmatrix}, \quad J_{xx}(t,x) = \frac{\partial^2 J(t,x)}{\partial x^2} = \begin{bmatrix} \frac{\partial^2 J}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 J}{\partial x_n \partial x_1} \\ \vdots & & \vdots \\ \frac{\partial^2 J}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 J}{\partial x_n \partial x_n} \end{bmatrix}$$

For a vector function $f : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^n$, we use the notation[2]

$$f(t,x) = \begin{bmatrix} f_1(t,x) \\ \vdots \\ f_n(t,x) \end{bmatrix}, \quad f_x(t,x) = \frac{\partial f(t,x)}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

The gradient of a function $g : \mathbf{R}^n \to \mathbf{R}$ is defined to be a column vector

$$\nabla g(x) = g_x(x) = \begin{bmatrix} \frac{\partial g(x)}{\partial x_1} \\ \vdots \\ \frac{\partial g(x)}{\partial x_n} \end{bmatrix}$$

and the Hessian is the second order derivative matrix

$$H_g(x) = g_{xx}(x) = \begin{bmatrix} \frac{\partial^2 g}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 g}{\partial x_n \partial x_1} \\ \vdots & & \vdots \\ \frac{\partial^2 g}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 g}{\partial x_n \partial x_n} \end{bmatrix}$$

For a vector function $G : \mathbf{R}^n \to \mathbf{R}^n$ we have

$$G_x(x) = \begin{bmatrix} \nabla g_1(x)^T \\ \vdots \\ \nabla g_n(x)^T \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{bmatrix}$$

If $H(x,u)$ is C^2 then a second order Taylor expansion around (x^0, u^0) is given as

$$H(x,u) = H(x^0,u^0) + H_x(x^0,u^0)\delta x + H_u(x^0,u^0)\delta u +$$
$$\frac{1}{2} \begin{bmatrix} \delta x \\ \delta u \end{bmatrix}^T \begin{bmatrix} H_{xx}(x^0,u^0) & H_{xu}(x^0,u^0) \\ H_{ux}(x^0,u^0) & H_{xx}(x^0,u^0) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta u \end{bmatrix} + H.O.T$$

where $\delta x = x - x^0$, and $\delta u = u - u^0$.

The Euclidean norm on $\mathbf{R}^n$ is defined as

$$\|x\| = \sqrt{\sum_{k=1}^{n} x_k^2}$$

1.4.1 Reference Literature

There are many books on optimal control. Some good references for the material in this course are the following ((A) advanced book that requires a higher level of mathematical sophistication than our course, (I) intermediate level, (E) elementary book)

Dynamic Programming: The classical book on this subject is the work of Bellman in

[2] $f_x(t,x)$ is sometimes in the literature denoted $J_f(t,x)$ (the Jacobi derivative).

(I) R. E. Bellman. *Dynamic Programming.* Princeton University Press, New Jersey, 1957.

Some recent good references are

(E) D. P. Bertsekas. *Dynamic Programming and Optimal Control (Vol 1).* Athena Scientific, 1995.

and chapter 5 in

(E) T. Basar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory.* SIAM, second edition, 1999.

Pontryagin Minimum Principle: The original book by Pontryagin and his colleagues is still one of the best treatments of the subject

(A) L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The Mathematical Theory of Optimal Processes.* Interscience Publishers, 1962.

Other good references are

(A) E. B. Lee and L. Marcus. *Foundations of Optimal Control.* Wiley, New York, 1967.

(I) M. Athans and P. Falb. *Optimal Control.* McGraw-Hill, New York, 1966.

(I) J. Macki and A. Strauss. *Introduction to Optimal Control Theory.* Springer-Verlag, 1982.

(I) G. Leitmann. *The Calculus of Variations and Optimal Control.* Plenum Press, New York, 1981.

(E) D. E. Kirk *Optimal Control Theory: An Introduction* Dover Publications, Mineola, New York, 1998.

(E) E. R. Pinch. *Optimal Control and the Calculus of Variations.* Oxford Science Publications, 1993.

A book that can be obtained free of charge on the internet is

P. Varaiya. *Lecture Notes on Optimization.* 1971.
http://paleale.eecs.berkeley.edu/~varaiya/

Numerical Methods: Numerical methods for optimal control are discussed in

(A) D. G. Luenberger. *Optimization by Vector Space Methods.* Wiley, New York, 1969.

(A) B. D. Craven. *Control and Optimization.* Chapman and Hall Mathematics, 1995

(E) A. E. Bryson. *Dynamic Optimization.* Addison-Wesley, Menlo Park, California, 1999.

(E) A. E. Bryson and Y. C. Ho. *Applied Optimal Control.* Blaisdell Waltham, 1969.

Optimal control SF2852, KTH

Optimal control SF2852, KTH

Chapter 2

Discrete Optimization

We will in the course mainly consider continuous-time optimal control problems. However, to get familiar with some of the key ideas we will in this chapter discuss optimal control problems in discrete time. In the first section on discrete dynamic programming we will even allow the state space to be discrete. Discrete dynamic programming has its application in a wide range of areas ranging from combinatorial optimization to coding theory and hybrid optimal control. It also helps us building the right intuitive understanding of the important *principle of optimality* and the associated *dynamic programming equation.*

We also introduce a discrete time version of the Pontryagin maximum principle (PMP). It shows that the first order optimality conditions has a certain structure for the class of dynamic optimization problems considered in this chapter. The continuous time PMP is more complicated but still derived in a "analogous fashion". We end the chapter with a discussion on optimal control problems on infinite time horizon.

2.1 Discrete Dynamic Programming

We begin with an example

Example 4 (Shortest path problem). Consider the directed graph in Figure 2.1. The initial node (or initial state) is connected to the terminal node (terminal state) through several possible paths. To each path is associated a cost obtained by adding the costs of each arc in the path. For example, the uppermost path $0 \to 1 \to 1 \to 1 \to 1 \to 0$ has the cost $1+5+4+1+2 = 13$. We are interested in finding the path with lowest cost. One way to solve this problem is simply to compute the cost of each path and then compare them. If N denotes the number of stages (in our case $N = 5$) then it is possible to verify that there are[1]

[1]The big ordo notation means that if $f(N) = O(\lambda^N)$ then there exists a constant such that $f(N) \leq c\lambda^N$ for large N.

$O((1+\sqrt{2})^N)$ paths. Hence, since we need to add N numbers to compute the cost of a single path, we need to do a total of $O((1+\sqrt{2})^N N)$ additions of numbers and then compare $O((1+\sqrt{2})^N)$ numbers in order to find the shortest path.

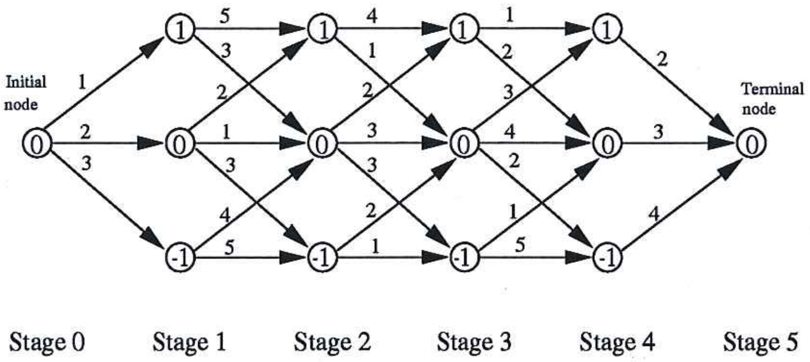

Figure 2.1: In a shortest path problem we want to compute the shortest (lowest cost) path from the initial node at stage 0 to the terminal node at stage 5. The cost of each arc is represented by a positive number above the arc. The numbers $\{1, 0, -1\}$ representing the three possible nodes at stage 1,2,3 and 4 represents the $\{\text{upper}, \text{middle}, \text{lower}\}$ positions.

Dynamic programming will provide us with a systematic and less expensive way of finding the shortest path. To develop this idea we first introduce some notation. We let $k \in \{0, 1, \ldots, 5\}$ denote the stage and for each stage the state vector $x_k := x(k) \in \{1, 0, -1\}$ tells us whether we are in the upper, middle, or lower node respectively. In this way we can represent the nodes of the graph with their "coordinates" (k, x_k). We let $J(k, x)$ be the shortest path (minimum cost path) from node (k, x) to the terminal node. The shortest path from stage 0 corresponds to $J(0, 0)$, which satisfies the following obvious relation

$$J(0,0) = \min(1 + J(1,1), 2 + J(1,0), 3 + J(1,-1)).$$

In words this means that the shortest path is the shortest of the following three paths

1. Go up to node $(1, 1)$ and then continue along the shortest path to the terminal node. This path has the cost $1 + J(1, 1)$.
2. Go forward to node $(1, 0)$ and then continue along the shortest path to the terminal node. This path has the cost $2 + J(1, 0)$.
3. Go down to node $(1, -1)$ and then continue along the shortest path to the terminal node. This path has the cost $3 + J(1, -1)$.

We can continue like this. For example, we have

$$J(1,1) = \min(5 + J(2,1), 3 + J(2,0)).$$

The basic principle behind the above formulas is the so called *principle of optimality*, which for our example says that

The shortest path has the property that for any initial part of the path from the initial node to some node $(k, x) \in \{1, \ldots, 5\} \times \{1, 0, -1\}$, *the remaining path must be the shorthest from the node* (k, x) *to the terminal node.*

We also notice that the cost corresponding to the terminal part of the shortest path is known in advance. Indeed, we have $J(5, 0) = 0$. This means that we can optimize backwards from stage 5 to stage 0 and this way recursively compute the *shortest path-to-go function* $J(k, x)$. Since, there is only one way of going from the nodes at stage 4 to the terminal node, we get

$$J(4,1) = 2, \quad J(4,0) = 3, \quad J(4,-1) = 4$$

In the next step, we get

$$\begin{aligned} J(3,1) &= \min(1 + J(4,1), 2 + J(4,0)) = \min(3,5) = 3 \\ J(3,0) &= \min(3 + J(4,1), 4 + J(4,0), 2 + J(4,-1)) = \min(5,7,6) = 5 \\ J(3,-1) &= \min(1 + J(4,0), 5 + J(4,-1)) = \min(4,9) = 4 \end{aligned} \tag{2.1}$$

If we continue like this we obtain the optimal solution in Figure 2.2. The cost of the shortest path is $J(0, 0) = 8$.

We can now determine the complexity of the dynamic programming approach. No addition and comparison is done while computing $J(5, x)$ and $J(4, x)$. To compute $J(3, x)$ for $x = 1, 0, -1$ using (2.1) we have to do a total of 7 additions and 4 comparisons. Hence, for arbitrary number of stages we need to add $3 + 7(N - 2)$ numbers and compare $2 + 4(N - 2)$ numbers. For large N this is much less expensive than computing the cost of all possible paths and then comparing them.

The shortest path problem in the previous example is a special case of a multistage decision problem. The general form is an optimal control problem

$$\text{minimize}\, \phi(x_N) + \Sigma_{k=0}^{N-1} f_0(k, x_k, u_k) \quad \text{subj. to} \quad \begin{cases} x_{k+1} = f(k, x_k, u_k) \\ x_0 \text{ given, } x_k \in X_k \\ u_k \in U(k, x_k) \end{cases} \tag{2.2}$$

This is similar to the continuous-time optimal control problem introduced in the previous chapter except that the system dynamics evolve in discrete *time* and the integral cost has been replaced by a sumation over the *time* axis. The reason we put *time* in italics is that the variable k may not have anything to do with time and should generally be viewed just as an enumeration of the stages of the optimization problem (2.2). Some comments about (2.2) are in order

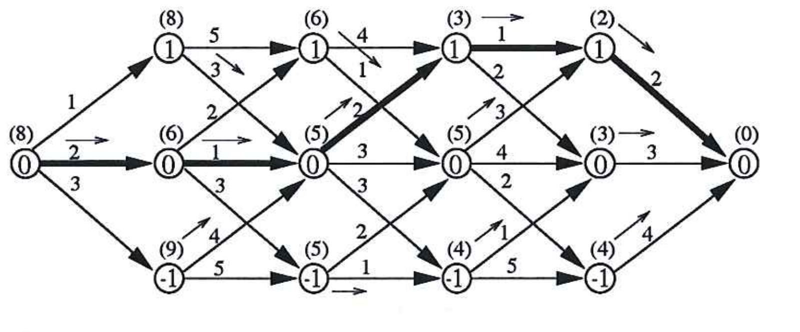

Figure 2.2: The optimal solution to the shortest path problem. The shortest path corresponds to the thick arcs in the graph. Above each node we have the *shortest path-to-go* $J(k, x)$ within parenthesis and the thin arrows corresponds to the optimal decision (direction) at each node.

System dynamics: The variable $k \in \{0, 1, \ldots, N\}$ represents an enumeration of the stages of the optimization problem (2.2). We generally call it the *time* variable and in many applications this is also the correct interpretation. The state vector x belongs to some state space that generally depends on k, i.e., we have $x_k := x(k) \in X_k$, where X_k denotes the state space at stage k. It is often the case that $X_k = \mathbf{R}^n$ for all $k = 0, 1, \ldots, N$ but it can also be a discrete set as in the shortest path problem. The evolution of the state vector is defined by the state equation

$$x_{k+1} = f(k, x_k, u_k)$$

where the *vector field* f generally depends on the time k, the state vector x_k, and the control variable u_k.

Control variable: The variable $u_k \in U(k, x_k)$ is the control (or decision) variable. The control constraint set $U(k, x_k)$, which constrains the range of the control variable, may depend on the present state of the system, i.e., (k, x_k).

Cost function: The cost function is additive and has one term corresponding to each stage. The terminal cost $\phi(x_N)$ penalizes deviation from a desired terminal state and the running cost adds a term $f_0(k, x_k, u_k)$ to the total cost at each stage.

The optimization problem (2.2) can be generalized by letting $x_0 \in S_0, x_N \in S_N$, where $S_0 \subset X_0$ and $S_N \subset X_N$ are given subsets of the state space. We can also let the number of stages N be variable.

Example 5 (Shortest path, continued). For the shortest path problem we have $N = 5$ and

- $x_0 = 0$, $X_k = \{1, 0, -1\}$ at $k = 1, \ldots, N-1$, and $X_N = \{0\}$.
- The control variable can in general take three values $1, 0, -1$ where $u_k = 1$ means go up, $u_k = 0$ means go forward, and $u_k = -1$ means go down. The control constraint set is

$$U(k,x) = \begin{cases} \{0,-1\}, & x = 1 \\ \{1,0,-1\}, & x = 0 \\ \{1,0\}, & x = -1 \end{cases}$$

  for $k = 0, \ldots, N-2$, and finally $U(N-1, 1) = -1$, $U(N-1, 0) = 0$, and $U(N-1,-1) = 1$.
- The state dynamics is given as

$$x_{k+1} = x_k + u_k,$$

  i.e., $f(k,x,u) = x + u$.
- The terminal cost $\phi(x) \equiv 0$ and the stage-wise additive costs $f_0(k,x,u) = c_{ij}^k$, where c_{ij}^k is the cost on the arrow from node (k,i) at stage k to node $(k+1, j)$ at stage $k+1$ in Figure 2.1. For example, $c_{0,1}^0 = 1$, $c_{1,1}^1 = 5$, $c_{1,0}^1 = 3$ and so on.

Next we will state the main result of this section. It shows that the optimal cost and the corresponding optimal control satisfies a dynamic programming recursion, which is a direct consequence of the principle of optimality:

If $\{u_k^*\}_{k=0}^{N-1}$ is an optimal control for (2.2), then $\{u_k^*\}_{k=n}^{N-1}$ is optimal for the subproblem obtained by considering an optimization on the form (2.2) but with initial condition $(n, x^*(n))$, i.e., we restart the optimization from somewhere along the optimal path.

Introduce the optimal *cost-to-go* function[2]

$$J^*(n,x) = \min \phi(x_N) + \Sigma_{k=n}^{N-1} f_0(k, x_k, u_k) \text{ subj. to } \begin{cases} x_{k+1} = f(k, x_k, u_k) \\ x_n = x, \ x_k \in X_k \\ u_k \in U(k, x_k) \end{cases}$$

for $n = 0, \ldots, N-1$ and $J^*(N,x) = \phi(x)$. In particular, the optimal solution of (2.2) is $J^*(0, x_0)$.

[2] We assume that the minimum exists otherwise min should be replaced by inf in the definition of $J^*(n,x)$. In particular, if no feasible solution exists then we would have $J*(n,x) = \infty$.

Theorem 1. *Suppose there exists a finite solution to the backwards dynamic programming recursion*

$$J(N,x) = \begin{cases} \phi(x), & x \in X_N \\ \infty, & x \notin X_N \end{cases}$$
$$J(n,x) = \min_{u \in U(n,x)} \{f_0(n,x,u) + J(n+1, f(n,x,u))\}, \quad n = N-1, N-2, \dots, 0$$

where the optimization over $U(n,x)$ is restricted to those control variables for which $f(n,x,u) \in X_{n+1}$. Then there exists an optimal solution to (2.2) and

(a) *$J^*(n,x) = J(n,x)$ for all $n = 0, \dots, N$, $x \in X_n$.*

(b) *The optimal feedback control in each stage is obtained as*

$$u_n^* = \mu(n,x) = \text{argmin}_{u \in U(n,x)} \{f_0(n,x,u) + J(n+1, f(n,x,u))\}.$$

Proof. We show by induction that $J^*(n,x) = J(n,x)$ for all n and x. By definition, we have $J^*(N,x) = J(N,x) = \phi(x)$. Assume, now that for some $n \in \{1, \dots, N-1\}$ we have $J^*(n+1,x) = J(n+1,x)$ for all $x \in X_{n+1}$. Then (here x_n is given by the state evolution $x_{n+1} = f(n, x_n, u_n)$)

$$J^*(n,x_n) = \min_{u_k \in U(k,x_k),\ k=n,\dots N-1} \left\{ \phi(x_N) + \sum_{k=n}^{N-1} f_0(k, x_k, u_k) \right\}$$
$$= \min_{u_n \in U(n,x_n)} \left\{ f_0(n,x_n,u_n) + \min_{u_k \in U(k,x_k),\ k=n+1,\dots N-1} \left\{ \phi(x_N) + \sum_{k=n+1}^{N-1} f_0(k,x_k,u_k) \right\} \right\}$$
$$= \min_{u_n \in U(n,x_n)} \{f_0(n,x_n,u_n) + J^*(n+1, f(n,x_n,u_n))\}$$
$$= \min_{u_n \in U(n,x_n)} \{f_0(n,x_n,u_n) + J(n+1, f(n,x_n,u_n))\}$$

where the first equality is by definition and the second follows since the cost function is additive over the different stages (which is the principle of optimality). In the third equality we used the definition of the optimal cost-to-go function and that $x_{n+1} = f(n, x_n, u_n)$. Finally, in the fourth equality we used the induction hypothesis. This proves statement (a).

Statement (b) follows since the optimum in the proof of (a) is obtained when using $u_n = \text{argmin}_{u \in U(n,x)} \{f_0(n,x,u) + J(n+1, f(n,x,u))\}$. □

Example 6 (Shortest path, continued). We now use Theorem 1 and the notation established in Example 5 to solve the shortest path problem in Figure 2.1. The first step of the recursion gives $J(5,x) = \phi(x) = 0$ (note that $X_5 = \{0\}$. At stage 4 we get

$$J(4,x) = \min_{u \in U(4,x)} \{f_0(4,x,u) + J(5,x)\} = \min_{u \in U(4,x)} \{c^4_{x,x+u}\}$$
$$= \begin{cases} c^4_{1,0} = 2, & x = 1 \\ c^4_{0,0} = 3, & x = 0 \\ c^4_{-1,0} = 4, & x = -1 \end{cases}$$

Next, for stage 3 we get

$$J(3,x) = \min_{u \in U(3,x)} \{f_0(3,x,u) + J(4,x+u)\} = \min_{u \in U(3,x)} \{c^3_{x,x+u} + J(4,x+u)\}$$
$$= \begin{cases} \min(c^3_{1,1} + J(4,1), c^3_{1,0} + J(4,0)) = \min(1+2, 2+3) = 3, & x = 1 \\ \min(c^3_{0,1} + J(4,1), c^3_{0,0} + J(4,0), c^3_{0,-1} + J(4,-1)) = \min(5,7,6) = 5, & x = 0 \\ \min(c^3_{-1,0} + J(4,0), c^3_{-1,-1} + J(4,-1)) = \min(4,9) = 4, & x = -1 \end{cases}$$

If we continue like this we get the optimal solution in Figure 2.2.

In the next example we consider the important linear quadratic control problem in discrete time.

Example 7. Consider the Linear Quadratic (LQ) control problem

$$\text{minimize}\, x_N^T Q_0 x_N + \sum_{k=0}^{N-1} \left(x_k^T Q x_k + u_k^T R u_k\right) \quad \text{subj. to} \quad \begin{cases} x_{k+1} = Ax_k + Bu_k \\ x_0 \text{ given} \end{cases} \tag{2.3}$$

where Q_0 and Q are symmetric positive semidefinite matrices and R is a symmetric and positive definite matrix. To find the optimal control we use the backwards recursion in Theorem 1, which for (2.3) becomes

$$J(N,x) = x^T Q_0 x$$
$$J(n,x) = \min_{u \in \mathbf{R}^m} \left\{x^T Q x + u^T R u + J(n+1, Ax + Bu)\right\}$$

The equations are quadratic with respect to x. Inspired by this we try $J(n,x) = x^T P_n x$. From the first equation we see that $P_N = Q_0$. The second equation becomes

$$x^T P_n x = \min_{u \in \mathbf{R}^m} \left\{x^T Q x + u^T R u + (Ax + Bu)^T P_{n+1} (Ax + Bu)\right\} \tag{2.4}$$

For the minimization in (2.4) we have

$$u_n^* = \mu(n,x) = \arg \min_{u \in \mathbf{R}^m} \left\{x^T Q x + u^T R u + (Ax + Bu)^T P_{n+1} (Ax + Bu)\right\}$$
$$= -(R + B^T P_{n+1} B)^{-1} B^T P_{n+1} A x,$$

which is obtained by differentiating with respect to u and then setting the derivative to zero. If we plug this into (2.4) we get

$$x^T P_n x = x^T \left(Q + A^T P_{n+1} A - A^T P_{n+1} B (R + B^T P_{n+1} B)^{-1} B^T P_{n+1} A \right) x$$

which must hold for all $x \in \mathbf{R}^n$. This means that the matrices P_n must satisfy the discrete-time Riccati equation

$$\begin{aligned} P_N &= Q_0 \\ P_n &= Q + A^T (P_{n+1} - P_{n+1} B (R + B^T P_{n+1} B)^{-1} B^T P_{n+1}) A \end{aligned} \tag{2.5}$$

for $n = N-1, N-2, \ldots, 0$. Note that $P_n \geq 0$ for all n, which implies that the inverse in (2.5) is well defined. To see that $P_n \geq 0$ we notice that $P_N \geq 0$ because $Q_0 \geq 0$. Furthermore, the minimum in (2.4) must be positive for $n = N-1$ because P_N, Q and R are all positive semidefinite. Induction proves the result.

To summarize, we have that the optimal cost-to-go and the optimal feedback control law are

$$\begin{aligned} J(n,x) &= x^T P_n x \\ u_n^* = \mu(n,x) &= -(R + B^T P_{n+1} B)^{-1} B^T P_{n+1} A x \end{aligned}$$

where P_n is the solution to the Riccati equation in (2.5).

2.2 Infinite Time Horizon Optimization

Let us consider multistage decision problems over an infinite time horizon. We consider the following general form of such problems

$$\min \Sigma_{k=0}^{\infty} f_0(x_k, u_k) \quad \text{subj. to} \quad \begin{cases} x_{k+1} = f(x_k, u_k) \\ x_0 \text{ given} \\ u_k \in U(x_k) \end{cases} \tag{2.6}$$

In order for the cost to be finite we need that $f_0(x_k, u_k) \to 0$ as $k \to \infty$. An interpretation is that (2.6) models problems where convergence to some particular set of values is desired. In our discussion we will set up the problem such that the state and the control must converge to zero.

The following assumptions are made

Assumption 1. We assume (w.l.o.g) that $0 \in X$, $U(0) = \{0\}$, $f(0,0) = 0$ and $f_0(0,0) = 0$.

The assumption implies that zero is an equilibrium point of the discrete dynamics

$$x_{k+1} = f(x_k, u_k)$$

This means that if $(x_0, u_0) = (0, 0)$ then the zero control $u_k = 0$, $\forall k$ implies that $x_k = 0$, $\forall k$.

We will always assume that $f(x_k, u_k) \in X$ for any $x_k \in X$ and $u_k \in U(x_k)$, i.e. that the state vector always remains in the presribed state space. A common situation is that $X = \mathbf{R}^n$ and then this is obviously satisfied.

In order to obtain the simplest possible result we will assume that the cost function and therefore the *value function* (the optimal cost of (2.6)) are positive definite and quadratically bounded.

Definition 1. A function $V : X \to \mathbf{R}$ is called *strictly positive definite*[3] if $V(0) = 0$ and there exists $c_1 > 0$ such that $V(x) \geq c_1\|x\|^2$ for all $x \in X$. It is called *strictly positive definite and quadratically bounded* if in addition there exists $c_2 > 0$ such that $V(x) \leq c_2\|x\|^2$ for all $x \in X$.

Example 8. A quadratic form $V(x) = x^T Px$, where $P = P^T$, is strictly positive definite if $P > 0$, i.e., if all eigenvalues of P are positive. It is clearly quadratically bounded.

Assumption 2. We assume that f_0 is strictly positive definite, i.e. there exists $\epsilon > 0$ such that $f_0(x, u) \geq \epsilon(\|x_k\|^2 + \|u_k\|^2)$.

Let us now define the optimal function (value function) corresponding to (2.6)

$$J^*(x_0) = \min_{u_k \in U} \sum_{k=0}^{\infty} f_0(x_k, u_k)$$

The value function is independent of time since the dynamics and cost function of (2.6) both are independent of time (the stage index k).

Theorem 2. *Suppose Assumption 1 and Assumption 2 hold. If there exists a strictly positive definite and quadratically bounded function* $V : X \to \boldsymbol{R}^+$ *that satisfies the Bellman equation*

$$V(x) = \min_{u \in U(x)} \{f_0(x, u) + V(f(x, u))\} \tag{2.7}$$

then

(a) $V(x) = J^*(x)$

(b) $u^* = \mu(x) = \operatorname{argmin}_{u \in U(x)} \{f_0(x, u) + V(f(x, u))\}$ *is an optimal feedback control that results in a globally convergent closed loop system, i.e. for any* $x_0 \in X$ *the optimal solution satisfies* $(x_k, \mu(x_k)) \to 0$ *as* $k \to \infty$.

[3]This definition of strictly positive definite is stronger than normal. It is used to simplify the understanding of this section.

Remark 2. The assumptions are stronger than necessary but it simplifies the proof. This allows us to prove a very strong form of convergence called exponential stability.

Proof. We first prove that $u_k = \mu(x_k)$ gives a globally convergent system. From the Bellman equation we get

$$V(x_{k+1}) = V(x_k) - f_0(x_k, \mu(x_k))$$

If we use Assumption 2 and the assumptions on V we get

$$\begin{aligned} V(x_{k+1}) &\leq V(x_k) - \epsilon(|x_k|^2 + |\mu(x_k)|^2) \\ &\leq (1 - \frac{\epsilon}{c_2})V(x_k) \end{aligned}$$

which implies $V(x_N) \leq (1-\epsilon/c_2)^N V(x_0)$. Finally, since $c_1\|x\|^2 \leq V(x) \leq c_2\|x\|^2$ we get

$$\|x_N\| \leq \sqrt{c_2/c_1}(1-\epsilon/c_2)^{N/2}\|x_0\| \tag{2.8}$$

which proves convergence. The control also converges because

$$\|\mu(x_N)\|^2 \leq \frac{1}{\epsilon}f_0(x_N, \mu(x_N)) \leq \frac{1}{\epsilon}V(x_N) \leq \frac{1}{\epsilon}c_2\|x_N\|^2$$

which due to (2.8) implies

$$|\mu(x_k)| \leq \sqrt{c_2^2/(\epsilon c_1)}(1-\epsilon/c_2)^{N/2}\|x_0\|$$

We have now proved that $u = \mu(x)$ is "stabilizing" in the sense that the closed loop state vector converges to zero. We will next see that it also gives the minimal cost. Consider an arbitrary control sequence $\{u_k\}_{k=0}^{\infty}$, which results in a convergent solution. By the Bellman equation we have

$$f_0(x_k, u_k) \geq V(x_k) - V(x_{k+1})$$

with equality if $u_k = \mu(x_k)$. Hence, we have

$$\begin{aligned} \lim_{N\to\infty}\sum_{k=0}^{N} f_0(x_k, u_k) &\geq \lim_{N\to\infty}\sum_{k=0}^{N} V(x_k) - V(x_{k+1}) \\ &= V(x_0) - \lim_{N\to\infty} V(x_N) = V(x_0) \end{aligned}$$

where we used that $x_N \to 0 \Rightarrow V(x_N) \to 0$. Since the first inequality becomes an equality when $u_k = \mu(x_k)$, we get

$$V(x_0) = \sum_{k=0}^{\infty} f_0(x_k, \mu(x_k)) \leq \sum_{k=0}^{\infty} f_0(x_k, u_k)$$

This proves the optimality. □

2.3 A Discrete Version of PMP

Consider the following discrete time optimal control problem

$$\text{minimize}\, \phi(x_N) + \sum_{k=0}^{N-1} f_0(k, x_k, u_k) \quad \text{subj. to} \quad \begin{cases} x_{k+1} = f(k, x_k, u_k), \\ x_0 \text{ is given, } G(x_N) = 0 \end{cases} \tag{2.9}$$

where $G(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{bmatrix}$ satisfies the usual regularity assumption, i.e. the gradients $\nabla g_k(x)$ are linearly independent. This is a special case of (2.2) in which $X = \mathbf{R}^n$ and $U = \mathbf{R}^m$. The dynamical programming approach to solving such problems is characterized by the following properties

- Feedback solutions are obtained. This means that we know the optimal control value for every position of the state vector x. This gives robustness to the closed loop system in the following sense: If the solution is perturbed by a disturbance then the controller still knows the optimal action.

- The solution is obtained using backwards recursion, which can be computationally demanding. One way to understand this is that we compute the optimal control value for every possible system state. What we win in robustness we loose in computational complexity.

- It is a sufficient condition for optimality.

We next use the Lagrange multiplier rule (also known as the Karush-Kuhn-Tucker conditions (KKT), or the first order optimality conditions) to obtain necessary conditions for optimality. The resulting conditions are the discrete version of the so-called Pontryagin minimization principle that we will study later in the course.

We recall from the optimization courses (the KKT conditions)

First order necessary condition (KKT): Suppose x^* is a (locally) optimal solution of

$$\text{minimize}\, f(x) \quad \text{subject to} \quad G(x) = 0$$

where $f : \mathbf{R}^n \to \mathbf{R}$ and $G : \mathbf{R}^n \to \mathbf{R}^p$ are continuously differentiable and the constraint set is regular, i.e., the gradients $\nabla g_k(x)$ are linearly independent. Then there exists a vector of *Lagrange multipliers* $\lambda \in \mathbf{R}^p$ such that

(*i*) $G(x^*) = 0$

(*ii*) $\nabla_x l(x^*, \lambda) = 0$, where $l(x, \lambda) = f(x) + \lambda^T G(x)$ is the *Lagrangian.*

We can use it to derive the following result.

Proposition 1. *Let* $\{u_k^*\}_{k=0}^{N-1}$ *be an optimal control for (2.9) and let* $\{x_k^*\}_{k=0}^{N}$ *be the corresponding trajectory. Then there exists an adjoint variable (Lagrange multiplier)* $\{\lambda_k\}_{k=1}^{N}$ *such that*

(*i*) *(adjoint equation)*

$$\lambda_k = \frac{\partial H}{\partial x}(k, x_k^*, u_k^*, \lambda_{k+1}), \quad k = 1, \dots, N-1$$

(*ii*) *("pointwise minimization")*

$$\frac{\partial H}{\partial u}(k, x_k^*, u_k^*, \lambda_{k+1}) = 0, \quad k = 0, 1, \dots, N-1$$

(*iii*) *(Boundary condition)*

$$\lambda_N = \frac{\partial \phi}{\partial x}(x_N^*) + G_x(x_N^*)^T \nu$$

for some $\nu \in \boldsymbol{R}^p$.

where the Hamiltonian is

$$H(k, x, u, \lambda) = f_0(k, x, u) + \lambda^T f(k, x, u)$$

Proof. Let

$$z = \begin{bmatrix} x_1^T & \dots & x_N^T & u_0^T & \dots & u_{N-1}^T \end{bmatrix}^T$$

$$\mathcal{F}(z) = \phi(x_N) + \sum_{k=0}^{N-1} f_0(k, x_k, u_k)$$

$$\mathcal{G}(z) = \begin{bmatrix} f(0, x_0, u_0) - x_1 \\ \vdots \\ f(N-1, x_{N-1}, u_{N-1}) - x_N \\ G(x_N) \end{bmatrix}$$

The Lagrange multiplier rule says that a necessary condition for optimality of

$$\min \mathcal{F}(z) \quad \text{subject to} \quad \mathcal{G}(z) = 0$$

is that there exists a Lagrange multiplier $\hat{\lambda}$ such that

$$\frac{\partial l}{\partial z}(z^*, \hat{\lambda}) = 0 \quad \text{where} \quad l(z, \hat{\lambda}) = \mathcal{F}(z) + \hat{\lambda}^T \mathcal{G}(z)$$

In our problem the Lagrange multiplier vector is $\hat{\lambda} = \begin{bmatrix}\lambda^T & \nu^T\end{bmatrix}^T$. We get

$$\begin{aligned}
\frac{\partial l}{\partial x_k}(z^*) &= \frac{\partial f_0}{\partial x}(k, x_k^*, u_k^*) + \lambda_{k+1}^T \frac{\partial f}{\partial x}(k, x_k^*, u_k^*) - \lambda_k, \quad k = 1, \ldots, N-1 \\
\frac{\partial l}{\partial x_N}(z^*) &= \frac{\partial \phi}{\partial x}(x_N^*) - \lambda_N + G_x(x_N^*)^T \nu \\
\frac{\partial l}{\partial u_k}(z^*) &= \frac{\partial f_0}{\partial u}(k, x_k^*, u_k^*) + \lambda_{k+1}^T \frac{\partial f}{\partial u}(k, x_k^*, u_k^*), \quad k = 0, \ldots, N-1
\end{aligned}$$

Hence, the condition $\frac{\partial l}{\partial z}(z^*, \hat{\lambda}) = 0$ together with the definition of the Hamiltonian function $H(k, x_k, u_k, \lambda_{k+1})$ proves the proposition. □

The propostion is often used in the following way

1. Define the Hamiltonian: $H(k, x, u, \lambda) = f_0(k, x, u) + \lambda^T f(k, x, u)$

2. Perform pointwise minimization, i.e. find a function $\mu(k, x, \lambda)$ such that $\frac{\partial H}{\partial u}(k, x, u, \lambda) = 0$. Hence the candidate optimal control is $u_k^* = \mu(k, x_k^*, \lambda_{k+1})$.

3. Solve the two point boundary value problem (TPBVP)

$$\begin{aligned}
x_{k+1} &= \frac{\partial H}{\partial \lambda}(k, x_k, \mu(k, x_k, \lambda_{k+1}), \lambda_{k+1}) = f(k, x_k, \mu(k, x_k, \lambda_{k+1})), \quad G(x_N) = 0 \\
\lambda_k &= \frac{\partial H}{\partial x}(k, x_k, \mu(k, x_k, \lambda_{k+1}), \lambda_{k+1}), \quad \lambda_N = \frac{\partial \phi}{\partial x}(x_N) + G_x(x_N)^T \nu
\end{aligned}$$

   We call this a two point boundary value problem because it involves boundary constraints both at the initial time and the final time (note that λ_0 is unknown). Proposition 1 reveals structure in the nonlinear program (2.9) that can be exploited.

The PMP approach is characterized by the following properties.

- It results in an open loop control program. This means that the optimal solution is only known for a particular initial condition x_0. If the solution is perturbed from the optimal by a disturbance then the optimal control may no longer be effective. The resulting system is therefore more sensitive to disturbances.

- It is generally easier to compute.

- It gives only a necessary condition for optimality.

2.4 Example: Feedback Versus Open Loop

Consider a particle moving along a 1-dimensional axis

$$x_{k+1} = x_k + u_k, \quad x_0 \text{ given.}$$

The state x denotes the position of the particle and the control u is the movement of the particle from one time instance to the next. We will compare open loop and feedback solutions for this problem with respect to their ability to reduce the effect of disturbances.

Let us assume that we want to bring the particle to the origin in two steps while using minimum control energy. This gives rise to the optimization problem

$$\begin{aligned} J^*(0, x_0) &= \min_u u_0^2 + u_1^2 \quad \text{s.t.} \quad \begin{cases} x_{k+1} = x_k + u_k \\ x_2 = 0 \end{cases} \\ &= \min u_0^2 + u_1^2 \quad \text{s.t.} \quad x_0 + u_0 + u_1 = 0 \\ &= \min u_0^2 + (x_0 + u_0)^2 = \frac{1}{2} x_0^2 \end{aligned}$$

and the optimal control sequence becomes

$$\begin{aligned} u_0^* &= \mu(0, x_0) = -\frac{1}{2} x_0, \\ u_1^* &= \mu(1, x_0) = -\frac{1}{2} x_0. \end{aligned}$$

The notation $u_k^* = \mu(k, x_0)$ is used to clarify that the control depends on the time k and the initial state x_0. Such control laws are called *open loop control.*

We obtain an alternative solution by using dynamic programming. The DynP algorithm gives

$$J(2, x) = \begin{cases} 0, & x = 0 \\ \infty, & x \neq 0 \end{cases}$$

Note that the cost is infinite unless the constraint $x = 0$ is satisfied. The next step of the DynP algorithm gives

$$J(1, x) = \min_u \{u^2 + J(2, x + u)\} = x^2$$

and the minimizing control is $u^* = \mu(1, x) = -x$. We used that $u = -x$ because otherwise $J(2, x + u) = \infty$, which would give $J(1, x) = \infty$. The final step of the DynP algorithm gives

$$J(0, x) = \min_u \{u^2 + J(1, x + u)\} = \min_u \{u^2 + (x + u)^2\} = \frac{1}{2} x^2$$

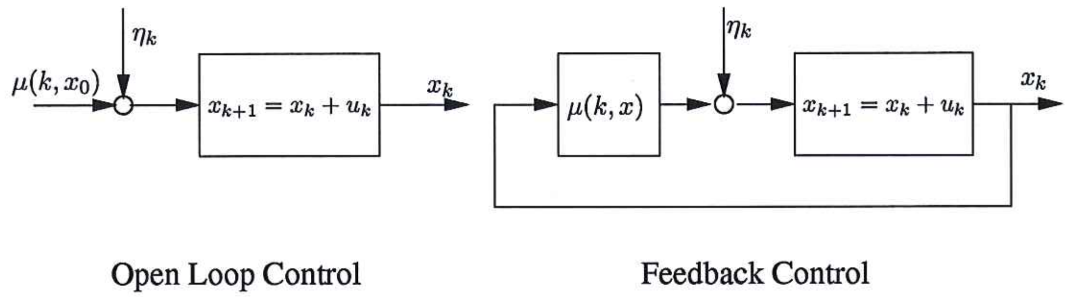

Figure 2.3: Open loop versus feedback control.

and the minimizing control is $u^* = \mu(0, x) = -\frac{1}{2}x$. Hence the optimal control is

$$u_0^* = \mu(0, x) = -\frac{1}{2}x$$
$$u_1^* = \mu(1, x) = -x$$

The notation $u_k^* = \mu(k, x_k)$ is used to clarify that the control depends on both time and the *current* state. Such control laws are called *feedback control.*

Disturbance Sensitivity

Consider the situatation in Fig. 2.3 where the left part illustrates the open loop situation and the right hand side illustrates the feedback control situation. The signal η_k denotes a disturbance and we first assume $\eta_0 \neq 0$ and $\eta_k = 0,\ k \geq 1$.

For the open loop case we get

$$x_1 = x_0 + \mu(0, x_0) + \eta_0 = \frac{1}{2}x_0 + \eta_0$$
$$x_2 = x_1 + \mu(1, x_0) = \eta_0$$

We do not reach the origin as desired!

In the closed loop case we get

$$x_1 = x_0 + \mu(0, x_0) + \eta_0 = \frac{1}{2}x_0 + \eta_0$$
$$x_2 = x_1 + \mu(1, x_1) = 0$$

The feedback compensated for the disturbance. Note however, that if the disturbance is persistent, i.e. $\eta_k \neq 0,\ k \geq 1$, then the position of the particle can still be disturbed and thus deviate from the desired position at the origin. In order to avoid such a situation we could consider optimal control over an infinite time horizon.

2.4.1 Infinite Horizon Optimal Control

Consider

$$\min_u \sum_{k=1}^{\infty} x_k^2 + u_k^2 \quad \text{s.t.} \quad x_{k+1} = x_k + u_k$$

The cost function forces the state and the control to converge to zero. We obtain a solution by solving the Bellman equation

$$J(x) = \min_u \left\{x^2 + u^2 + J(x+u)\right\}$$

Let us try the form $J(x) = px^2$, where $p > 0$ in order for J to be positive definite. This gives

$$\begin{aligned} px^2 &= \min_u \left\{x^2 + u^2 + p(x+u)^2\right\} \\ &= \min_u (1+p)(u + \frac{p}{1+p}x)^2 + x^2(1 + p - \frac{p^2}{1+p}) \\ &= x^2(1+p - \frac{p^2}{1+p}) \end{aligned}$$

Hence, the optimal feedback control is

$$u^* = \mu(x) = -\frac{p}{1+p}x$$

where p is the positive solution to the Riccati equation

$$p^2 = 1 + p \tag{2.10}$$

Let us consider the situation in Fig 2.4. We obtain the solution

$$\begin{aligned} x_{k+1} = x_k + \mu(x_k) + \eta_k &= \frac{1}{1+p}x_k + \eta_k \\ &= \frac{1}{1+p}(x_{k-1} + \mu(x_{k-1}) + \eta_{k-1}) + \eta_k \\ &= \frac{1}{(1+p)^2}x_{k-1} + \frac{1}{1+p}\eta_{k-1}) + \eta_k \\ &= \ldots = \frac{1}{(1+p)^{k+1}}x_k + \sum_{l=0}^{k} \frac{1}{(1+p)^{k-l}}\eta_l \end{aligned}$$

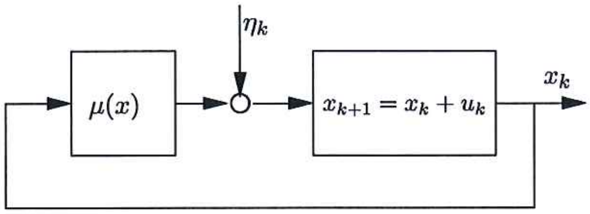

Figure 2.4: Infinite horizon control.

We see that the influence of the initial condition decays to zero since $1/(1+p) < 1$. We also see that the contribution from old disturbances is reduced as time evolves. Hence, the infinite time horizon optimal control problem results in a stabilizing (convergence) feedback controller that also gives robustness to the disturbance.

We have the following conclusions

- Feedback solutions have the advantage that the effect of disturbances can be compensated for.
- Infinite time horizon optimal control gives both convergence and disturbance compensation. Note that the convergence to the desired value (zero in our example) in general is slower than if a finite time horizon optimal criterion is used for the control design.
- It is usually easier to derive an open loop controller than a feedback controller.

Optimal control SF2852, KTH

Chapter 3

Dynamic Programming

We will in this chapter discuss dynamic programming and the related Hamilton-Jacobi-Bellman equation. We use the ideas from the previous chapter to build the right intuitive understanding of the important *principle of optimality* and the associated *dynamic programming equation.*

3.1 Continuous Time Dynamic Programming

We consider the optimal control problem

$$\text{minimize}\, \phi(x(t_f)) + \int_{t_i}^{t_f} f_0(t, x(t), u(t))dt \quad \text{subj. to} \quad \begin{cases} \dot{x}(t) = f(t, x(t), u(t)) \\ x(t_i) = x_i, \ u(t) \in U \end{cases} \tag{3.1}$$

where t_i and t_f are fixed initial and terminal times and x_i is a fixed initial point. The end point $x(t_f)$ is free and can take any value in $\mathbf{R}^n$. The control is a piecewise continuous function, which satisfies the constraint $u(t) \in U$, for $t \in [t_i, t_f]$. It should be remarked that we could extend everything in this section to the case when $U = U(t, x)$, i.e., the control constraint set depends on time and the state.

We will embed the optimization problem (3.1) in a larger class of problems by considering optimization from any initial point (t_0, x_0), where $t_0 \in [t_i, t_f]$ and $x_0 \in \mathbf{R}^n$. It turns out that these optimization problems are related via a partial differential equation, viz. the Hamilton-Jacobi-Bellman equation (HJBE).

Let $u(\cdot)$ be an admissible control on $[t_0, t_f]$, i.e., it is piecewise continuous with

$u(t) \in U$ for $t \in [t_0, t_f]$. Then the *cost-to-go function* is defined as[1]

$$J(t_0, x_0, u(\cdot)) = \phi(x(t_f)) + \int_{t_0}^{t_f} f_0(t, x(t), u(t))dt$$

where $\dot{x}(t) = f(t, x(t), u(t))$, $x(t_0) = x_0$. Next, we define the *optimal cost-to-go function* as[2] (this is also called the *value function*)

$$J^*(t_0, x_0) = \min_{u(\cdot)} J(t_0, x_0, u(\cdot))$$

where the minimization is performed with respect to all admissible controls. This means in particular that the optimization problem (3.1) can be written

$$\min_{u(\cdot)} J(t_i, x_i, u(\cdot)),$$

where again the minimization is performed with respect to all admissible controls $u(t) \in U$. Moreover, if $u^*(\cdot)$ is the optimal control function then

$$J^*(t_i, x_i) = J(t_i, x_i, u^*(\cdot)),$$

which of course means that $J^*(t_i, x_i)$ is the optimal value of (3.1). Note also that J^* must satisify the boundary condition $J^*(t_f, x) = \phi(x)$.

We next prove the principle of optimality, which immediately leads to the dynamic programming equation. This equation is then used to show that J^* must satisfy HJBE.

The Principle of Optimality

The principle of optimality states a fundamental property that holds for all optimization problems considered in this course. It simply says that the restriction of an optimal control to a subinterval is optimal for the corresponding restricted optimization problem. This is an obvious observation but we still give a formal proof.

The principle of optimality is illustrated in Figure 3.1.

Proposition 2. *Let $u^* : [t_0, t_f] \to \boldsymbol{R}^m$ be an optimal control for $\min_{u(\cdot)} J(t_0, x_0, u(\cdot))$ that generates the optimal trajectory $x^* : [t_0, t_f] \to \boldsymbol{R}^n$. Then, for any $t' \in (t_0, t_f]$, the restriction of the optimal control to $[t', t_f]$, $u^*|_{[t',t_f]}$, is optimal for $\min_{u(\cdot)} J(t', x^*(t'), u(\cdot))$ and the corresponding optimal trajectory is $x^*|_{[t',t_f]}$.*

[1] We implicitly assume that for each admissible control $u(\cdot) : [t_0, t_f] \to U$, there exists a unique solution to the differential equation $\dot{x} = f(t, x, u)$, $x(t_0) = x_0$. (This is true if the vector field is sufficiently regular, see the next chapter for a discussion.). This assumption makes J dependent only on t_0, x_0, and $u(\cdot)$, and not the full trajectory $x(\cdot)$, since it is completely specified by the other three.

[2] We have assumed there exists a minimizing solution otherwise min should be replaced by inf. This assumption is made throughout the section.

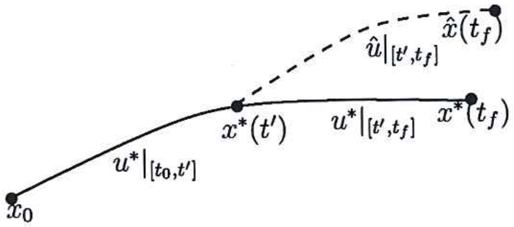

Figure 3.1: Illustration of the principle of optimality. The solid trajectory corresponds to the optimal control function $u^*(\cdot)$ and the dashed trajectory corresponds to another admissible control $\hat{u}(\cdot)$. The principle of optimality says that if $u^*(\cdot)$ is optimal over $[t_0, t_f]$ then it is also optimal over any subinterval. This means in particular that the dashed trajectory corresponds to a higher cost (or possibly equal if the optimal control is not unique) than the solid trajectory.

Proof. By additivity of the cost function we have

$$J^*(t_0, x_0) = \int_{t_0}^{t'} f_0(t, x^*(t), u^*(t))dt + J(t', x^*(t'), u^*|_{[t',t_f]})$$

Suppose that $u^*|_{[t',t_f]}$ is not optimal over the interval $[t', t_f]$ when the initial point is $x(t') = x^*(t')$. Then there exists an admissible control function $\hat{u}(\cdot)$ defined on $[t', t_f]$ such that

$$J(t', x^*(t'), \hat{u}(\cdot)) < J(t', x^*(t'), u^*|_{[t',t_f]}).$$

The control

$$u(t) = \begin{cases} u^*(t), & t \in [t_0, t') \\ \hat{u}(t), & t \in [t', t_f] \end{cases} \tag{3.2}$$

is admissible for (3.1) and gives the cost

$$\begin{aligned} J(t_0, x_0, u(\cdot)) &= \int_{t_0}^{t'} f_0(t, x^*(t), u^*(t))dt + J(t', x^*(t'), \hat{u}(\cdot)) \\ &< \int_{t_0}^{t'} f_0(t, x^*(t), u^*(t))dt + J(t', x^*(t'), u^*|_{[t',t_f]}) = J^*(t_0, x_0) \end{aligned}$$

This contradicts the optimality of $u^*(\cdot)$ and we conclude that the restriction $u^*|_{[t',t_f]}$ is optimal over $[t', t_f]$. □

The Dynamic Programming Equation

The dynamic programming equation is a direct consequence of the principle of optimality. Let (t_0, x_0) be given. From the principle of optimality it follows that the optimal cost-to-go function satisfies the relation

$$\begin{aligned} J^*(t_0, x_0) &= \int_{t_0}^{t'} f_0(s, x^*(s), u^*(s))ds + J^*(t', x^*(t')) \\ &= \min_{u(\cdot)} \left\{ \int_{t_0}^{t'} f_0(s, x(s), u(s))ds + J^*(t', x(t')) \right\}. \end{aligned}$$

where the minimization is with respect to all admissible controls, i.e., $u(s) \in U$ for $s \in [t_0, t']$. If we let the starting point be $(t, x(t))$ and $t' = t + \Delta t$ then the above relation becomes

$$J^*(t, x(t)) = \min_{u(\cdot)} \left\{ \int_{t}^{t+\Delta t} f_0(s, x(s), u(s))ds + J^*(t + \Delta t, x(t + \Delta t)) \right\}, \quad (3.3)$$

This is the dynamic programming equation and it shows that the optimal control can be computed in a backward direction. In other words, if you know $J^*(t + \Delta t, x)$ for any $x \in \mathbf{R}$ then you can compute $J^*(t, x)$ using this formula. This is particularly useful for discrete time systems with a discrete state space as we saw in the previous section. For continuous time systems we take one more step and derive the HJBE, which in general gives a more constructive way to determine J^*.

Hamilton-Jacobi-Bellman Equation (HJBE)

We will here derive the Hamilton-Jacobi-Bellman Equation using the dynamic programming equation. We have (for an arbitrary initial point)

$$\begin{aligned} J^*(t_0, x_0) &= \min_{u(\cdot)} \left\{ \int_{t_0}^{t_0+\Delta t} f_0(s, x(s), u(s))ds + J^*(t_0 + \Delta t, x(t_0 + \Delta t)) \right\} \\ &= \min_{u(\cdot)} \left\{ \int_{t_0}^{t_0+\Delta t} f_0(s, x(s), u(s))ds + J^*(t_0 + \Delta t, x_0 + f(t_0, x(t_0), u(t_0))\Delta t + o(\Delta t)) \right\} \\ &= \min_{u(\cdot)} \left\{ f_0(t_0, x_0, u(t_0))\Delta t + J^*(t_0, x_0) + \left(\frac{\partial J^*}{\partial t}(t_0, x_0) \right.\right. \\ &\qquad \left.\left. + \frac{\partial J^*}{\partial x}(t_0, x_0)^T f(t_0, x_0, u(t_0)) \right) \Delta t + o(\Delta t) \right\} \end{aligned}$$

where in the second equality we used the Euler approximation of the system equation, $x(t_0 + \Delta t) = x_0 + f(t_0, x_0, u(t_0))\Delta t + o(\Delta t)$, and in the third equality we made a Taylor expansion of J^* around (t_0, x_0). If we divide by Δt and use the fact that $\lim_{\Delta t \to 0} o(\Delta t)/\Delta t = 0$ then we get[3]

$$-\frac{\partial J^*}{\partial t}(t_0, x_0) = \min_{u \in U} \left\{ f_0(t_0, x_0, u) + \frac{\partial J^*}{\partial x}(t_0, x_0)^T f(t_0, x_0, u) \right\},$$

[3]This is not as trivial as it looks since one should justify that $\min_u$ and $\lim_{\Delta t \to 0}$ can be permuted.

where the optimization over u is pointwise. Since this partial differential equation was derived for arbitrary $x_0 \in \mathbf{R}^n$ and $t_0 \in [t_i, t_f]$ it follows that

$$-\frac{\partial J^*}{\partial t}(t,x) = \min_{u \in U} \left\{ f_0(t,x,u) + \frac{\partial J^*}{\partial x}(t,x)^T f(t,x,u) \right\}, \quad \forall (t,x) \in [t_i, t_f] \times \mathbf{R}^n \tag{3.4}$$

which is the Hamilton-Jacobi-Bellman Equation (HJBE). The boundary condition to HJBE is $J^*(t_f, x) = \phi(x)$.

What we have shown is the following: Assume that

- there exists an optimal control $u^*(\cdot)$
- the optimal cost to go function J^* is C^1 in both arguments

then

- J^* solves the HJBE in (3.4).
- $u^*(t)$ is the minimizing argument in (3.4) (pointwise).

This necessary condition for optimality is not as useful as it may appear since it assumes that the value function J^* is C^1, which is not always the case. The Pontryagin minimum principle gives much weaker and more useful necessary conditions. However, the great thing about the HJBE in (3.4) is that it also gives a sufficient condition for optimality.

The Verification Theorem for Dynamic Programming

We will next state and prove one of the main results of this course. The theorem shows that the HJBE in (3.4) gives a sufficient condition for an optimal control to exist. This gives us our first systematic method to synthesize an optimal control.

Theorem 3. *Suppose*

(*i*) $V : [t_i, t_f] \times \boldsymbol{R}^n \to \boldsymbol{R}$ *is* C^1 *(in both arguments) and solves HJBE*

$$\begin{aligned} -\frac{\partial V}{\partial t}(t,x) &= \min_{u \in U} \left\{ f_0(t,x,u) + \frac{\partial V}{\partial x}(t,x)^T f(t,x,u) \right\} \\ V(t_f, x) &= \phi(x) \end{aligned} \tag{3.5}$$

(*ii*) $\mu(t,x) = \arg\min_{u \in U} \left\{ f_0(t,x,u) + \frac{\partial V}{\partial x}(t,x)^T f(t,x,u) \right\}$ *is admissible*[4]

[4] This means that $u^*(t) = \mu(t, x(t))$ is a piecewise continuous function for any closed loop solution. The closed loop system equation $\dot{x}(t) = f(t, x(t), \mu(t, x(t)))$ must therefore be well defined on $[t_i, t_f]$ for all initial conditions. This topic is discussed in Chapter 4.

Then

(a) $V(t,x) = J^*(t,x)$ *for all* $(t,x) \in [t_i, t_f] \times \boldsymbol{R}^n$.

(b) $\mu(t,x)$ *is the optimal feedback control law, i.e.* $u^*(t) = \mu(t, x(t))$.

Remark 3. The second condition (ii) means

$$\min_{u \in U} \left\{ f_0(t,x,u) + \frac{\partial V}{\partial x}(t,x)^T f(t,x,u) \right\}$$
$$= f_0(t,x,\mu(t,x)) + \frac{\partial V}{\partial x}(t,x)^T f(t,x,\mu(t,x)) \quad (3.6)$$

Remark 4. The theorem says that

$$\begin{aligned} \dot{x}^*(t) &= f(t, x^*(t), u^*(t)), \quad x^*(t_i) = x_i \\ u^*(t) &= \mu(t, x^*(t)) \end{aligned}$$

is the optimal solution to (3.1).

Proof. Let $u(\cdot)$ be an admissible control on $[t_0, t_f]$ and let $x(\cdot)$ be the corresponding solution to $\dot{x} = f(t, x(t), u(t))$, $x(t_0) = x_0$. We have

$$V(t_f, x(t_f)) - V(t_0, x_0) = \int_{t_0}^{t_f} \dot{V}(t, x(t))dt$$
$$= \int_{t_0}^{t_f} [\frac{\partial V}{\partial t}(t, x(t)) + \frac{\partial V}{\partial x}(t, x(t))^T f(t, x(t), u(t))]dt \geq -\int_{t_0}^{t_f} f_0(t, x(t), u(t))dt \quad (3.7)$$

where the inequality follows since (3.5) implies

$$\frac{\partial V}{\partial t}(t, x(t)) + \frac{\partial V}{\partial x}(t, x(t))^T f(t, x(t), u(t)) \geq -f_0(t, x(t), u(t)).$$

Using $V(t_f, x(t_f)) = \phi(x(t_f))$ in (3.7) gives

$$V(t_0, x_0) \leq \phi(x(t_f)) + \int_{t_0}^{t_f} f_0(t, x(t), u(t))dt = J(t_0, x_0, u(\cdot))$$

This inequality holds for all admissible controls, in particular the optimal $u^*(\cdot)$. Hence, we have shown that

$$V(t_0, x_0) \leq J^*(t_0, x_0) \quad (3.8)$$

for all initial points. We will next see that equality is obtained by using $u(t) = \mu(t, x(t))$. Indeed, by combining (3.5) and (3.6), and then rearranging terms we get

$$\frac{\partial V}{\partial t}(t,x) + \frac{\partial V}{\partial x}(t,x)^T f(t,x,\mu(t,x)) = -f_0(t,x,\mu(t,x))$$

Integration of this equation then gives (where we again use that $V(t_f, x(t_f)) = \phi(x(t_f))$ and $u(t) = \mu(t, x(t))$)

$$V(t_0, x_0) = \phi(x(t_f)) + \int_{t_0}^{t_f} f_0(t, x(t), u(t))dt = J(t_0, x_0, u(\cdot)) \geq J^*(t_0, x_0) \quad (3.9)$$

where the last inequality follows since $J^*(t_0, x_0) = \min_{u(\cdot)} J(t_0, x_0, u(\cdot))$. Combining the inequalities in (3.8) and (3.9) gives $V(t_0, x_0) = J^*(t_0, x_0)$. This also shows that $(x(t), \mu(t, x(t)))$ is the optimal state and control trajectory, i.e., $x^*(t) = x(t)$ and $u^*(t) = \mu(t, x(t))$. Since (t_0, x_0) are arbitrary, this proves the theorem. □

Before we give some examples we make a short comment on the use of Theorem 3. For a given optimal control problem on the form (3.1) we take the following steps[5]

1. Define the Hamiltonian

$$H(t,x,u,\lambda) = f_0(t,x,u) + \lambda^T f(t,x,u).$$

   Here $\lambda \in \mathbf{R}^n$ is a parameter vector.

2. Optimize pointwise over u to obtain

$$\tilde{\mu}(t,x,\lambda) = \arg\min_{u\in U} H(t,x,u,\lambda) = \arg\min_{u\in U} \left\{ f_0(t,x,u) + \lambda^T f(t,x,u) \right\}.$$

3. Solve the partial differential equation

$$-\frac{\partial V}{\partial t}(t,x) = H\left(t, x, \tilde{\mu}(t,x,\frac{\partial V}{\partial x}(t,x)), \frac{\partial V}{\partial x}(t,x)\right)$$

   subject to the initial condition $V(t_f, x) = \phi(x)$.

Then $\mu(t,x) = \tilde{\mu}(t,x,\frac{\partial V}{\partial x}(t,x))$ is the optimal feedback control law, i.e. $u^*(t) = \mu(t, x(t))$.

[5]The same optimization as in step 2 is a part of the conditions in PMP.

3.2 Examples

We next give some examples.

Example 9. Consider the mass and spring system in Figure 3.2. The system equation is

$$m\ddot{z} = -kz + F;$$

We assume for simplicity of notation that $k/m = 1$, and that the force is bounded by $F/m \in [-1, 1]$. Our goal is to find a control law for the force function such

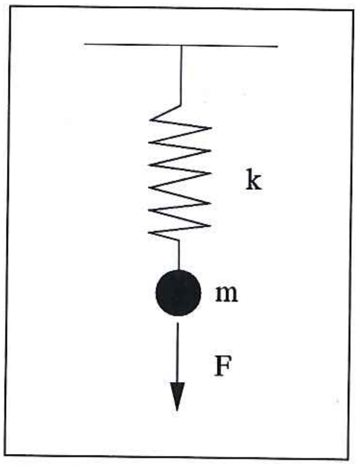

Figure 3.2: A mass hanging in a spring.

that the deviation from zero is as large as possible after time t_f. It is assumed that the initial condition is known $(z(0), \dot{z}(0))$.

The first step is to convert the system into state space form. Let $x_1 = z$, $x_2 = \dot{z}$. Then

$$\begin{aligned} \dot{x}_1 &= x_2, \quad x_1(0) = z(0) \\ \dot{x}_2 &= -x_1 + u, \quad x_2(0) = \dot{z}(0) \end{aligned}$$

If we define

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad x_0 = \begin{bmatrix} z(0) \\ \dot{z}(0) \end{bmatrix}$$

then the optimization problem can be stated as

$$\text{maximize}\, Cx(t_f) \quad \text{subj. to} \quad \begin{cases} \dot{x} = Ax + Bu, & x(0) = x_0 \\ |u| \leq 1 \end{cases}$$

We follow the solution strategy:

1. $H(t, x, u, \lambda) = \lambda^T(Ax + Bu)$
2. $\widetilde{\mu}(t, x, \lambda) = \arg\max_{|u|\leq 1}\{\lambda^T(Ax + Bu)\} = \text{sign}(\lambda^T B)$, which gives
$$H(t, x, \widetilde{\mu}(t, x, \lambda), \lambda) = \lambda^T Ax + |\lambda^T B|.$$
3. As the last step we need to solve the HJBE
$$-\frac{\partial V}{\partial t}(t, x) = \frac{\partial V}{\partial x}(t, x)^T Ax + |\frac{\partial V}{\partial x}(t, x)^T B|$$
subject to $V(t_f, x) = Cx$.

This HJBE is affine (linear+constant) with respect to x. Inpsired by this we try $V(t, x) = \psi(t)^T x + \alpha(t)$, where ψ and α are functions we need to determine.

The end condition gives $V(t_f, x) = \psi(t_f)^T x + \alpha(t_f) = Cx(t_f)$, which holds if $\psi(t_f) = C^T$ and $\alpha(t_f) = 0$. The HJBE becomes
$$-\dot{\psi}^T x - \dot{\alpha} = \psi^T Ax + |\psi^T B|$$
which holds if $\dot{\alpha} = -|\psi^T B|$ and $\dot{\psi} = -A^T\psi$. Putting everything together gives the optimal control
$$u^*(t) = \text{sign}(\psi(t)^T B) \quad \text{where} \quad \begin{cases} \dot{\psi} = -A^T\psi, \\ \psi(t_f) = C^T \end{cases}$$
It is now possible to explicitly compute this control function and then determine the maximum deviation $z(t_f)$.

Example 10. Solve the linear quadratic optimal control problem
$$\text{minimize}\, x(t_f)^T Q_0 x(t_f) + \int_0^{t_f} [x(t)^T Qx(t) + u(t)^T Ru(t)]dt$$
$$\text{subj. to} \quad \begin{cases} \dot{x} = Ax + Bu \\ x(0) = x_0 \end{cases}$$
where Q_0 and Q are symmetric positive semi-definite matrices and R is a symmetric positive definite matrix. The solution strategy gives

1. $H(t, x, u, \lambda) = x^T Qx + u^T Ru + \lambda^T(Ax + Bu)$
2. $\widetilde{\mu}(t, x, \lambda) = \arg\min_u\{x^T Qx + u^T Ru + \lambda^T(Ax + Bu)\} = -\frac{1}{2}R^{-1}B^T\lambda$. To see this we differentiate with respect to u and look for a stationary point. This gives $2Ru + \lambda^T B = 0$. Hence, $u = -\frac{1}{2}R^{-1}B^T\lambda$, which must be the unique minimizing argument since R is positive definite.

   We get
$$H(t, x, \widetilde{\mu}(t, x, \lambda), \lambda) = x^T Qx - \frac{1}{4}\lambda^T BR^{-1}B^T\lambda + \lambda^T Ax$$

3. Inspired by the quadratic form of $H(t, x, \widetilde{\mu}(t, x, \lambda), \lambda)$ and the boundary condition $V(t_f, x) = x^T Q_0 x$, we try $V(t, x) = x^T P(t) x$, where P is a symmetric matrix function (it must also be positive semidefinite since the cost function is positive).

Using $V_t(t, x) = x^T \dot{P}(t) x$ and $V_x(t, x) = 2P(t)x$, the HJBE PDE and its boundary condition become

$$x^T[\dot{P} + Q - PBR^{-1}B^T P + PA + A^T P]x = 0,$$
$$x^T P(t_f) x = x^T Q_0 x$$

for all $(t, x) \in [0, t_f] \times \mathbf{R}^n$. Hence, P is the solution to the Riccati equation

$$\dot{P} + Q - PBR^{-1}B^T P + PA + A^T P = 0, \quad P(t_f) = Q_0.$$

The optimal feedback law is $\mu(t, x) = -R^{-1}B^T P(t) x$ and the corresponding optimal cost is $V(0, x_0) = x_0^T P(0) x_0$.

Remark 5. We can let Q and R be piecewise continuous functions of time and still get the same result.

Remark 6. It can be established using the results in the next chapter that there exists a solution to the above Riccati equation in the case when $R > 0$ and $Q_0, Q \geq 0$.

3.3 Practical Aspects

Continuous time dynamic programming gives

- sufficient conditions for optimality
- optimal solutions in feedback form

These are both extremely satisfactory properties. However, there are also several drawbacks

- Analytic solutions can only be obtained in a few cases (In particular so called linear quadratic, or LQ, problems).
- The HJB partial differential equation is in general very hard to solve numerically. The main problem is that the full state space must be discretized and a huge number of samples are needed in order to get reasonable solutions. This is called the *curse of dimensionality.*
- It is required that the value function is continuously differentiable, which is not always the case.

3.4 Terminal State Constraint (Optional)

It is possible to consider variations of the optimization problem in (3.1). We will here consider optimization problems where there is a terminal boundary condition $G(t_f, x) = 0$, where $G : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^p$ $(p < n)$. The optimal control problem can be stated as

$$\text{minimize}\, \phi(t_f, x(t_f)) + \int_{t_i}^{t_f} f_0(t, x(t), u(t))dt$$
$$\text{subj. to} \quad \begin{cases} \dot{x} = f(t, x(t), u(t)) \\ x(t_i) = x_i, \;\; G(t_f, x(t_f)) = 0 \\ u(t) \in U, \;\; t_f \geq 0 \end{cases} \tag{3.10}$$

Note that t_f is a free variable in this problem (If t_f is fixed then the t_f dependence of G and ϕ can be removed). Again, let the value function is defined as

$$J^*(t, x) = \min_{u(\cdot)} \left\{ \phi(t_f, x(t_f)) + \int_{t}^{t_f} f_0(t, x(t), u(t))dt \right\}$$

which satisfies the boundary condition $J^*(t_f, x) = \phi(t_f, x)$ on the manifold $G(t_f, x) = 0$. We have the following result

Proposition 3. *Suppose*

(i) V is C^1 (in both arguments) and solves HJBE

$$\begin{aligned} -\frac{\partial V}{\partial t}(t, x) &= \min_{u \in U} \left\{ f_0(t, x, u) + \frac{\partial V}{\partial x}(t, x)^T f(t, x, u) \right\} \\ V(t_f, x) &= \phi(t_f, x) \quad \text{when} \quad G(t_f, x) = 0 \end{aligned} \tag{3.11}$$

(ii) $\mu(t, x) = \arg \min_{u \in U} \left\{ f_0(t, x, u) + \frac{\partial V}{\partial x}(t, x)^T f(t, x, u) \right\}$ *is admissible*

Then

(a) $J^*(t, x) = V(t, x)$ for all $(t, x) \in [t_i, t_f] \times \mathbf{R}^n$.

(b) $\mu(t, x)$ is the optimal feedback control law, i.e. $u^*(t) = \mu(t, x(t))$.

Proof. Completely analogous with the proof of Theorem 3. □

Remark 7 (Stationary systems). If the system dynamics, the cost function, and the boundary conditions are independent of time (and if the final time is still an independent variable) then the optimal control and the value function are also independent of time. The HJBE in (3.11) becomes

$$\begin{aligned} 0 &= \min_{u \in U} \left\{ f_0(x, u) + \frac{\partial V}{\partial x}(x)^T f(x, u) \right\} \\ V(x) &= \phi(x) \quad \text{when} \quad G(x) = 0 \end{aligned} \tag{3.12}$$

The above results appears simple to use but there is a catch. It turns out that even very simple problems may give rise to value functions that do not belong to C. We illustrate this with a simple example.

Example 11. Consider the optimal control problem

$$J^*(x_0) = \text{minimize}\, t_f \quad \text{subj.to} \quad \begin{cases} \dot{x} = -x + u, & x(0) = x_0 \\ x(t_f) = 0 \\ |u| \le 1, & t_f > 0 \end{cases}$$

It is easy to verify that

$$\begin{aligned} \mu(x) &= -\text{sign}(x) \\ J^*(x) &= \ln(1 + \text{sign}(x)x) \end{aligned}$$

We note that $V(x) = J^*(x)$ satisfies HJBE

$$0 = 1 - V'(x)x - |V(x)|, \text{ when } x \neq 0 \text{ and } V(0) = 0.$$

However, $J^* \notin C^1$ so our theory is not valid for this example.

It is possible to relax the assumptions on the value function in order to treat the situation in Example 11 rigorously. We refer to [13, 5] for treatments on this topic.

Chapter 4

Mathematical Preliminaries

We start with some theory for Ordinary Differential Equations (ODE). The chapter also contains a brief review of the first and second order optimality conditions for constrained nonlinear optimization problems.

4.1 ODE Theory

Consider the differential equation $\dot{x}(t) = f(t, x(t))$, $x(t_0) = x_0$. Several questions arise

- When does there exist a solution?
- When is a solution unique?

If there exists a solution then further questions arise

- How does the solution depend on the boundary condition (in this case the initial condition)?
- How does the solution depend on the right hand side of the equation?
- How can we obtain estimates on the size of the solution?

Before we answer any of these questions we illustrate by means of a few examples that these questions are justified.

Example 12 (No solution). Consider the system

$$\dot{x}(t) = \begin{cases} -1, & x(t) \geq 0 \\ 1, & x(t) < 0 \end{cases}$$

The solution disappears to exist when $x(t) = 0$. The problem is that the right hand side is discontinuous. It is not always the case that a discontinuous right

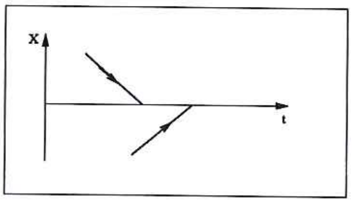

Figure 4.1: There is no solution at $x = 0$

hand side causes any problem and even if it does, it may still be possible to define a solution. However, this requires a more sophisticated machinery than will be treated in this course.

Example 13 (Lack of uniqueness). Consider

$$\dot{x}(t) = x(t)^{1/3}, \quad x(0) = 0$$

Then $x(t) \equiv 0$ is a solution. There are, however, more solutions. Separation of variables gives

$$\frac{dx}{x^{1/3}} = dt \quad \Leftrightarrow \quad \frac{3}{2}x^{2/3} = t - c \quad \Leftrightarrow \quad x = (\frac{2}{3}(t-c))^{3/2} \text{ (if } t - c \geq 0)$$

Hence, the initial value problem $\dot{x}(t) = x(t)^{1/3}, \quad x(0) = 0$ has infinitely many solutions

$$x(t) = \begin{cases} 0, & 0 \leq t \leq c \\ (\frac{2}{3}(t-c))^{3/2}, & t > c \end{cases}$$

where $c \geq 0$. See, Figure 4.2.

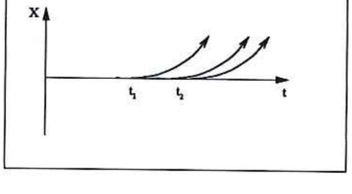

Figure 4.2: There are infinitely many solutions to $\dot{x}(t) = x(t)^{1/3}$, $x(0) = 0$.

The next example shows that sometimes solutions only exist on a finite time interval.

Example 14 (Finite escape time). Consider the ODE

$$\dot{x}(t) = x(t)^{\alpha}, \quad x(0) = x_0 > 0$$

When $x(t) \neq 0$, separation of variables gives

$$\frac{dx}{x^\alpha} = dt \quad \Leftrightarrow \quad x^{1-\alpha}\frac{1}{1-\alpha} = t + c$$

The choice $c = \frac{x_0^{1-\alpha}}{1-\alpha}$ gives $x(0) = x_0$. Hence, if $\alpha > 1$ we get

$$x(t) = \frac{x_0}{(1-\beta x_0^\beta t)^{1/\beta}}, \qquad \beta = \alpha - 1.$$

We see that $x(t) \to \infty$ as $t \to 1/(\beta x_0^\beta)$; See Figure 4.3.

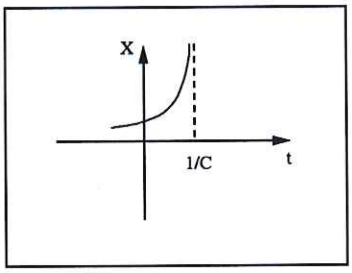

Figure 4.3: Finite escape time. Here $C = \beta x_0^\beta$

From the basic course on differential equations it is known that if the vector field is C^1 then there exists a solution (at least locally). It turns out that Lipschitz continuity of the vector field is a less restrictive condition for existence and uniqueness of a solution to the ODE.

Definition 2. Let $\Omega \subset \mathbf{R}^n$. We say that f is Lipschitz continuous in x on $[t_0, t_1] \times \Omega$ with Lipschitz constant L ($f \in \text{Lip}(L)$) if

$$\|f(t, x_1) - f(t, x_2)\| \leq L\|x_1 - x_2\|$$

for all $t \in [t_0, t_1]$ and $x_1, x_2 \in \Omega$.

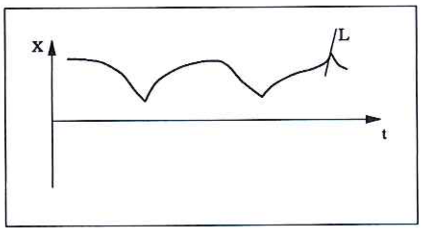

Figure 4.4: The largest slope corresponds to the Lipschitz constant L

We have the following propositions

Proposition 4. *A vector valued function f is Lipschitz continuous if and only if all its components f_k are Lipschitz continuous, i.e.,*

$$f \in Lip \Leftrightarrow f_k \in Lip, \; k = 1, \ldots, n$$

Proof. ($\Leftarrow$)

$$\begin{aligned} |f_k(t,x) - f_k(t,y)| &\leq L_k||x-y|| \Rightarrow \\ ||f(t,x) - f(t,y)||^2 &= \Sigma_1^n |f_k(t,x) - f_k(t,y)|^2 \\ &\leq \Sigma_1^n L_k^2 ||x-y||^2 \\ &= \left(\Sigma_1^n L_k^2\right) ||x-y||^2 \\ &= L^2 ||x-y||^2 \end{aligned}$$

($\Rightarrow$)

$$\begin{aligned} |f_k(t,x) - f_k(t,y)|^2 &\leq \Sigma_1^n |f_k(t,x) - f_k(t,y)|^2 \\ &= ||f(t,x) - f(t,y)||^2 \\ &\leq L^2 ||x-y||^2 \end{aligned}$$

□

The next proposition shows that continuous differentiability of the vector field is sufficient for it to be Lipschitz continuous.

Proposition 5. *Assume that $f(t,x)$ and $\frac{\partial f}{\partial x}(t,x)$ are continuous in a closed and bounded convex set $[t_0, t_1] \times \Omega \subset \boldsymbol{R} \times \boldsymbol{R}^n$. Then f is Lipschitz continuous in x on $[t_0, t_1] \times \Omega$.*

Proof. Let

$$M \equiv \max_{\{i,j=1,\ldots,n\}} \max_{(t,x)\in[t_0,t_1]\times\Omega} \left\{ \left| \frac{\partial f_i}{\partial x_j} \right| \right\}$$

It is clear that $M < \infty$ since all $\frac{\partial f_i}{\partial x_j}$ are continuous on the closed and bounded set $[t_0, t_1] \times \Omega$. For given t, x, y and a variable $s \in [0,1]$ we have for all $k \in \{1, \ldots, n\}$

$$\begin{aligned} \frac{d}{ds} f_k(t, sx + (1-s)y) &= \sum_{j=1}^n \frac{\partial f_k}{\partial x_j}(t, sx + (1-s)y) \frac{d}{ds}(sx_j + (1-s)y_j) \\ &= \sum_{j=1}^n \frac{\partial f_k}{\partial x_j}(t, sx + (1-s)y)(x_j - y_j) \end{aligned}$$

The mean value theorem of calculus says that

$$g(b) - g(a) = g'(c)(b-a)$$

for some $c \in (a, b)$ given that $g \in C[a, b]$ and $g \in C^1(a, b)$. Using this gives

$$\begin{aligned} f_k(t,x) - f_k(t,y) &= f_k(t, 1x + 0y) - f_k(t, 0x + 1y) \\ &= \frac{d}{ds} f_k(t, sx + (1-s)y)|_{s=s_0}(1-0) \\ &= \Sigma_{j=1}^n \frac{\partial f_k}{\partial x_j}(t, s_0 x + (1-s_0)y)(x_j - y_j) \end{aligned}$$

for some $s_0 \in (0, 1)$. Square this expression and use Schwartz inequality to get

$$\begin{aligned} |f_k(t,x) - f_k(t,y)|^2 &\leq \Sigma_{j=1}^n |\frac{\partial f_k}{\partial x_j}(t, s_0 x + (1-s_0)y)|^2 \Sigma_{j=1}^n |(x_j - y_j)|^2 \\ &\leq nM^2 ||x - y||^2 \end{aligned}$$

After summation, we obtain

$$||f(t,x) - f(t,y)||^2 \leq n \cdot nM^2 ||x - y||^2 = L^2 ||x - y||^2$$

which proves the proposition. □

Uniqueness and existence theorems

The following local existence and uniqueness results can be found in, for example, [11, 6, 2].

Theorem 4 (Local existence and uniqueness). *Assume f is piecewise continuous in t and Lipschitz continuous in x on $[t_0, t_1] \times \Omega$, where $\Omega \subset \boldsymbol{R}^n$ is an open and connected set. Then there exists $\delta > 0$ such that*

$$\dot{x}(t) = f(t, x(t)), \qquad x(t_0) = x_0 \quad \text{where } x_0 \in \Omega$$

has a unique solution over $[t_0, t_0 + \delta]$.

Remark 8. The solution can be continued until we reach the boundary of $[t_0, t_1] \times \Omega$.

Remark 9. The vector field in Example 13 is not Lipschitz continuous since we have $|x^{1/3}|/|x| \to \infty$ as $x \to 0$.

Theorem 5 (Global existence and uniqueness). *Assume f is piecewise continuous in t and globally Lipschitz continuous in x for $t \in [t_0, t_1]$ (here $\Omega = \boldsymbol{R}^n$). Assume further that $\|f(t, x_0)\| \leq M$ for $t \in [t_0, t_1]$. Then*

$$\dot{x}(t) = f(t, x(t)), \qquad x(t_0) = x_0$$

has a unique solution over $[t_0, t_1]$.

Remark 10. The system in Example 14 is not globally Lipschitz since $|x^\alpha|/|x| \to \infty$ as $|x| \to \infty$ (when $\alpha > 1$).

Dependence on the right hand side

We will first state a very useful result.

Lemma 1 (Grönwall-Bellman). *Let $\alpha(t), \beta(t)$, where $\beta(t) \geq 0$, and $x(t)$ be continuous scalar valued functions on $[a,b]$ and let $x(t)$ satisfy the inequality*

$$x(t) \leq \alpha(t) + \int_a^t \beta(\tau)x(\tau)d\tau, \;\; t \in [a,b]$$

Then for all $t \in [a,b]$ we have

$$x(t) \leq \alpha(t) + \int_a^t \beta(\tau)g(t,\tau)d\tau$$

where

$$g(t,\tau) = \alpha(\tau)e^{\int_\tau^t \beta(\sigma)d\sigma}$$

Proof. Define

$$F(t) \;\; = \;\; \int_a^t \beta(\tau)x(\tau)d\tau \quad (\Rightarrow \; F(a) = 0)$$

Differentiation with respect to t gives

$$\frac{dF}{dt} = \beta(t)x(t) \leq \beta(t)[\alpha(t) + F(t)]$$

i.e., $\frac{dF}{dt} - \beta(t)F(t) \leq \beta(t)\alpha(t)$. Multiplication with the integrating factor $e^{-\int_a^t \beta(\sigma)d\sigma}$ gives

$$\frac{d}{dt}[e^{-\int_a^t \beta(\sigma)d\sigma}F(t)] \leq \beta(t)\alpha(t)e^{-\int_a^t \beta(\sigma)d\sigma}$$

Integration of both sides now gives

$$e^{-\int_a^t \beta(\sigma)d\sigma}F(t) - e^0F(a) \leq \int_a^t \beta(\tau)\alpha(\tau)e^{-\int_a^\tau \beta(\sigma)d\sigma}$$

which implies

$$F(t) \leq \int_a^t \beta(\tau)\alpha(\tau)e^{\int_\tau^t \beta(\sigma)d\sigma}d\tau$$

where we have used $F(a) = 0$. This shows that

$$x(t) \leq \alpha(t) + \int_a^t \beta(\tau)g(t,\tau)d\tau$$

□

We will next address the question of continuous dependence on the initial conditions and continuous dependence on the right hand side [11].

Theorem 6. *Let $f(t,x)$ be piecewise continuous in t and Lipschitz continuous in x on $[t_0,t_1]\times\Omega$, where $\Omega\subset \boldsymbol{R}^n$ is an open connected set ($f\in \mathrm{Lip}(L)$). Assume $x(\cdot)$ is a solution of the system*

$$\dot{x}=f(t,x), \qquad x(t_0)=x_0.$$

and that $z(\cdot)$ is a solution of the perturbed system

$$\dot{z}=f(t,z)+g(t,z), \qquad z(t_0)=z_0$$

where

$$\|g(t,z)\|\le M, \qquad \forall(t,z)\in[t_0,t_1]\times\Omega$$

If $x(t)\in\Omega$ and $z(t)\in\Omega$ for $t\in[t_0,t_1]$ then

$$\|z(t)-x(t)\|\le e^{L(t-t_0)}\|z_0-x_0\|+\frac{M}{L}(e^{L(t-t_0)}-1)$$

for all $t\in[t_0,t_1]$.

Proof. The solutions satisfy

$$\begin{aligned} x(t)&=x_0+\int_{t_0}^{t_1} f(\tau,x(\tau))d\tau\\ z(t)&=z_0+\int_{t_0}^{t_1}[f(\tau,z(\tau))+g(\tau,z(\tau))]d\tau \end{aligned}$$

Subtract and take norms to get

$$\begin{aligned} \|z(t)-x(t)\| &\le \|z_0-x_0\|+\|\int_{t_0}^{t}[f(\tau,z(\tau))+g(\tau,z(\tau))-f(\tau,x(\tau))]d\tau\|\\ &\le \|z_0-x_0\|+\int_{t_0}^{t}\|(f(\tau,z(\tau))-f(\tau,x(\tau))+g(\tau,z(\tau)\|d\tau\\ &\le \|z_0-x_0\|+\int_{t_0}^{t}\|f(\tau,z(\tau))-f(\tau,x(\tau))\|d\tau+\int_{t_0}^{t}\|g(\tau,z(\tau))\|d\tau\\ &\le \|z_0-x_0\|+\int_{t_0}^{t}L\|z(\tau)-x(\tau)\|d\tau+M(t-t_0) \end{aligned}$$

We can now use Grönwall-Bellman Lemma with

$$\begin{aligned} \alpha(t)&=\|z_0-x_0\|+M(t-t_0)\\ \beta(t)&=L \end{aligned}$$

This gives

$$\|z(t)-x(t)\|\le e^{L(t-t_0)}\|z_0-x_0\|+\frac{M}{L}(e^{L(t-t_0)}-1)$$

which proves the theorem. □

The following corollary will be used in the next chapter.

Corollary 1. *Let $f(t,x,u)$ be piecewise continuous in t and Lipschitz continuous in x and u, i.e.,*

$$\|f(t,x_1,u_1)-f(t,x_2,u_2)\| \leq L(\|x_1-x_2\|+\|u_1-u_2\|)$$

for $[t_0,t_1]\times\Omega_1\times\Omega_2$, where $\Omega_1\subset \boldsymbol{R}^n$ and $\Omega_2\subset\boldsymbol{R}^m$ are open connected sets. Assume $x(\cdot)$ is a solution of the system

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

for a given piecewise continuous $u(\cdot)$ and that $z(\cdot)$ is a solution when we use $u(\cdot)+\delta u(\cdot)$, where $\delta u(t)$ is piecewise continuous with $\|\delta u(t)\|\leq M$. If $x(t)\in\Omega_1$, $z(t)\in\Omega_1$, $u(t)\in\Omega_2$ and $u(t)+\delta u(t)\in\Omega_2$ for $t\in[t_0,t_1]$, then

$$\|z(t)-x(t)\|\leq M(e^{L(t-t_0)}-1)$$

for all $t\in[t_0,t_1]$.

Proof. The perturbed system has the vector field

$$f(t,z,u+\delta u)=f(t,z,u)+g(t,z)$$

where $g(t,z)=f(t,z,u+\delta u)-f(t,z,u)$. Hence,

$$\|g(t,z)\|\leq L\|\delta u\|\leq LM, \quad \forall(t,z)\in[t_0,t_1]\times\Omega_1$$

due to the Lipschitz condition. The result now follows from the previous theorem. □

4.2 Linear Differential Equations

Consider a system of homogeneous linear differential equations of first order:

$$\dot{x}(t)=A(t)x(t), \quad x(t_0)=x_0 \tag{4.1}$$

where A is an $n\times n$ matrix valued (piecewise) continuous function of time. From the above theory we know that this set of equations has a unique solution defined for all time. The solution to (4.1) can be written $x(t)=\Phi(t,t_0)x_0$, where the *transition matrix* is defined by the (matrix) differential equation

$$\frac{\partial\Phi(t,s)}{\partial t}=A(t)\Phi(t,s), \quad \Phi(s,s)=I$$

For time-invariant systems (A is a constant matrix) we have $\Phi(t,s)=e^{A(t-s)}$. The transition matrix satisfies the following properties [14, 3]

1. $\Phi(t,s) = \Phi(t,\tau)\Phi(\tau,s)$, for all t, s, τ
2. $\Phi(t,s)^{-1} = \Phi(s,t)$

The solution to the system equation

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad x(t_0) = x_0$$

is given by the variation of constants formula

$$x(t) = \Phi(t,t_0)x_0 + \int_{t_0}^{t} \Phi(t,s)B(s)u(s)ds.$$

We give two important examples

Example 15. Consider a first order linear differential equation

$$\dot{x} = a(t)x, \quad x(0) = x_0$$

In this case the transition matrix becomes

$$\Phi(t,s) = e^{\int_s^t a(\tau)d\tau}, \tag{4.2}$$

and the solution of the differential equation is $x(t) = \Phi(t,0)x_0$.

Example 16. Consider a linear time invariant system

$$\dot{x} = Ax, \quad x(0) = x_0 \tag{4.3}$$

where A is a constant matrix. Then $\Phi(t,s) = e^{A(t-s)}$. It can be computed in several ways. Two simple ways are

1. Use the Taylor expansion

$$e^{At} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}.$$

   This is useful, for example, if A is diagonal or nilpotent, which means that $A^N = 0$ for some N. One example is $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, which has the matrix exponential

$$e^{At} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.$$

2. Another useful formula is

$$e^{At} = \mathcal{L}^{-1}\{(sI - A)^{-1}\}$$

   where $\mathcal{L}^{-1}$ means the inverse Laplace transform.

4.3 Nonlinear Programming

The first and second order necessary conditions for optimality of nonlinear programs will be reviewed here. We also introduce Newton's method for solving such problems. Many optimal control problems can be treated with completely analogous methods. We will see this in Chapters 4, 7, and 8.

Let

$$G(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{bmatrix}, \qquad (p < n)$$

and consider the nonlinear program

$$\text{minimize}\, f(x) \quad \text{subject to} \quad G(x) = 0 \tag{4.4}$$

where $f : \mathbf{R}^n \to \mathbf{R}$ and $G : \mathbf{R}^n \to \mathbf{R}^p$ are twice continuously differentiable. We assume that the constraint set is regular, i.e., the gradients $\nabla g_k(x)$ are linearly independent for all[1] x such that $G(x) = 0$. This means that the constraint set is an $(n - p)$-dimensional smooth manifold.

The following results can be found in, for example, [15, 21].

First order necessary condition: If x^* is a (locally) optimal solution of (4.4) then there exists a vector of *Lagrange multipliers* $\lambda \in \mathbf{R}^p$ such that

(*i*) $G(x^*) = 0$

(*ii*) $\nabla l(x^*, \lambda) = 0$, where $l(x, \lambda) = f(x) + \lambda^T G(x)$ is the *Lagrangian.*

The first order necessary condition is also called the Lagrange's multiplier rule.

Second order necessary condition: If x^* is a (locally) optimal solution of (4.4) then there exists a vector of *Lagrange multipliers* $\lambda \in \mathbf{R}^p$ such that

(*i*) $G(x^*) = 0$

(*ii*) $\nabla l(x^*, \lambda) = 0$

(*iii*) the Hessian $L(x^*, \lambda) = f_{xx}(x^*) + \sum_{k=1}^{p} \lambda_k \cdot (g_k)_{xx}(x^*)$ is positive semidefinite on the tangent space $M = \{y \in \mathbf{R}^n : \nabla g_k(x^*)^T y = 0, \; k = 1, \ldots, p\}$, i.e.,

$$y^T L(x^*, \lambda) y \geq 0 \quad \forall y \in \{y \in \mathbf{R}^n : \nabla g_k(x^*)^T y = 0, \; k = 1, \ldots, p\}$$

Second order sufficient condition: Suppose there exists x^* and $\lambda \in \mathbf{R}^p$ such that

[1] It is enough that this holds in a neighborhood around the optimal point.

(*i*) $G(x^*) = 0$

(*ii*) $\nabla l(x^*, \lambda) = 0$

(*iii*) $L(x^*, \lambda) = f_{xx}(x^*) + \sum_{k=1}^{p} \lambda_k (g_k)_{xx}(x^*)$ is positive definite on $M = \{y \in \mathbf{R}^n : \nabla g_k(x^*)^T y = 0,\ k = 1, \ldots, p\}$, i.e.,

$$y^T L(x^*, \lambda) y > 0 \quad \forall\, y \in \{y \in \mathbf{R}^n \backslash \{0\} : \nabla g_k(x^*)^T y = 0,\ k = 1, \ldots, p\}.$$

Then x^* is a local minimum of (4.4).

Newton's method

The idea behind Newton's method for solving (4.4) is to solve the linearized version of the first order necessary conditions recursively. This gives the following algorithm:

Step 1 Guess initial values (x^0, λ^0)

Step 2 Solve the system

$$\begin{aligned} \nabla l(x^k, \lambda^k) + L(x^k, \lambda^k)\delta x^k + G_x(x^k)^T \delta\lambda^k &= 0 \\ G(x^k) + G_x(x^k)\delta x^k &= 0 \end{aligned} \tag{4.5}$$

Step 3 Update $x^{k+1} = x^k + \delta x^k$, $\lambda^{k+1} = \lambda^k + \delta\lambda^k$, $k = k + 1$.

Step 4 Stop if $|\nabla l(x^k, \lambda^k)|^2 + |G(x^k)|^2 \leq \epsilon$, otherwise go to step 2.

The second step of the algorithm corresponds to the first order necessary conditions for the quadratic program

$$\min \nabla l(x^k, \lambda^k)\delta x^k + (\delta x^k)^T L(x^k, \lambda^k)\delta x^k \quad \text{subj. to} \quad G(x^k) + G_x(x^k)\delta x^k = 0$$

Optimal control SF2852, KTH

Chapter 5

PMP: A Special Case

We will here derive the Pontryagin minimum principle for the special case when there is no constraint on the control (except that it should - as always - be piecewise continuous), the final time is fixed, and when there is no terminal constraint. The optimal control problem can be stated as

$$\begin{aligned} &\text{minimize}\, \phi(x(t_f)) + \int_{t_i}^{t_f} f_0(t, x(t), u(t))dt \\ &\text{subj. to} \begin{cases} \dot{x}(t) = f(t, x(t), u(t)) \\ x(t_i) = x_i \end{cases} \end{aligned} \tag{5.1}$$

where ϕ is continuously differentiable and the cost function f_0 and the vector field f are continuously differentiable with respect to all variables t, x, and u.

Our derivation will be informal and based on variational arguments of the same type as are used in the classical calculus of variations. We first notice that the constraint can be written as

$$f(t, x(t), u(t)) - \dot{x}(t) = 0$$

This implies that

$$\lambda(t)^T[f(t, x(t), u(t)) - \dot{x}(t)] = 0$$

for an arbitrary differentiable function $\lambda(\cdot)$, which is called the *Lagrange multiplier* or *adjoint variable.* Then also the integral evaluates to zero

$$\int_{t_i}^{t_f} \lambda(t)^T[f(t, x(t), u(t)) - \dot{x}(t)]dt = 0$$

If we add this integral to the cost function then we get the *Lagrangian*

$$\begin{aligned} \tilde{J}(u(\cdot)) =& \phi(x(t_f)) + \int_{t_i}^{t_f} [f_0(t, x(t), u(t)) + \lambda(t)^T(f(t, x(t), u(t)) - \dot{x}(t))]dt \\ =& \phi(x(t_f)) + \int_{t_i}^{t_f} [H(t, x(t), u(t), \lambda(t)) - \lambda(t)^T\dot{x}(t)]dt \end{aligned}$$

where the *Hamiltonian* is defined as

$$H(t,x,u,\lambda) = f_0(t,x,u) + \lambda^T f(t,x,u).$$

Inspired by the Lagrange's multiplier rule we expect to get a necessary condition for optimality by letting the first variation ("first order derivative") of $\tilde{J}$ be zero. So assume that $u^*(\cdot)$ is an optimal control for[1] (5.1) and make a small perturbation

$$u(t) = u^*(t) + \delta u(t)$$

where δu is "small"

$$||\delta u(t)|| < \epsilon, \quad \forall t \in [t_0, t_f]$$

It follows from Corollary 1 in the previous chapter that the distance between the optimal and the perturbed trajectories can be made arbitrarily small by making ϵ sufficiently small, i.e., if $x^*(\cdot)$ and $x^*(\cdot) + \delta x(\cdot)$ are the optimal and perturbed trajectories, respectively, then $\|\delta x(t)\|$ can be made as small as we like in $[t_i, t_f]$. We have

$$\begin{aligned}
\Delta\tilde{J} &= \tilde{J}(u^*(\cdot)+\delta u(\cdot)) - \tilde{J}(u^*(\cdot)) \\
&= \phi(x^*(t_f)+\delta x(t_f)) - \phi(x^*(t_f)) \\
&\quad + \int_{t_i}^{t_f} [H(t,x^*(t)+\delta x(t),u^*(t)+\delta u(t),\lambda(t)) - H(t,x^*(t),u^*(t),\lambda(t))]dt \\
&\quad - \int_{t_i}^{t_f} \lambda(t)^T \frac{d}{dt}(x^*(t)+\delta x(t))dt + \int_{t_i}^{t_f} \lambda(t)^T \frac{d}{dt} x^*(t)dt
\end{aligned}$$

If we make a Taylor series expansion then, (all arguments of H_x and H_u are supressed)

$$\begin{aligned}
\Delta\tilde{J} =& \phi_x(x^*(t_f))^T\delta x(t_f) + \int_{t_i}^{t_f} [H_x^T\delta x(t) + H_u^T\delta u(t) - \lambda(t)^T\frac{d}{dt}\delta x(t)]dt + o(\epsilon) \\
=& \left\{ \left[\lambda(t)^T\delta x(t)\right]_{t_i}^{t_f} = \int_{t_i}^{t_f} [(\frac{d}{dt}\lambda(t)^T)\delta x(t) + \lambda(t)^T\frac{d}{dt}\delta x(t)]dt \right\} \\
=& \phi_x(x^*(t_f))^T\delta x(t_f) + \int_{t_i}^{t_f} [H_x^T\delta x(t) + H_u^T\delta u(t) + (\frac{d}{dt}\lambda(t))^T\delta x(t)]dt \\
&- \left[\lambda(t)^T\delta x(t)\right]_{t_i}^{t_f} + o(\epsilon)
\end{aligned}$$

where we have used integration by parts. Since $x(t_i)$ is fixed we have $\delta x(t_i) = 0$. The first variation $\delta\tilde{J}$ is now obtained by removing the higher order terms in $o(\epsilon)$.

[1] Everything below holds also for the case when $u^*(\cdot)$ is a locally optimal control.

This gives

$$\delta\tilde{J} = (\phi_x(x^*(t_f)) - \lambda(t_f))^T\delta x(t_f) + \int_{t_i}^{t_f} [H_u(t, x^*(t), u^*(t), \lambda(t))]^T \delta u(t)dt$$
$$+ \int_{t_i}^{t_f} [H_x(t, x^*(t), u^*(t), \lambda(t)) + \frac{d}{dt}(\lambda(t))]^T \delta x(t)dt = 0$$

This clearly holds if

$$\dot{\lambda}(t) = -H_x(t, x^*(t), u^*(t), \lambda(t)), \qquad \lambda(t_f) = \phi_x(x^*(t_f))$$
$$H_u(t, x^*(t), u^*(t), \lambda(t)) = 0$$

The arguments we have used can be made rigorous by using functional analysis, see [16]. We now state this version of the Pontryagin minimum principle formally. First recall the definition of the Hamiltonian:

$$H(t, x, u, \lambda) = f_0(t, x, u) + \lambda^T f(t, x, u).$$

We have

Proposition 6. *Let $u^*(\cdot)$ be an optimal control for (5.1) and let $x^*(\cdot)$ be the corresponding trajectory. Then there exists an adjoint variable $\lambda(\cdot)$ such that*

(i) $\dot{\lambda}(t) = -\frac{\partial H}{\partial x}(t, x^*(t), u^*(t), \lambda(t)), \quad \lambda(t_f) = \frac{\partial \phi}{\partial x}(x^*(t_f))$

(ii) $\frac{\partial H}{\partial u}(t, x^*(t), u^*(t), \lambda(t)) = 0$ *for* $t \in [t_i, t_f]$

(iii) $H^*(t) = H(t, x^*(t), u^*(t), \lambda(t))$ *satisfies the relation*

$$H^*(t) = H^*(t_f) - \int_t^{t_f} \frac{\partial H}{\partial t}(t, x^*, u^*, \lambda)|_{t=s} ds$$

for all $t \in [t_i, t_f]$.

Remark 11. The adjoint equation can alternatively be written

$$\dot{\lambda}(t) = -\frac{\partial f_0}{\partial x}(t, x^*(t), u^*(t)) - \frac{\partial f}{\partial x}(t, x^*(t), u^*(t))^T \lambda(t)$$
$$= -\frac{\partial f_0}{\partial x}(t, x^*(t), u^*(t)) - \sum_{k=1}^{n} \frac{\partial f_k}{\partial x}(t, x^*(t), u^*(t))\lambda_k(t)$$

It can also be noted that

$$\dot{x}(t) = \frac{\partial H}{\partial \lambda}(t, x(t), u(t), \lambda(t)).$$

Remark 12. The third condition is particularly useful when f_0 and f are autonomous, i.e., when they do not depend on time ($f_0(x,u)$, $f(x,u)$). Then it reduces to (note that now H do not depend explicitly on time)

$$H(x^*(t), u^*(t), \lambda(t)) = const, \quad t \in [t_i, t_f]$$

To motivate condition *(iii)* we consider the special case when the optimal control is differentiable with respect to time. The result then follows from a simple derivation (all arguments are suppressed for sake of brevity)

$$\begin{aligned}\dot{H}^* &= H_t^* + (H_x^*)^T f + (H_u^*)^T \dot{u} + \dot{\lambda}^T f \\ &= H_t^* + (H_u^*)^T \dot{u} + (H_x^* + \dot{\lambda})^T f = H_t^*\end{aligned}$$

where we used the first and the second equations of the theorem. Integration over time gives the desired result.

Remark 13. The proposition does not distinguish between a maximum or minimum. It merely recognizes stationary points (*extremals*).

Conservative Mechanical Systems

To illustrate the fundamental importance of the above optimality conditions we consider conservative mechanical systems which 'behave optimally" in the sense that a particular integral functional is stationary with respect to the dynamics, see [1] for more details. We will show that the conditions in Proposition 6 lead to the famous Euler-Lagrange equation and a characterization of the motion which states that the sum of the potential and the kinetic energy is constant over time. The example is taken from [4].

Example 17. The following notation is standard in mechanics

- q — Generalized coordinates (positions and angles). This is the state of the mechanical system.
- $\dot{q}$ — Generalized velocity which we also denote u ($\dot{q} = u$).
- $V(q)$ — The potential energy of the system
- $T(q,\dot{q})$ — The kinetic energy, which we assume to be of the form $T(q,\dot{q}) = \dot{q}^T A(q)\dot{q}$, where $A(q)$ is a symmetric matrix.
- $L(q,\dot{q})$ — The Lagrangian $L(q,\dot{q}) = T(q,\dot{q}) - V(q)$

One of the fundamental principles of mechanics ("the principle of least action") states that the equations of motion must be an extremal (stationary solution) of the cost function

$$J(u(\cdot)) = \int_{t_i}^{t_f} L(q(t), u(t))dt \quad \text{subject to} \quad \dot{q}(t) = u(t)$$

Hence, it must satisfy the conditions of Proposition 6. The Hamiltonian becomes

$$H(q, u, \lambda) = L(q, u) + \lambda^T u.$$

The following equations must hold

$$(i)\ \ \dot{\lambda}(t) = -H_q(q(t), u(t), \lambda(t)) = -\frac{\partial L}{\partial q}(q(t), u(t))$$

$$(ii)\ \ H_u(q(t), u(t), \lambda(t)) = 0 \text{ for } t \in [t_i, t_f], \text{ which gives } \lambda(t) = -\frac{\partial L}{\partial u}(q(t), u(t))$$

If we use (ii) in (i) and the fact that $\dot{q} = u$ then we get the famous Euler-Lagrange equations from the Calculus of Variations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}(q(t), \dot{q}(t))\right) - \frac{\partial L}{\partial q}(q(t), \dot{q}(t)) = 0. \tag{5.2}$$

We will next use condition (iii) of Proposition 6 to show that the sum of the potential and kinetic energies is constant over time. To do this note that

$$\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) = \frac{\partial T}{\partial \dot{q}}(q, \dot{q}) = 2A(q)\dot{q}.$$

Hence, by (ii) above we get $\lambda = -\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) = -2A(q)\dot{q}$ and the Hamiltonian becomes

$$H(q, \dot{q}, \lambda) = L(q, \dot{q}) + \lambda^T \dot{q} = T(q, \dot{q}) - V(q) - 2\dot{q}^T A(q)\dot{q} = -(T(q, \dot{q}) + V(q)).$$

Since the system is autonomous (time-independent) condition (iii) of Proposition 6 becomes $H(q(t), \dot{q}(t), \lambda(t)) = const$, which thus means that $T(q(t), \dot{q}(t)) + V(q(t)) = const$.

5.1 Linear Quadratic Control

We will here solve the same LQ control problem as in Chapter 3 but now by using Proposition 6. Let us state the problem again (here we have multiplied with by factor 1/2, which makes the resulting equations cleaner but does not affect the final result)

$$\begin{aligned} &\text{minimize}\ \frac{1}{2}x(t_f)^T Q_0 x(t_f) + \frac{1}{2}\int_0^{t_f} [x^T Q x + u^T R u] dt \\ &\text{subj. to} \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ x(0) = x_0 \end{cases} \end{aligned}$$

where Q_0 and Q are symmetric positive semi-definite matrices and R is a symmetric positive definite matrix.

We start by writing out the Hamiltonian

$$H(t,x,u,\lambda) = \frac{1}{2}(x^TQx + u^TRu) + \lambda^T(Ax+Bu)$$

It is usually wise to start with condition (ii) since it allows us to find an expression for the optimal control. We have

$$H_u(t,x,u,\lambda) = Ru + B^T\lambda = 0$$

which has the unique solution $u = \widetilde{\mu}(t,x,\lambda) = -R^{-1}B^T\lambda$. If we insert this into the system equation and the adjoint equation then we get

$$\begin{aligned} \dot{x}(t) &= H_\lambda(t,x(t),\widetilde{\mu}(t,x(t),\lambda(t)),\lambda(t)) = Ax(t) - BR^{-1}B^T\lambda(t), \quad x(0) = x_0 \\ \dot{\lambda}(t) &= -H_x(t,x(t),\widetilde{\mu}(t,x(t),\lambda(t)),\lambda(t)) = -Qx(t) - A^T\lambda(t), \quad \lambda(t_f) = Q_0x(t_f) \end{aligned}$$

These equations can equivalently be written

$$\begin{bmatrix}\dot{x}(t)\\ \dot{\lambda}(t)\end{bmatrix} = \underbrace{\begin{bmatrix}A & -BR^{-1}B^T\\ -Q & -A^T\end{bmatrix}}_{\mathcal{H}}\begin{bmatrix}x(t)\\ \lambda(t)\end{bmatrix}, \quad \begin{bmatrix}x(0)\\ \lambda(t_f)\end{bmatrix} = \begin{bmatrix}x_0\\ Q_0x(t_f)\end{bmatrix}, \tag{5.3}$$

where the matrix $\mathcal{H}$ is called the *Hamiltonian matrix.* Such matrices have some special properties that will be discussed below. The differential equation in (5.3) is called a Two Point Boundary Value Problem (TPBVP), since some of the boundary conditions pertain to t_i and some to t_f. PMP often gives rise to TPBVPs that in general cannot be solved analytically. We discuss one solution method (shooting) for TPBVPs in Chapter 10.

Fortunately, the TPBVP in (5.3) is easy to solve. Let us introduce the transition matrix $\Phi(t,s) = e^{\mathcal{H}(t-s)}$, with block structure

$$\Phi(t,s) = \begin{bmatrix}\Phi_{11}(t,s) & \Phi_{12}(t,s)\\ \Phi_{21}(t,s) & \Phi_{22}(t,s)\end{bmatrix}$$

where each block matrix has size $n \times n$. Then it follows from (5.3) that the unkown parameters $\lambda(0)$ and $x(t_f)$ must satisfy the linear equation

$$\begin{bmatrix}x_0\\ \lambda(0)\end{bmatrix} = \begin{bmatrix}\Phi_{11}(0,t_f) & \Phi_{12}(0,t_f)\\ \Phi_{21}(0,t_f) & \Phi_{22}(0,t_f)\end{bmatrix}\begin{bmatrix}I\\ Q_0\end{bmatrix}x(t_f) = \begin{bmatrix}\Phi_{11}(0,t_f) + \Phi_{12}(0,t_f)Q_0\\ \Phi_{21}(0,t_f) + \Phi_{22}(0,t_f)Q_0\end{bmatrix}x(t_f).$$

It can be proven that $\Phi_{11}(0,t_f) + \Phi_{12}(0,t_f)Q_0$ is invertible (we will discuss this in more detail in the next subsection), which means that we can solve for $x(t_f)$

in the first equation: $x(t_f) = (\Phi_{11}(0,t_f) + \Phi_{12}(0,t_f)Q_0)^{-1}x_0$. This in turn gives the unknown initial condition for the adjoint variable

$$\lambda(0) = (\Phi_{21}(0,t_f) + \Phi_{22}(0,t_f)Q_0)(\Phi_{11}(0,t_f) + \Phi_{12}(0,t_f)Q_0)^{-1}x_0. \tag{5.4}$$

So far PMP has given us the following information: *If λ is the solution of the adjoint differential equation with the initial condition in (5.4), then $u(t) = -R^{-1}B^T\lambda(t)$ is a candidate for the optimal control.* This is much less than we get from using dynamic programming where we learnt that $u(t) = -R^{-1}B^TP(t)x(t)$ *is the optimal feedback solution. Here P is the solution to a Riccati equation.* It is fortunately possible to show that the two solutions are identical. We do so in the next section by using standard arguments.

Derivation of the Riccati Equation (Optional)

It can be proven that[2] $\Phi_{11}(t,t_f) + \Phi_{12}(t,t_f)Q_0$ is invertible for all $t \in [0,t_f]$. In complete analogy with the derivation of (5.4) we can then obtain

$$\lambda(t) = \underbrace{(\Phi_{21}(t,t_f) + \Phi_{22}(t,t_f)Q_0)(\Phi_{11}(t,t_f) + \Phi_{12}(t,t_f)Q_0)^{-1}}_{P(t)}x(t). \tag{5.5}$$

where we have denoted the total matrix $P(t)$ anticipating that we will obtain the optimal feedback solution $u(t) = -R^{-1}B^T\lambda(t) = -R^{-1}B^TP(t)x(t)$. In order to do this we need to show that $P(t)$ satisfies the Riccati equation. If we differentiate both sides of (5.5) and use (5.3) then we get

$$(-Q - A^TP(t))x(t) = (\dot{P}(t) + PA - P(t)BR^{-1}B^TP(t))x(t)$$

From this we can conclude that[3]

$$\dot{P}(t) + A(t)^TP(t) + P(t)A(t) + Q(t) = P(t)BR^{-1}B^TP(t), \quad P(t_f) = Q_0 \tag{5.6}$$

where the boundary condition is obtained from

$$\begin{aligned} P(t_f) &= (\Phi_{21}(t_f,t_f) + \Phi_{22}(t_f,t_f)Q_0)(\Phi_{11}(t_f,t_f) + \Phi_{12}(t_f,t_f)Q_0)^{-1} \\ &= (0 + Q_0)(I + 0)^{-1} = Q_0 \end{aligned}$$

We have thus proven that PMP gives the same result as HJBE. We have also learnt that the solution to the Riccati equation can be obtained in two ways.

[2]A quick way to prove this is as follows (This is essentially Kalmans argument from his famous paper [10]). The LQ optimal control problem is strictly convex since $R > 0$, $Q \geq 0$, and $Q_0 \geq 0$. This means that there exists a unique optimal solution, which must satisfy PMP. This in turn means that the transition matrix of the closed loop system will be $\Psi(t,t_f) = \Phi_{11}(t,t_f) + \Phi_{12}(t,t_f)Q_0$. A transition matrix is invertible!

[3]To see this we notice that our derivation of the optimal control is not dependent on the initial condition (t_0,x_0). Hence, the same equation would have been obtained for any initial point x_0 and this means that $x(t)$ can be arbitrary.

1. Solve the Riccati differential equation in (5.6)
2. Solve the Hamiltonian system (a linear matrix differential equation)

$$\begin{bmatrix} \dot{X}(t) \\ \dot{\Psi}(t) \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^T \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} X(t) \\ \Psi(t) \end{bmatrix}, \quad \begin{bmatrix} X(t_f) \\ \Psi(t_f) \end{bmatrix} = \begin{bmatrix} I \\ Q_0 \end{bmatrix} \tag{5.7}$$

backwards on $t \in [0, t_f]$. Then $P(t) = \Psi(t)X(t)^{-1}$.

Example 18. Consider the Rocket car problem in Figure 1.3. The problem is to drive the rocket car from rest at position z_0 to rest at position 0. An approximate solution is obtained by considering the following problem with a terminal penalty

$$\text{minimize } \|x(t_f)\|^2 + \int_0^{t_f} u(t)^2 dt \quad \text{subj to} \quad \begin{cases} \dot{x} = Ax + Bu, \\ x(0) = \begin{bmatrix} z_0 \\ 0 \end{bmatrix} \end{cases}$$

where

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The first term in the objective function penalizes deviation of the final state from the zero vector while the second is used to ensure low energy consumption. We use the Hamiltonian system in (5.7) to compute the solution to the Riccati equation

$$\begin{bmatrix} \dot{X} \\ \dot{\Psi} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}}_{\mathcal{H}} \begin{bmatrix} X \\ \Psi \end{bmatrix}, \quad \begin{bmatrix} X(t_f) \\ \Psi(t_f) \end{bmatrix} = \begin{bmatrix} I \\ I \end{bmatrix}$$

which gives

$$\begin{aligned} \begin{bmatrix} X(t) \\ \Psi(t) \end{bmatrix} &= e^{\mathcal{H}(t-t_f)} \begin{bmatrix} X(t_f) \\ \Psi(t_f) \end{bmatrix} \\ &= (I + \mathcal{H}(t-t_f) + \mathcal{H}^2(t-t_f)^2/2 + \mathcal{H}^3(t-t_f)^3/6) \begin{bmatrix} I \\ I \end{bmatrix} \end{aligned}$$

Some calculations gives

$$\begin{bmatrix} X(t) \\ \Psi(t) \end{bmatrix} = \begin{bmatrix} 1 + (t-t_f)^3/6 & t - t_f - (t-t_f)^2/2 \\ (t-t_f)^2/2 & 1 - (t-t_f) \\ 1 & 0 \\ -(t-t_f) & 1 \end{bmatrix}$$

and the optimal feedback solution becomes $u(t) = -R^{-1}B^T P(t)x(t)$, where

$$P(t) = \begin{bmatrix} 1 & 0 \\ -(t-t_f) & 1 \end{bmatrix} \begin{bmatrix} 1+(t-t_f)^3/6 & t-t_f-(t-t_f)^2/2 \\ (t-t_f)^2/2 & 1-(t-t_f) \end{bmatrix}^{-1}$$

$$= \frac{\begin{bmatrix} 1+t_f-t & t_f-t+(t_f-t)^2/2 \\ t_f-t+(t_f-t)^2/2 & 1+(t_f-t)^2+\frac{1}{3}(t_f-t)^3 \end{bmatrix}}{1+t_f-t+(t_f-t)^3/3+(t_f-t)^4/12}$$

$$-R^{-1}B^T P(t) = -\frac{\begin{bmatrix} t_f-t+(t_f-t)^2/2 & 1+(t_f-t)^2+\frac{1}{3}(t_f-t)^3 \end{bmatrix}}{1+t_f-t+(t_f-t)^3/3+(t_f-t)^4/12}$$

A property of Hamiltonian matrices (Optional)

We will show that the eigenvalues of the Hamiltonian matrix introduced above are symmetric with respect to the imaginary axis. This can have severe consequences for the solution of two point boundary problems. The reason is that the transition matrix $\Phi(t_f, 0) = e^{\mathcal{H}t_f}$ will have eigenvalues that are symmetric with respect to the unit circle (if λ is an eigenvalue of $e^{\mathcal{H}t_f}$ then also $1/\lambda$ is an eigenvalue). In particular, if λ is an eigenvalue of $\mathcal{H}$ with very large real part (Re $\lambda >> 0$) then $e^{\mathcal{H}t_f}$ will have one eigenvalue with a very large absolute value, $e^{\lambda t_f}$, and one with a very small absolute value, $e^{-\lambda t_f}$. This tends to make $e^{\mathcal{H}t_f}$ numerically ill conditioned.

Define $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$, which satisfies the property $J^{-1} = J^T$. A matrix is called *Hamiltonian* if it satisfies

$$\mathcal{H}^T J + J\mathcal{H} = 0.$$

It is straightforward to verify that $\mathcal{H}$ in (5.3) satisfies this property.

The Hamiltonian property will now be used to prove that if λ is an eigenvalue of $\mathcal{H}$ then also $-\lambda$ is an eigenvalue of $\mathcal{H}$. Indeed, if we have

$$\mathcal{H}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

then

$$J\mathcal{H}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \lambda \begin{bmatrix} v_2 \\ -v_1 \end{bmatrix}, \quad \text{but also} \quad J\mathcal{H}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = -\mathcal{H}^T J \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = -\mathcal{H}^T \begin{bmatrix} v_2 \\ -v_1 \end{bmatrix}$$

where in the second equation we used $J\mathcal{H} = -\mathcal{H}^T J$. The two right hand sides show that $-\lambda$ is an eigenvalue of $\mathcal{H}^T$. But $\mathcal{H}$ and $\mathcal{H}^T$ have the same eigenvalues so this also means that $-\lambda$ is an eigenvalue of $\mathcal{H}$.

5.2 Derivation Using Dynamic Programming (Optional)

We will in this section use the value function obtained using dynamic programming in order to give an alternative derivation of PMP. We will make the strong assumption that the value function $J^*(t,x)$ is twice continuously differentiable (which not always is the case). What we will learn from this approach is that the adjoint variables can be interpreted as the gradient of the value function with respect to the state vector.

The value function satisfies HJBE

$$J_t^*(t,x) + f_0(t,x,\mu(t,x)) + J_x^*(t,x)^T f(t,x,\mu(t,x)) = 0, \quad \forall (t,x) \in [t_i, t_f] \times \mathbf{R}^n.$$

Differentiation with respect to x and t, respectively, gives

$$J_{xt}^*(t,x) + f_{0x}(t,x,\mu(t,x)) + J_{xx}^*(t,x)^T f(t,x,\mu(t,x)) + f_x(t,x,\mu(t,x))^T J_x^*(t,x) = 0 \tag{5.8}$$

$$J_{tt}^*(t,x) + f_{0t}(t,x,\mu(t,x)) + J_{xt}^*(t,x)^T f(t,x,\mu(t,x)) + J_x^*(t,x)^T f_t(t,x,\mu(t,x)) = 0 \tag{5.9}$$

where

$$J_{xx}^* = \begin{bmatrix} \frac{\partial^2 J}{\partial x_1 x_1} & \cdots & \frac{\partial^2 J}{\partial x_n x_1} \\ \vdots & & \vdots \\ \frac{\partial^2 J}{\partial x_1 x_n} & \cdots & \frac{\partial^2 J}{\partial x_n x_n} \end{bmatrix}, \quad f_x = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

where $J_{xx}^* = (J_{xx}^*)^T$ and $J_{tx}^* = J_{xt}^*$ since J^* was assumed to be in C^2. Let us now evaluate the equations for a particular optimal control and state trajectory $u^*(t) = \mu(t, x^*(t))$, where $\dot{x}^* = f(t, x^*(t))$, $x^*(t_i) = x_i$. If we introduce

$$\lambda(t) = J_x^*(t, x^*(t))$$
$$\lambda_0(t) = J_t^*(t, x^*(t))$$

then

$$\dot{\lambda}(t) = J_{tx}^*(t, x^*(t)) + J_{xx}^*(t, x^*(t))^T f(t, x^*(t), u^*(t)), \quad \lambda(t_f) = \phi_x(x^*(t))$$
$$\dot{\lambda}_0(t) = J_{tt}^*(t, x^*(t)) + J_{xt}^*(t, x^*(t))^T f(t, x^*(t), u^*(t)), \quad \lambda_0(t_f) = 0$$

Using this in (5.8) and (5.9) we get

$$\dot{\lambda}(t) = -f_{0x}(t, x^*(t), u^*(t)) - f_x(t, x^*(t), u^*(t))^T \lambda(t) = -H_x(t, x^*(t), u^*(t), \lambda(t))$$
$$\dot{\lambda}_0(t) = -f_{0t}(t, x^*(t), u^*(t)) - \lambda(t)^T f_t(t, x^*(t), u^*(t)) = -H_t(t, x^*(t), u^*(t), \lambda(t))$$

where the Hamiltonian is defined as in Theorem 6. From this discussion we draw the following conclusions

- the adjoint variable is the gradient of the value function with respect to the state vector.
- In dynamic programming the value function is obtained by solving a partial differential equation (HJBE). This is a consequence of the approach of looking for an optimal control from any initial point.
- In PMP we only solve for the value function (or rather its gradient which is the adjoint variable) for a special initial condition. This gives a two point boundary problem for ODEs, which is normally much simpler to solve than the HJBE.

Chapter 6

PMP: General Results

We will in this chapter discuss the Pontryagin minimum principle for several cases. The derivation will be done in exactly the same spirit as in the original book by Pontryagin and his colleagues [23]. We start with their simplest case when the system is autonomous and when the initial and final states are fixed.

6.1 Autonomous Systems: Fixed Initial and Final States

We consider the optimization problem

$$J^* = \text{minimize} \int_0^{t_f} f_0(x(t), u(t))dt \quad \text{subj. to} \quad \begin{cases} \dot{x}(t) = f(x(t), u(t)) \\ x(0) = x_i, \ x(t_f) = x_f \\ u(t) \in U, \ t_f \geq 0 \end{cases} \tag{6.1}$$

where the following assumptions hold

- $f_0 : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$ and $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n$ are continuous and continuously differentiable with respect to x.
- The final time t_f is free. This means that it is a variable that must be optimized.

It is important to note that it is no restriction to let the initial time be $t_i = 0$. The reason is that an autonomous system allows arbitrary time translation of its solutions. In other words, if $u(t) \in U$ transfers $x(0) = x_i$ to $x(t_f) = x_f$, then $\tilde{u}(t) = u(t - t_i)$ transfers $\tilde{x}(t_i) = x_i$ to $\tilde{x}(t_i + t_f) = x_f$ and the trajectories are related as $\tilde{x}(t) = x(t - t_i)$. Moreover, for the cost integral we have the relation

$$\int_{t_i}^{t_f + t_i} f_0(\tilde{x}(t), \tilde{u}(t))dt = \int_0^{t_f} f_0(x(t), u(t))dt.$$

We will next follow Pontryagin and reformulate (6.1) by introducing an additional state variable

$$x_0(t) = \int_0^t f_0(x(s), u(s))ds,$$

with initial condition $x_0(0) = 0$. This implies that $\dot{x}_0(t) = f_0(x(t), u(t))$. Now define the extended state vector and the extended vector field as

$$\tilde{x} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad \tilde{f} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix}. \tag{6.2}$$

The initial extended state is $\tilde{x}_i = \begin{bmatrix} 0 & x_i^T \end{bmatrix}^T$ and the final extended state must belong to the line $\pi = \mathbf{R} \times \{x_f\}$. This shows that an equivalent formulation of (6.1) is

$$J = \text{minimize}\, x_0(t_f) \quad \text{subj. to} \quad \begin{cases} \dot{\tilde{x}}(t) = \tilde{f}(\tilde{x}(t), u(t)) \\ \tilde{x}(0) = \tilde{x}_i, \ \tilde{x}(t_f) \in \pi = \mathbf{R} \times \{x_f\} \\ u(t) \in U, \ t_f \geq 0 \end{cases}$$

which is illustrated in Figure 6.1. Let us introduce the extended vector $\tilde{\lambda} = \begin{bmatrix} \lambda_0 & \lambda_1 & \dots & \lambda_n \end{bmatrix}^T$, which satisfies the adjoint system

$$\dot{\lambda}_l(t) = -\sum_{k=0}^{n} \frac{\partial f_k(x(t), u(t))}{\partial x_l} \lambda_k(t) \quad l = 0, 1, \dots, n. \tag{6.3}$$

In particular, we have $\dot{\lambda}_0 = 0$, which shows $\lambda_0(t) = const.$ If we define the *Hamiltonian* function

$$H(x, u, \tilde{\lambda}) = \tilde{\lambda}^T \tilde{f}(x, u) = \sum_{k=0}^{n} \lambda_k f_k(x, u) = \lambda_0 f_0(x, u) + \lambda^T f(x, u)$$

then we have

$$\begin{aligned} \dot{\tilde{x}}(t) &= H_{\tilde{\lambda}}(x(t), u(t), \tilde{\lambda}(t)) = \tilde{f}(x(t), u(t)) \\ \dot{\tilde{\lambda}}(t) &= -H_{\tilde{x}}(x(t), u(t), \tilde{\lambda}(t)) = -\tilde{f}_{\tilde{x}}(x(t), u(t))^T \tilde{\lambda}(t) \end{aligned} \tag{6.4}$$

This is the so called *Hamiltonian* system of differential equations. It is important in the following Pontryagin minimum principle for problem (6.1).

Theorem 7 (PMP). *Suppose $(x^*(t), u^*(t), t_f^*)$ is an optimal solution of (6.1), i.e. it transfers x_i to x_f with minimum cost at the optimal transition time t_f^*. Then there exists a nonzero extended adjoint function $\tilde{\lambda}(\cdot)$ such that*

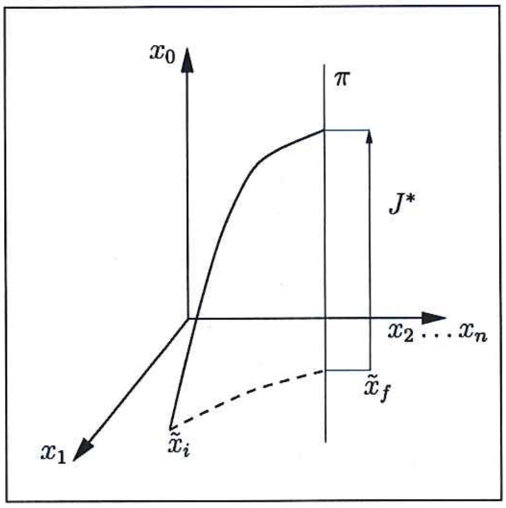

Figure 6.1: The dashed line corresponds to the state trajectory $x(t)$ and the solid line corresponds to the extended state trajectory $\tilde{x}(t)$. The goal of the optimization is to find a control such that the corresponding extended state trajectory intersects the line $\pi = \mathbf{R} \times \{x_f\}$ at a point with as low x_0 coordinate as possible. The minimum x_0 coordinate corresponds to the optimal cost J^*.

(*i*) $\dot{\tilde{\lambda}}(t) = -H_{\tilde{x}}(x^*(t), u^*(t), \tilde{\lambda}(t))$

(*ii*) $H(x^*(t), u^*(t), \tilde{\lambda}(t)) = \min_{v \in U} H(x^*(t), v, \tilde{\lambda}(t)) = 0$ *for all* $t \in [0, t_f^*]$

(*iii*) $\lambda_0(t) = const \geq 0$

Remark 14. Note that

- the Hamiltonian function does not depend on x_0.
- the adjoint equation is a linear differential equation.

Remark 15. In [22] and the classical references [23, 18] the sign of λ_0 is changed and the minimization in (*i*) is replaced by a maximization. This gives a completely analogous result. Our sign convention is chosen in consistency with the results in Chapter 2 and Chapter 4.

Remark 16. If we change the optimization problem (6.1) such that t_f is fixed but everything else remains the same then we obtain an identical theorem except that condition (*ii*) is replaced by

(*ii*) $H(\tilde{x}^*(t), u^*(t), \tilde{\lambda}(t)) = \min_{v \in U} H(\tilde{x}^*(t), v, \tilde{\lambda}(t)) = const$ for all $t \in [0, t_f]$.

where the constant can be any real number. This modification is easily seen from the last step of the proof below.

Remark 17. The λ_0 coordinate did not appear in the PMP theorem in the previous section. In most situations we will have $\lambda_0 > 0$ and then it is no restriction to let $\lambda_0 = 1$. This follows since the adjoint equation is linear so we can multiply it by any positive number. The cases when $\lambda_0 = 0$ are pathological and often due to lack of controllability or other related problems. Note also that $\lambda_0 = 0$ means that the cost integral does not affect the criterion. For the most part of the course we simply ignore the case when $\lambda_0 = 0$.

Remark 18. Since we know that λ_0 is a constant, we often replace the extended adjoint equation in (i) with

$$\dot{\lambda}(t) = -H_x(x^*(t), u^*(t), \tilde{\lambda}(t)) = -\frac{\partial f_0}{\partial x}(x^*(t), u^*(t))\lambda_0 - \frac{\partial f}{\partial x}(x^*(t), u^*(t))^T\lambda(t)$$

In our examples we often assume $\lambda_0 = 1$.

Remark 19. The conditions are necessary conditions for stationarity and we can use them to find candidates for optimality. We discuss this in more detail in section 6.4

Before we embark on a discussion of the proof of Theorem 7 we consider an example.

Example 19. We will solve the linear quadratic optimal control problem for the rocket car in Chapter 1. The optimization problem is (note that t_f is fixed in this example)

$$J^* = \text{minimize} \int_0^{t_f} u(t)^2 dt \quad \text{subj to} \quad \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ x(0) = x_0, \ x(t_f) = 0 \end{cases}$$

This problem was solved in the basic course [14] using completion of squares. Here we solve it using PMP. The final time is fixed so we need to use Remark 16.

We start by minimizing the Hamiltonian with respect to u. The Hamiltonian becomes

$$H(x, u, \tilde{\lambda}) = u^2 + \lambda^T(Ax + Bu),$$

(normalizing $\lambda_0 = 1$). We now have $\text{argmin}_u H(x, u, \tilde{\lambda}) = \text{argmin}_u\{u^2 + \lambda^T(Ax + Bu)\} = -\frac{1}{2}B^T\lambda$. The adjoint equation is

$$\dot{\lambda}(t) = -A^T\lambda(t)$$

which has the solution $\lambda(t) = e^{-A^Tt}\lambda(0)$. The initial value $\lambda(0)$ can be determined using the terminal condition for x. When $u(t) = -\frac{1}{2}B^T\lambda(t)$ is substituted into the state equation we get

$$\dot{x}(t) = Ax(t) - \frac{1}{2}BB^Te^{-A^Tt}\lambda(0), \quad x(0) = x_0$$

By the variation of constants formula we have

$$\begin{aligned} x(t_f) &= e^{At_f}x_0 - \frac{1}{2}\int_0^{t_f} e^{A(t_f-s)}BB^Te^{-A^Ts}ds\lambda(0) \\ &= e^{At_f}x_0 - \frac{1}{2}W(t_f,0)e^{-A^Tt_f}\lambda(0) = 0 \end{aligned}$$

where the reachability Grammian is

$$W(t_f,0) = \int_0^{t_f} e^{A(t_f-s)}BB^Te^{A^T(t_f-s)}ds.$$

In our case the system is controllable and therefore $W(t_f,0)$ is positive definite and thus invertible. We can solve for $\lambda(0)$, which gives

$$\lambda(0) = 2e^{A^Tt_f}W(t_f,0)^{-1}e^{At_f}x_0$$

This gives the optimal control

$$u(t) = -\frac{1}{2}B^Te^{-A^Tt}\lambda(0) = -B^Te^{A^T(t_f-t)}W(t_f,0)^{-1}e^{At_f}x_0$$

and the optimal cost becomes (after some calculations)

$$\begin{aligned} J^* &= \int_0^{t_f} u(t)^2dt = \frac{1}{4}\int_0^{t_f} \lambda(t)^TBB^T\lambda(t)dt \\ &= x_0^Te^{A^Tt_f}W(t_f,0)^{-1}e^{At_f}x_0. \end{aligned}$$

Proof of Theorem 7

We will in this section discuss the proof of Theorem 7. A quick look at the proof in one of the references [23, 18, 22] may give the impression that it is very complicated. However, it turns out that the ideas behind the proof are very elegant and quite easy to understand. What makes things complicated is that $u(\cdot)$ is allowed to have discontinuities since it is piecewise continuous. This will mess things up quite a bit since we loose smoothness at the points of discontinuity. Here we ignore all these technical details in order to highlight the main ideas. We refer to [23] for the additional technical arguments.

First let us recall what we have set out to prove (Consider Figure 6.1): *Among the admissible controls, $u^*(\cdot)$ is one that transfers $\tilde{x}(0) = (0, x_i)$ to a point on the line $\pi = \boldsymbol{R} \times \{x_f\}$ with lowest possible x_0 coordinate. Show that the conditions of Theorem 7 are necessary for $(x^*(\cdot), u^*(\cdot), t_f^*)$ to be optimal in this sense.*

Step 1: Perturbation of optimal control and transition time: Suppose $(x^*(\cdot), u^*(\cdot), t_f^*)$ is an optimal solution. We will make small variations in the optimal control $u^*(\cdot)$ and the optimal transfer time t_f^*. Any such change must lead to a larger (or equal) value of the cost (i.e., an intersection of π at a larger (or equal) x_0 coordinate).

Basic perturbation of control: First consider a small (in time) perturbation at some time τ, where $u^*(\cdot)$ is continuous.

$$u(t) = \begin{cases} u^*(t), & t \leq \tau - \Delta\sigma \\ v, & \tau - \Delta\sigma \leq t \leq \tau \\ u^*(t), & t \geq \tau \end{cases} \tag{6.5}$$

where $v \in U$. Note here that only $\Delta\sigma$ is assumed to be small and v may be very different from $u^*(\tau)$.

The question is how much such a perturbation can affect the cost $x_0(t_f)$. Let us first study the local effect of the perturbation. A first order approximation gives

$$\begin{aligned} \delta\tilde{x}(\tau) = \tilde{x}(\tau) - \tilde{x}^*(\tau) &= \int_{\tau-\Delta\sigma}^{\tau} [\tilde{f}(x(t), v) - \tilde{f}(x^*(t), u^*(t))]dt \\ &= (\tilde{f}(x^*(\tau), v) - \tilde{f}(x^*(\tau), u^*(\tau)))\Delta\sigma + o(\Delta\sigma) \end{aligned}$$

We will next show how this perturbation can be transported to the optimal end point $x^*(t_f^*)$ using the linearized dynamics.

Transportation of basic perturbation: We have seen that a perturbation of the optimal control on the form (6.5) to first order corresponds to a change $\tilde{x}(\tau) = \tilde{x}^*(\tau) + \delta\tilde{x}(\tau)$ of the state vector at time τ, where

$$\delta\tilde{x}(\tau) = (\tilde{f}(x^*(\tau), v) - \tilde{f}(x^*(\tau), u^*(\tau)))\Delta\sigma.$$

We now want to show how this perturbation is transported to the optimal final time t_f^*. We can view $\delta\tilde{x}(\tau)$ as a perturbation in the initial condition at time τ. It then follows from Theorem 6 in Chapter 4 that $\delta\tilde{x}(t_f^*) = \tilde{x}(t_f^*) - \tilde{x}^*(t_f^*) = O(\delta\tilde{x}(\tau))$. This motivates us to consider the linearized dynamics. A Taylor expansion gives

$$\begin{aligned} \delta\dot{\tilde{x}}(t) = \dot{\tilde{x}}(t) - \dot{\tilde{x}}^*(t) &= \tilde{f}(x(t), u^*(t)) - \tilde{f}(x^*(t), u^*(t)) \\ &= \tilde{f}_{\tilde{x}}(x^*(t), u^*(t))\delta\tilde{x}(t) + o(\delta\tilde{x}(t)), \quad t \in [\tau, t_f^*] \end{aligned}$$

We have thus shown that to first order approximation, the perturbation is transported by the linear system

$$\delta\dot{\tilde{x}}(t) = A(t)\delta\tilde{x}(t), \quad \delta\tilde{x}(\tau) = (\tilde{f}(x^*(\tau), v) - \tilde{f}(x^*(\tau), u^*(\tau)))\Delta\sigma$$

where $A(t) = \tilde{f}_{\tilde{x}}(x^*(t), u^*(t))$. Hence, we have $\delta\tilde{x}(t_f^*) = \Phi(t_f^*, \tau)\delta\tilde{x}(\tau)$, where Φ is the transition matrix corresponding to $A(t)$. Figure 6.2 illustrates how the

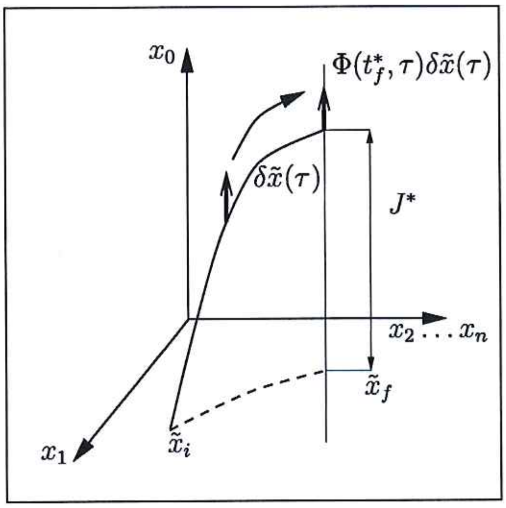

Figure 6.2: Transportation of the perturbation (more precisely, the first order approximation) $\delta\tilde{x}(\tau)$ to the final point using the linearized dynamics. The perturbations are illustrated as thick arrows.

perturbation $\delta\tilde{x}(\tau)$ is moved to a perturbation $\Phi(t_f^*, \tau)\delta\tilde{x}(\tau)$, which has its vertex at the optimal point $x^*(t_f^*)$.
Perturbation of the final time: We can perturb the final time by either extending the optimal control to time $t_f^* + \Delta t$ (if $\Delta t \geq 0$)

$$u(t) = \begin{cases} u^*(t), & t \in [0, t_f^*] \\ u^*(t_f^*), & t \in [t_f^*, t_f^* + \Delta t] \end{cases}$$

or by stopping earlier (if $\Delta t \leq 0$)

$$u(t) = u^*(t_f^* + \Delta t), \quad t \in [0, t_f^* + \Delta t]$$

Hence, a perturbation of the final time to $t_f^* + \Delta t$ gives rise to the following perturbation of the end point

$$\delta\tilde{x}(t_f^*) = \tilde{x}(t_f^* + \Delta t) - \tilde{x}^*(t_f^*) = \int_{t_f^*}^{t_f^* + \Delta t} \tilde{f}(x(t), u(t))dt = \tilde{f}(x^*(t_f^*), u^*(t_f^*))\Delta t + o(\Delta t)$$

where Δt may be positive or negative.
Combined perturbation: Let us combine perturbations on the form above, see Figure 6.3. Each of the individual perturbations will contribute to the perturbation at the end point in the way described above. The first order contribution at the end point can then be obtained by adding up the individual contributions.

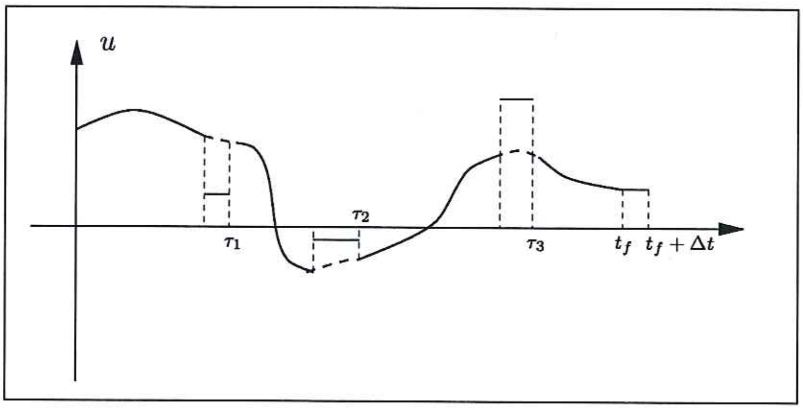

Figure 6.3: A more general perturbation of $u^*(\cdot)$. Here $u^*(\cdot)$ is drawn in dashed lines and the perturbed control is in solid line.

Let us assume that we perturb the control at times $0 < \tau_1 \leq \tau_2 \leq \ldots \leq \tau_{p-1} \leq \tau_p < t_f^*$ and that the control value at each perturbation point is $v_k \in U$. If we choose the intervals of perturbation to be $I_k = [\tau_k - \alpha_k \Delta\sigma, \tau_k]$, where the $\alpha_k \geq 0$ are such that the intervals are disjoint then we get the perturbed control

$$u(t) = \begin{cases} v_k, & t \in I_k, \quad v_k \in U \\ u^*(t), & t \notin I_k \cup [t_f^*, t_f^* + \Delta t] \\ u^*(t_f^*), & t \in [t_f^*, t_f^* + \Delta t] \end{cases}$$

and similarly if $\Delta t < 0$. It can be shown that all possible perturbed controls of this type give rise to the following set of end point perturbations (follows from superposition of the individual perturbations at point τ_k and at the end point t_f^*).

$$\mathcal{K}(t_f^*) = \{\tilde{f}(x^*(t_f^*), u^*(t_f^*))\Delta t + \sum_{k=1}^{p} \alpha_k \Phi(t_f^*, \tau_k)\delta\tilde{x}_k : \Delta t \in \mathbf{R};\ \alpha_k \geq 0;\ \tau_k \in (0, t_f^*)$$
$$p \text{ is an integer, } \delta\tilde{x}_k = (\tilde{f}(x^*(\tau_k), v_k) - \tilde{f}(x^*(\tau_k), u^*(\tau_k)))\Delta\sigma,\ v_k \in U\}$$

This set is a convex cone i.e., if $z_1, z_2 \in \mathcal{K}(t_f^*)$, then also $\beta_1 z_1 + \beta_2 z_2 \in \mathcal{K}(t_f^*)$ for any $\beta_1, \beta_2 \geq 0$. We will pictorially illustrate such cones as "ice cream cones" but it should be understood that they can have quite different appearance in reality.

Step 2: Separating hyperplane at the endpoint: We have seen that the set of end point perturbations is a convex cone $\mathcal{K}(t_f^*)$ with vertex at $\tilde{x}^*(t_f^*)$ (i.e, we consider the set $\mathcal{K}(t_f^*) + \tilde{x}^*(t_f^*)$. Now consider the ray $r = \{x_0 : x_0 \leq$

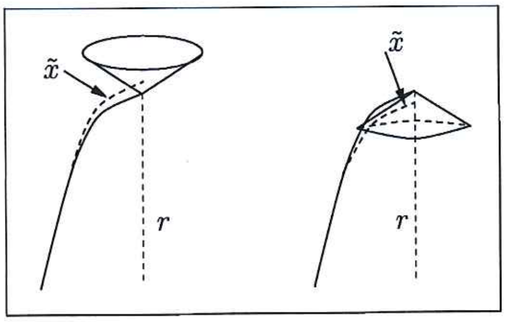

Figure 6.4: In the left hand side the ray r does not contain any points of the cone $\mathcal{K}(t_f^*)$. Then every admissible perturbation of the control and/or transition time will give a trajectory $\tilde{x}(\cdot)$ with endpoint above the optimal point. In the right hand side the cone $\mathcal{K}(t_f^*)$ intersects the ray r and then it is possible to perturb the control and/or the transition time such that the corresponding trajectory $\tilde{x}(\cdot)$ intersects the ray below $\tilde{x}^*(t_f^*)$. However, this contradicts the optimality of $\tilde{x}^*(\cdot)$ so this case is impossible.

$x_0^*(t_f^*)\} \times \{x_f\}$, which consists of all points below the optimal end point. It is intuitively clear that this ray cannot contain any points of $\mathcal{K}(t_f^*) + \tilde{x}^*(t_f^*)$, because that would contradict optimality. It is possible to give a mathematical proof of this but we just give a pictorial illustration of this fact in Figure 6.4. Since the ray r and the cone $\mathcal{K}(t_f^*)$ are both convex and without any common points it follows that there exists a separating hyperplane. In other words, there exists a nonzero vector a such that

$$\begin{aligned} &a^T z \geq 0, \quad \forall z \in \mathcal{K}(t_f^*) \\ &a^T z \leq 0, \quad \forall z \in r - \tilde{x}^*(t_f^*) = \{(z_0, 0, \ldots, 0) : z_0 \leq 0\}. \end{aligned} \tag{6.6}$$

This is illustrated in Figure 6.5. It is important to note that the first coordinate of a must be positive, i.e., $a_0 \geq 0$. This follows since otherwise the second condition in (6.6) cannot hold.

Step 3: Proof of $H(\tilde{x}^*(t), u^*(t), \tilde{\lambda}(t)) = \min_{v \in U} H(\tilde{x}^*(t), v, \tilde{\lambda}(t))$: Let the end condition for the adjoint differential equation be $\tilde{\lambda}(t_f^*) = a$, i.e.,

$$\dot{\tilde{\lambda}}(t) = -A(t)^T \tilde{\lambda}(t), \quad \tilde{\lambda}(t_f^*) = a$$

where $A(t) = \tilde{f}_{\tilde{x}}(x^*(t), u^*(t))$. Then we can easily show that $\tilde{\lambda}(t) = \Phi(t_f^*, t)^T a$.

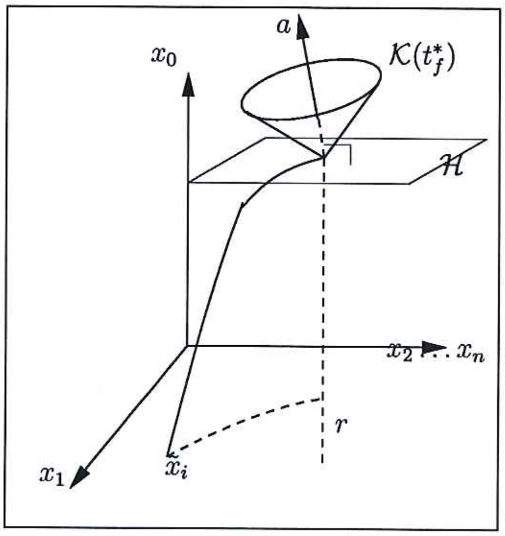

Figure 6.5: The perturbation cone $\mathcal{K}(t_f^*)$ has its vertex at the optimal point $x^*(t_f^*)$. The vector a is the normal of the separating hyperplane.

With this choice of $\tilde{\lambda}$ we get

$$\begin{aligned} &H(x^*(\tau_k), v_k, \tilde{\lambda}(\tau_k)) - H(x^*(\tau_k), u^*(\tau_k), \tilde{\lambda}(\tau_k)) \\ &\quad = \tilde{\lambda}(\tau_k)^T[\tilde{f}(x^*(\tau_k), v_k) - \tilde{f}(x^*(\tau_k), u^*(\tau_k))] \\ &\quad = a^T\Phi(t_f^*, \tau_k)[\tilde{f}(x^*(\tau_k), v_k) - \tilde{f}(x^*(\tau_k), u^*(\tau_k))] = a^T z \geq 0 \end{aligned}$$

since $z = \Phi(t_f^*, \tau_k)[\tilde{f}(x^*(\tau_k), v_k) - \tilde{f}(x^*(\tau_k), u^*(\tau_k))] \in \mathcal{K}(t_f^*)$. This proves that

$$H(x^*(t), u^*(t), \tilde{\lambda}(t)) = \min_{v \in U} H(x^*(t), v, \tilde{\lambda}(t))$$

because $v_k \in U$ and $\tau_k \in (0, t_f^*)$ are arbitrary.

Note also that the condition $\lambda_0 \geq 0$ also holds since we have already concluded that $a_0 \geq 0$.

Step 4: Proof of $H(x^*(t), u^*(t), \tilde{\lambda}(t)) = 0$ for $t \in [0, t_f^*]$: As a first step we show that the condition holds at the endpoint, i.e., $H(x^*(t_f^*), u^*(t_f^*), \tilde{\lambda}(t_f^*)) = 0$. To do this we notice that $z = \tilde{f}(x^*(t_f^*), u^*(t_f^*))\Delta t \in \mathcal{K}(t_f^*)$, for all Δt. Hence, in order for the separating hyperplane condition

$$a^T z = a^T \tilde{f}(x^*(t_f^*), u^*(t_f^*))\Delta t \geq 0$$

to hold for all Δt, we need $a^T \tilde{f}(x^*(t_f^*), u^*(t_f^*)) = 0$. In other words, the normal to the hyperplane in Figure 6.5 must be perpendicular to $\tilde{f}(x^*(t_f^*), u^*(t_f^*))$. This gives,

$$H(x^*(t_f^*), u^*(t_f^*), \tilde{\lambda}(t_f^*)) = \tilde{\lambda}(t_f^*)^T \tilde{f}(x^*(t_f^*), u^*(t_f^*)) = a^T \tilde{f}(x^*(t_f^*), u^*(t_f^*)) = 0.$$

The next step is to show why $H(x^*(t), u^*(t), \tilde{\lambda}(t))$ must be a constant function of time. This constant must of course be 0 since we just proved that this is the value at t_f^*. We only prove this for a special case when

1. $u^*(\cdot)$ is continuously differentiable in a neighborhood around t and it belongs to the interior of U.

2. f_0 and f are C^1 also with respect to u.

For this special case differentiation gives (we suppress the arguments for brevity)

$$\frac{d}{dt} H(x^*(t), u^*(t), \tilde{\lambda}(t)) = \frac{\partial H}{\partial \tilde{x}}^T \dot{\tilde{x}}^* + \frac{\partial H}{\partial \tilde{\lambda}}^T \dot{\tilde{\lambda}} + \frac{\partial H}{\partial u}^T \dot{u}^* = -(\dot{\tilde{\lambda}})^T \dot{\tilde{x}} + (\dot{\tilde{x}})^T \dot{\tilde{\lambda}} = 0$$

where we used (6.4) and that $H_u(x^*(t), u^*(t), \tilde{\lambda}(t)) = 0$ (which corresponds to the pointwise minimization). This shows that the Hamiltonian $H(x^*(t), u^*(t), \tilde{\lambda}(t))$ is constant along the optimal trajectory in this special case. The proof is more complicated in general and we refer to any of [23, 18, 22] for the details.

6.2 Optimal Control to a Manifold

We will in this section consider what happens with the optimality conditions in Theorem 7 when the terminal state is required to belong to a manifold S_f. An $(n-p)$-dimensional smooth *manifold* in $\mathbf{R}^n$ is an intersection of p *hypersurfaces* (each described by a set $S_k = \{x \in \mathbf{R}^n : g_k(x) = 0\}$):

$$S_f = \{x \in \mathbf{R}^n : G(x) = 0\} \quad \text{where} \quad G(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{bmatrix}$$

where the gradients $\nabla g_k(x)$ are linearly independent for all points on the manifold[1]. This is equivalent to requiring that the functional matrix

$$G_x(x) = \begin{bmatrix} \frac{\partial g_1(x)}{\partial x_1} & \cdots & \frac{\partial g_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_p(x)}{\partial x_1} & \cdots & \frac{\partial g_p(x)}{\partial x_n} \end{bmatrix}$$

[1]It is actually enough that the gradients are linearly independent in a neighborhood around the optimal terminal point on S_f.

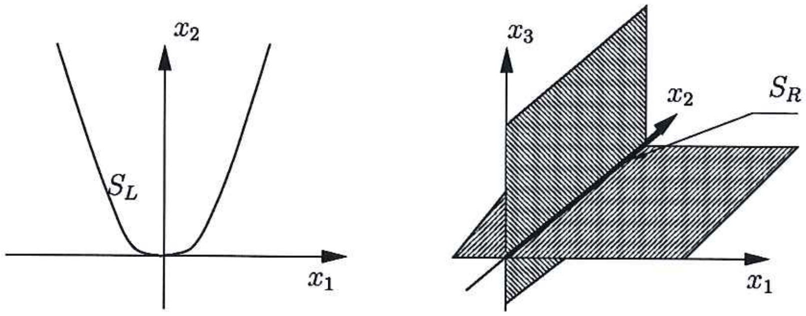

Figure 6.6: The left figure shows the manifold $S_L = \{x \in \mathbf{R}^2 : x_2 - x_1^2 = 0\}$. The right figure shows the manifold $S_R = \{x \in \mathbf{R}^3 : x_1 = 0,\ x_3 = 0\}$, which is the x_2-axis. It can be viewed as the intersection of the manifolds (planes) $S_1 = \{x \in \mathbf{R}^3 : x_1 = 0\}$ and $S_3 = \{x \in \mathbf{R}^3 : x_3 = 0\}$.

has rank p for all $x \in S_f$.

Two examples of manifolds are given in Figure 6.6. The left figure shows the manifold $S_L = \{x \in \mathbf{R}^2 : g(x) = 0\}$, where $g(x) = x_2 - x_1^2$. In this case we have $\nabla g(x) = \begin{bmatrix} -2x_1 \\ 1 \end{bmatrix} \neq 0$ for all $x \in S_L$, which means that the rank condition on the functional matrix is satisfied. The right part of Figure 6.6 shows the manifold $S_R = \{x \in \mathbf{R}^3 : g_1(x) = 0,\ g_2(x) = 0\}$, where $g_1(x) = x_1$ and $g_2(x) = x_3$. The functional matrix becomes

$$\begin{bmatrix} \frac{\partial g_1(x)}{\partial x_1} & \frac{\partial g_1(x)}{\partial x_2} & \frac{\partial g_1(x)}{\partial x_3} \\ \frac{\partial g_2(x)}{\partial x_1} & \frac{\partial g_2(x)}{\partial x_2} & \frac{\partial g_2(x)}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which has full rank (rank=2).

The optimal control problem is formulated as

$$J^* = \text{minimize} \int_0^{t_f} f_0(x(t), u(t))dt \quad \text{subj. to} \quad \begin{cases} \dot{x}(t) = f(x(t), u(t)) \\ x(0) = x_i,\ x(t_f) \in S_f \\ u(t) \in U,\ t_f \geq 0 \end{cases} \tag{6.7}$$

where everything is defined as in the previous section. The new problem differs from the old one in (6.1) only in that we now need to determine the optimal position at the terminal manifold S_f. It turns out that the only change to the optimality conditions in Theorem 7 is that the boundary value of the adjoint vector must be perpendicular to the manifold S_f. In order to show that this is true we take the same approach as before and introduce the extended state vector $\tilde{x}$ and the extended vector field $\tilde{f}$, see equation (6.2). We also introduce

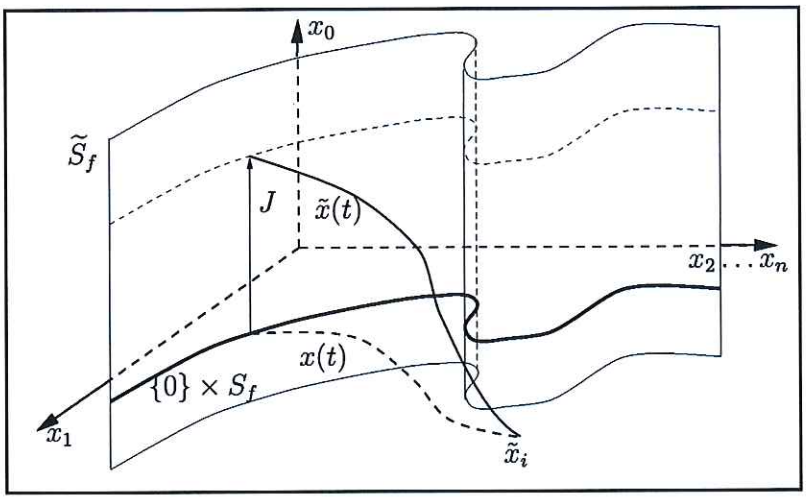

Figure 6.7: The state trajectory $x(t)$ is the dashed curve and the corresponding extended state trajectory $\tilde{x}(t)$ is the solid curve. The goal of the optimization is to find an admissible control such that the corresponding extended state trajectory intersects the manifold (in this case it is a surface) $\widetilde{S}_f = \mathbf{R} \times S_f$ at a point with as low x_0 coordinate as possible. The minimum x_0 coordinate corresponds to the optimal cost J^*. The manifold S_f on the $x_1, \dots x_n$-plane is drawn in thick solid line.

an extended terminal state manifold

$$\widetilde{S}_f = \mathbf{R} \times S_f = \{\tilde{x} \in \mathbf{R}^{n+1} : \tilde{x} = \begin{bmatrix} x_0 & x^T \end{bmatrix}^T ; G(x) = 0\}$$

Then the optimal control problem can equivalently be stated as follows: *Find an admissible control such that the extended state vector is transfered from $\tilde{x}_i = (0, x_i)$ to a point on the extended manifold $\widetilde{S}_f$ with as low x_0 coordinate as possible,* see Figure 6.7.

Recall the idea behind the proof of Theorem 7 in the previous section. We perturbed the optimal control and the optimal transition time t_f^*, which gave rise to a perturbation cone $\mathcal{K}(t_f^*)$ at the optimal end point $\tilde{x}^*(t_f^*)$. We could then argue that there must be a hyperplane that separates the perturbation cone $\mathcal{K}(t_f^*)$ from terminal points with lower values of the x_0-coordinate than the optimal. The final value $\tilde{\lambda}(t_f^*)$ of the extended adjoint vector was the normal to the hyperplane[2]. Transportation of the hyperplane to points corresponding to times in the interval $[0, t_f^*]$ was then used to derive the conditions on the Hamiltonian function. The situation is similar now but the hyperplane must be tangential to the "curve" $\{x_0^*(t_f^*)\} \times S_f$ at the optimal point, see Figure 6.8. An arbitrary tangent vector of

[2]The hyperplane is not necessarily unique.

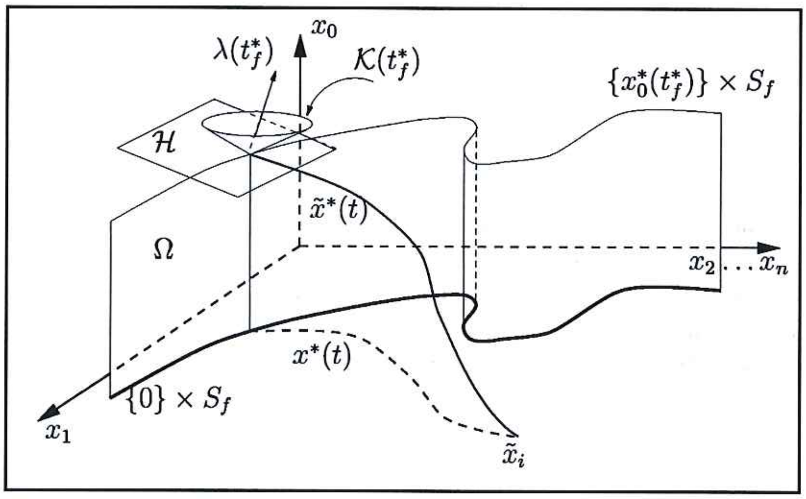

Figure 6.8: Define the surface $\Omega = \{x_0 : x_0 \leq x_0^*(t_f^*)\} \times S_f$ of points on the extended terminal manifold with lower x_0 coordinate than the optimal point $x^*(t_f^*)$. In order for $\tilde{x}^*(t_f^*)$ to be the optimal terminal point, it is necessary that the perturbation cone $\mathcal{K}(t_f^*)$ does intersect the interior of Ω (locally around the optimal point). It can be proven that this implies the existence of a separating hyperplane that must be tangential to the boundary of Ω, i.e., tangential to the "curve" $\{x_0^*(t_f^*)\} \times S_f$. For a careful proof, see [23].

$\{x_0^*(t_f^*)\} \times S_f$ at $\tilde{x}^*(t_f^*)$ has the form $\tilde{v} = \begin{bmatrix} 0 & v^T \end{bmatrix}^T$, where $v \in \mathbf{R}^n$ is perpendicular to all gradients $\nabla g_k(x^*(t_f^*))$, i.e.,

$$\begin{bmatrix} \frac{\partial g_1(x^*(t_f^*))}{\partial x_1} & \cdots & \frac{\partial g_1(x^*(t_f^*))}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_p(x^*(t_f^*))}{\partial x_1} & \cdots & \frac{\partial g_p(x^*(t_f^*))}{\partial x_n} \end{bmatrix} v = 0$$

Hence, for a hyperplane to be tangential to the "curve" $\{x_0^*(t_f^*)\} \times S_f$ we need its normal $\tilde{\lambda}(t_f^*)$ to be perpendicular to $\tilde{v}$, i.e., $\tilde{v}^T \tilde{\lambda}(t_f^*) = 0$. This condition implies that the following *transversality condition* must hold

$$\lambda(t_f^*)^T v = 0 \quad \text{for all } v \text{ such that} \quad \begin{bmatrix} \frac{\partial g_1(x^*(t_f^*))}{\partial x_1} & \cdots & \frac{\partial g_1(x^*(t_f^*))}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_p(x^*(t_f^*))}{\partial x_1} & \cdots & \frac{\partial g_p(x^*(t_f^*))}{\partial x_n} \end{bmatrix} v = 0 \qquad (6.8)$$

This transversality condition will be denoted $\lambda(t_f^*) \perp S_f$. It can equivalently be stated as $\lambda(t_f^*) = G_x(x^*(t_f^*))^T \nu = \sum_{k=1}^p \nu_k \nabla g_k(x^*(t_f^*))$, for some appropriate vector $\nu \in \mathbf{R}^p$. We can now summarize the optimality conditions for (6.7)

Theorem 8. *Suppose $(x^*(\cdot), u^*(\cdot), t_f^*)$ is an optimal solution of (6.7), i.e. it transfers x_i to S_f with minimum cost at the optimal transition time t_f^*. Then there exists a nonzero extended adjoint function such that*

(*i*) $\dot{\tilde{\lambda}}(t) = -H_{\tilde{x}}(x^*(t), u^*(t), \tilde{\lambda}(t))$

(*ii*) $H(x^*(t), u^*(t), \tilde{\lambda}(t)) = \min_{v \in U} H(x^*(t), v, \tilde{\lambda}(t)) = 0$ *for all* $t \in [0, t_f^*]$

(*iii*) $\lambda_0(t) = const \geq 0$

(*iv*) $\lambda(t_f^*) \perp S_f$

Proof. We have already discussed the main idea behind the proof above. For the complete details, see [23]. □

We illustrate with an example. This example also shows that the "time"-variable can mean other things than time.

Example 20. Determine the shape of the curve of minimum length that connects a point x_i in the plane with a smooth curve (manifold) $S_f = \{x \in \mathbf{R}^2 : g(x) = 0\}$, see Figure 6.9. Let s denote the arc length. The optimal control problem can then be stated

$$\text{minimize}\, s_f \quad \text{subj. to} \quad \begin{cases} \frac{dx_1}{ds} = \cos(\theta(s)) \\ \frac{dx_2}{ds} = \sin(\theta(s)) \\ x(0) = x_i, \; x(s_f) \in S_f \\ \theta \in [0, 2\pi], \; s_f \geq 0 \end{cases}$$

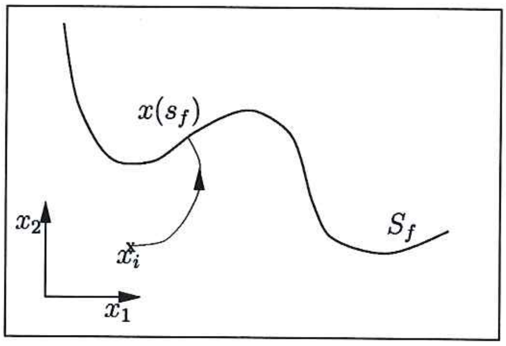

Figure 6.9: We want to determine the minimum length curve that connects the point x_i with the curve S_f.

In other words, we want to minimize the length s_f of a curve that connects the point x_i with the curve S_f. The dynamical equations just determine the direction of the curve at each $s \in [0, s_f]$.

The Hamiltonian is $H(x, \theta, \tilde{\lambda}) = \lambda_0 + \lambda_1 \cos(\theta) + \lambda_2 \sin(\theta)$. This gives the adjoint equation

$$\frac{d\lambda_1}{ds} = -H_{x_1}(x, \theta, \tilde{\lambda}) = 0 \quad \Rightarrow \quad \lambda_1(s) = \lambda_1^0$$
$$\frac{d\lambda_2}{ds} = -H_{x_2}(x, \theta, \tilde{\lambda}) = 0 \quad \Rightarrow \quad \lambda_2(s) = \lambda_2^0$$

i.e., all the adjoint variables are constant. The optimality condition (ii) becomes

$$\min_{\theta} H(x(s), \theta(s), \tilde{\lambda}(s)) = \lambda_0 - \sqrt{(\lambda_1^0)^2 + (\lambda_2^0)^2}$$

and the corresponding optimal direction is constant and such that

$$\begin{bmatrix} \cos\theta(s) \\ \sin\theta(s) \end{bmatrix} = - \begin{bmatrix} \lambda_1^0 \\ \lambda_2^0 \end{bmatrix} \frac{1}{\sqrt{(\lambda_1^0)^2 + (\lambda_2^0)^2}}$$

i.e., the vector field is aligned with the adjoint variable. This in turn implies that

$$\frac{dx_2}{dx_1} = \frac{\sin(\theta)}{\cos(\theta)} = \frac{\lambda_2^0}{\lambda_1^0} \tag{6.9}$$

which shows that the the slope of the the optimal joining curve is constant. Hence, we are looking for a straight line of shortest length between the initial point x_i and the curve S_f.

The transversality condition will tell us even more about this straight line. Indeed, it becomes

$$\lambda_1^0 v_1 + \lambda_2^0 v_2 = 0 \quad \text{for all } v \text{ such that} \quad g_{x_1}(x(s_f^*))v_1 + g_{x_2}(x(s_f^*))v_2 = 0$$

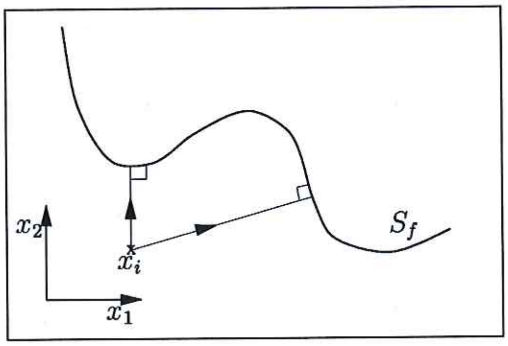

Figure 6.10: There are two candidates for optimality. PMP does not distinguish between them.

which is equivalent to

$$\lambda_1^0 g_{x_2}(x^*(s_f^*)) - \lambda_2^0 g_{x_1}(x^*(s_f^*)) = 0,$$

From (6.9) it now follows that the optimal straight line from x_i to S_f must be orthogonal to S_f:

$$\frac{x_2^*(s_f^*) - x_{2i}}{x_1^*(s_f^*) - x_{1i}} = \frac{g_{x_2}(x^*(s_f^*))}{g_{x_1}(x^*(s_f^*))}$$

We have thus proven that the curve of minimum length that joins x_i with S_f must be a straight line orthogonal to S_f. There are in general several candidate optimal solutions as illustrated in Figure 6.10.

Relationship with Dynamic Programming (Optional)

Let us introduce the value function[3] (optimal cost-to-go function)

$$J^*(x) = \int_t^{t_f^*} f_0(x^*(s), u^*(s))ds$$

where $(x^*(\cdot), u^*(\cdot), t_f^*)$ is an optimal solution to (6.7) for the case when we start at position (t, x), i.e., $x^*(t) = x$. It satisfies the boundary condition $J^*(x) = 0$ for $x \in S_f$.

An optimal isocost surface consist of initial states (in $\mathbf{R}^n$) with the same optimal cost. The extended state trajectories starting from an optimal isocost surface and terminating at the extended terminal manifold $\widetilde{S}_f = S_f \times \mathbf{R}$ makes

[3] J^* is time independent since $t_f - t$ is a variable which is optimized. The optimization of $t_f - t$ is independent of the particular starting time t since f_0 and f are independent of time (autonomous).

up another surface, the so called level surface[4]. Figure 6.11 illustrates an optimal isocost surface and the corresponding level surface. All extended states on the level surface end up at a terminal state with the same x_0 coordinate (cost). Let us define a function Ψ on the level surface as $\Psi(\tilde{x}) = x_0 + J^*(x)$. This is a constant function, since $\Psi(\tilde{x}) = x_0 + J^*(x) = J^*(x_i)$ for every point $\tilde{x} = (x_0, x)$ on the surface, see Figure 6.11. Note that this relation is nothing but the dynamic programming equation

$$J^*(x_i) = \min_{u(\cdot)} \left\{ \int_0^t f_0(x(s), u(s))ds + J^*(x(t)) \right\}.$$

since $x_0(t) = \int_0^t f_0(x(s), u(s))$.

Assume that J^* is continuously differentiable and consider an optimal solution $(\tilde{x}^*(\cdot), u^*(\cdot))$ in the extended state space. Then since, $\Psi(\cdot)$ is a constant function on the level surface we get

$$\begin{aligned}\dot{\Psi}(\tilde{x}^*(t)) &= \nabla\Psi(\tilde{x}^*(t))^T \tilde{f}(x^*(t), u^*(t)) \\ &= f_0(x^*(t), u^*(t)) + \frac{\partial J^*}{\partial x}(x^*(t))^T f(x^*(t), u^*(t)) = 0\end{aligned}$$

On the other hand, from the optimality of the states on the level surface it also follows that no admissible control will move the extended state below this surface. Hence, we obtain

$$\min_{u \in U} \left\{ f_0(x^*(t), u) + \frac{\partial J^*}{\partial x}(x^*(t))^T f(x^*(t), u) \right\} = 0 \tag{6.10}$$

If we compare with (*ii*) in Theorem 8 then we see that *the adjoint variables correspond to the gradient of the value function* $\lambda(t) = \nabla J^*(x^*(t))$*:* and[5] $\lambda_0 = 1$. Note that with $x^*(t) = x$ (6.10) is the HJBE.

In order to clearly see the similarities and differences between PMP and HJBE we illustrate how these methods are used in an application of the type in this section. For simplicity assume that we have a nonpathological case when $\lambda_0 > 0$, i.e., we can take $\lambda_0 = 1$. We take the following steps

Step 1 Define the Hamiltonian $H(x, u, \lambda) = f_0(x, u) + \lambda^T f(x, u)$

Step 2 Let $\widetilde{\mu}(x, \lambda) = \text{argmin}_{u \in U} H(x, u, \lambda)$

Step 3a In dynamic programming we now need to solve the HJBE

$$\begin{aligned}&H(x, \widetilde{\mu}(x, \nabla V(x)), \nabla V(x)) = 0 \\ &V(x) = 0 \quad \text{when} \quad x \in S_f\end{aligned}$$

[4]A derivation of PMP using isocost surfaces and level surfaces is done in [13].

[5]The reason for $\lambda_0 = 1$ is that, in Figure 6.11, the level surface is not vertical when intersecting the extended manifold $\widetilde{S}_f$. For such cases we can surely put $\lambda_0 = 1$.

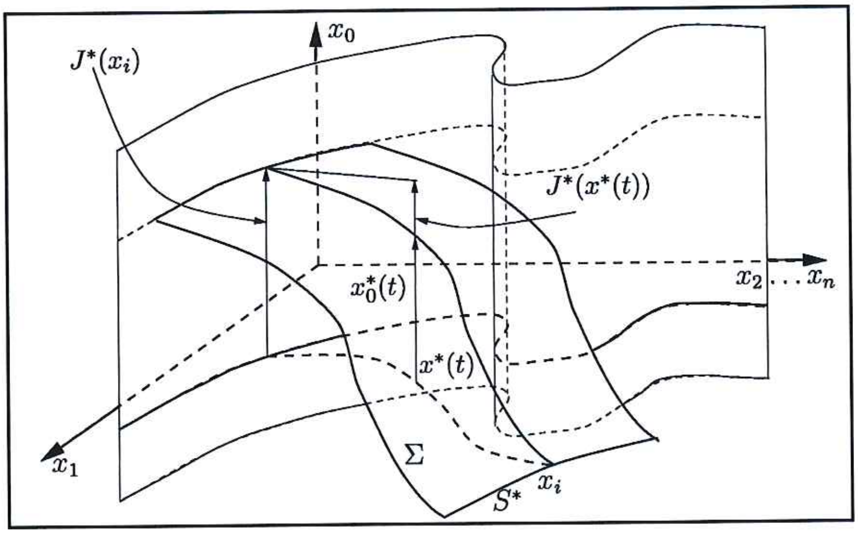

Figure 6.11: The "curve" S^* is the optimal isocost surface. It contains states in the $x_1 \ldots x_n$ plane with the same optimal cost as x_i. The level surface, Σ, consists of states in the extended state space with the same terminal x_0 coordinate (cost) as the points on the optimal isocost surface S^*. In particular, its intersection with the $x_1 \ldots x_n$ plane is the optimal isocost surface. The figure indicates the relation $J^*(x_i) = x_0^*(t) + J^*(x^*(t))$, $t \in [0, t_f^*]$, for an optimal trajectory on the level surface.

- This corresponds to computing the entire isocost surface.
- The optimal feedback control is $u(x) = \widetilde{\mu}(x, \nabla V(x))$.

Step 3b In PMP we solve the two point boundary value problem

$$\dot{x}(t) = H_\lambda(x(t), \mu(x(t), \lambda(t)), \lambda(t)), \quad x(0) = x_0, \quad x(t_f) \in S_f$$
$$\dot{\lambda}(t) = -H_x(x(t), \widetilde{\mu}(x(t), \lambda(t)), \lambda(t)), \quad \lambda(t_f) \perp S_f$$

- This corresponds to computing a trajectory on the isocost surface.
- The optimal control function is $u(t) = \widetilde{\mu}(x(t), \lambda(t))$

6.3 Some Generalizations

We will now discuss generalizations of the Pontryagin principle to cases where:

1. The initial point belongs to a smooth manifold.

2. There is a terminal cost term.

3. The system is nonautonomous.

We will only indicate how the derivation can be done by using the already established results. We refer to [23, 13] for further details.

Autonomous Systems

We consider the optimization problem

$$\text{minimize}\, \phi(x(t_f)) + \int_0^{t_f} f_0(x(t),u(t))dt \quad \text{subj. to} \quad \begin{cases} \dot{x}(t) = f(x(t),u(t)) \\ x(0) = S_i,\ x(t_f) \in S_f \\ u(t) \in U,\ t_f \geq 0 \end{cases} \tag{6.11}$$

where ϕ is assumed to be continuously differentiable, S_i is a smooth manifold, and everything else is as before. In order to derive necessary optimality conditions we transform the problem so that the previous results can be applied. Introduce additional control and state variables, u_{m+1} and x_{n+1}, respectively. Then the following optimization problem is equivalent to (6.11)

$$\text{minimize} \int_0^{t_f} [f_0(x(t),u(t)) + u_{m+1}(t)]dt$$

$$\text{subj. to} \quad \begin{cases} \dot{x}(t) = f(x(t),u(t)) \\ \dot{x}_{n+1}(t) = u_{m+1} \\ x(0) \in S_i,\ x(t_f) \in S_f, \\ x_{n+1}(0) = 0,\ x_{n+1}(t_f) = \phi(x(t_f)) \\ u(t) \in U,\ u_{m+1} \in \mathbf{R},\ t_f \geq 0 \end{cases}$$

This follows since $\int_0^{t_f} u_{m+1}(t)dt = x_{n+1}(t_f) = \phi(x(t_f))$. The initial condition gives rise to an initial transversality condition on the adjoint variable, see [23, 13]. In order to derive the remaining optimality conditions we apply Theorem 8. The Hamiltonian becomes

$$\begin{aligned} \tilde{H}\left(\begin{bmatrix} x \\ x_{n+1} \end{bmatrix}, \begin{bmatrix} u \\ u_{m+1} \end{bmatrix}, \begin{bmatrix} \tilde{\lambda} \\ \lambda_{n+1} \end{bmatrix}\right) &= \lambda_0(f_0(x,u) + u_{m+1}) + \lambda^T f(x,u) + \lambda_{n+1}u_{m+1} \\ &= H(x,u,\tilde{\lambda}) + (\lambda_0 + \lambda_{n+1})u_{m+1} \end{aligned}$$

where $H(x,u,\tilde{\lambda}) = \lambda_0 f_0(x,u) + \lambda^T f(x,u)$. Pointwise minimization gives rise to the condition

$$\begin{aligned} \min_{u\in U, u_{m+1}\in \mathbf{R}} \tilde{H}\left(\begin{bmatrix} x^* \\ x^*_{n+1} \end{bmatrix}, \begin{bmatrix} u \\ u_{m+1} \end{bmatrix}, \begin{bmatrix} \tilde{\lambda} \\ \lambda_{n+1} \end{bmatrix}\right) &= \begin{cases} \min_{u\in U} H(x^*,u,\tilde{\lambda}), & \lambda_{n+1} = -\lambda_0 \\ -\infty, & \lambda_{n+1} \neq -\lambda_0 \end{cases} \\ &= 0 \end{aligned}$$

Hence, we conclude that $\lambda_{n+1} = -\lambda_0$ and

$$\min_{u \in U} H(x^*, u, \tilde{\lambda}) = 0.$$

The adjoint equation becomes

$$\dot{\lambda}(t) = -\tilde{H}_x\left(\begin{bmatrix} x^*(t) \\ x^*_{n+1}(t) \end{bmatrix}, \begin{bmatrix} u^*(t) \\ u^*_{m+1}(t) \end{bmatrix}, \begin{bmatrix} \tilde{\lambda}(t)) \\ \lambda_{n+1}(t) \end{bmatrix}\right) = -H_x(x^*(t), u^*(t), \tilde{\lambda}(t))$$

$$\dot{\lambda}_{n+1}(t) = -\tilde{H}_{x_{n+1}}\left(\begin{bmatrix} x^*(t) \\ x^*_{n+1}(t) \end{bmatrix}, \begin{bmatrix} u^*(t) \\ u^*_{m+1}(t) \end{bmatrix}, \begin{bmatrix} \tilde{\lambda}(t)) \\ \lambda_{n+1}(t) \end{bmatrix}\right) = 0$$

The last equation is consistent with the fact that $\lambda_{n+1} = -\lambda_0 = const$. It remains to derive the boundary value of the adjoint function. Condition (*iv*) in Theorem 8 becomes

$$\begin{bmatrix} \lambda(t_f^*) \\ \lambda_{n+1} \end{bmatrix} \perp \tilde{S}_f = \left\{\begin{bmatrix} x \\ x_{n+1} \end{bmatrix} : \tilde{G}\left(\begin{bmatrix} x \\ x_{n+1} \end{bmatrix}\right) = 0\right\},$$

where

$$\tilde{G}\left(\begin{bmatrix} x \\ x_{n+1} \end{bmatrix}\right) = \begin{bmatrix} G(x) \\ x_{n+1} - \phi(x) \end{bmatrix}.$$

The condition can equivalently be written

$$\begin{bmatrix} \lambda(t_f^*) \\ \lambda_{n+1} \end{bmatrix}^T \begin{bmatrix} v \\ v_{n+1} \end{bmatrix} = 0, \quad \forall \begin{bmatrix} v \\ v_{n+1} \end{bmatrix} \quad \text{s.t.} \begin{bmatrix} G_x(x^*(t_f^*)) & 0 \\ -\nabla\phi(x^*(t_f^*))^T & 1 \end{bmatrix} \begin{bmatrix} v \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \lambda(t_f^*)^T v - \lambda_0 \nabla\phi(x^*(t_f))^T v = 0, \quad G_x(x^*(t_f^*))v = 0$$

$$\Leftrightarrow \lambda(t_f^*) - \lambda_0 \nabla\phi(x^*(t_f^*)) \perp S_f$$

where we used that $\lambda_{n+1} = -\lambda_0$. An equivalent formulation of this transversality condition is

$$\begin{aligned} \lambda(t_f^*) &= \lambda_0 \nabla\phi(x^*(t_f^*)) + G_x(x(t_f^*))^T \nu \\ &= \lambda_0 \nabla\phi(x^*(t_f^*)) + \sum_{k=1}^{p} \nu_k \nabla g_k(x^*(t_f^*)) \end{aligned}$$

for some vector $\nu \in \mathbf{R}^p$.

To summarize we have the following optimality conditions for problem (6.11): Note: We IGNORE the pathological case and USE $\lambda_0 = 1$.

PMP: Autonomous Systems: Define the Hamiltonian

$$H(x, u, \lambda) = f_0(x, u) + \lambda^T f(x, u)$$

Assume that $(x^*(t), u^*(t), t_f^*)$ is an optimal solution to (6.11). Then there exists an adjoint function $\lambda(\cdot)$ that satisfies the following conditions (*i*) $\dot{\lambda}(t) = -H_x(x^*(t), u^*(t), \lambda(t))$

(*ii*) $H(x^*(t), u^*(t), \lambda(t)) = \min_{v \in U} H(x^*(t), v, \lambda(t)) = 0$ for all $t \in [0, t_f^*]$

(*iii*) $\lambda(0) \perp S_i$

(*iv*) $\lambda(t_f^*) - \nabla\phi(x^*(t_f^*)) \perp S_f$

Special Case 1: It is reasonable to assume that the terminal cost and the terminal manifold involve two disjoint set of states. For example, $\phi(x) = \phi(x_{p+1}, \ldots, x_n)$ and $g_k(x) = g_k(x_1, \ldots, x_p)$, $k = 1, \ldots, p$. Then the transversality condition reduces to

$$\begin{bmatrix} \lambda_{p+1}(t_f^*) \\ \vdots \\ \lambda_n(t_f^*) \end{bmatrix} = \begin{bmatrix} \frac{\partial\phi(x(t_f^*))}{\partial x_{p+1}} \\ \vdots \\ \frac{\partial\phi(x(t_f^*))}{\partial x_n} \end{bmatrix} \tag{6.12}$$

and the remaining variables $(\lambda_1(t_f^*), \ldots, \lambda_p(t_f^*))$ remain undetermined.
Special Case 2: If $S_i = \{x_i\}$ (a given point) then there is no constraint on $\lambda(0)$.
Special Case 3: If $S_f = \mathbf{R}^n$ then $\lambda(t_f) = \nabla\phi(x^*(t_f^*))$.
Special Case 4: If $S_f = \mathbf{R}^n$ and $\phi = 0$ then $\lambda(t_f^*) = 0$.
Special Case 5: If $S_f = \{x_f\}$ (a given point) and $\phi = 0$ then there is no constraint on $\lambda(t_f)$.
Special Case 6: If the final time is fixed then (*ii*) is replaced by $H(x^*(t), u^*(t), \lambda(t)) = \min_{v \in U} H(x^*(t), v, \lambda(t)) = const$ for all $t \in [0, t_f]$.

Nonautonomous systems

We consider the optimization problem

$$\begin{aligned} &\text{minimize}\, \phi(t_f, x(t_f)) + \int_{t_i}^{t_f} f_0(t, x(t), u(t))dt \\ &\text{subj. to} \begin{cases} \dot{x}(t) = f(t, x(t), u(t)) \\ x(t_i) = x_i, \;\; x(t_f) \in S_f(t_f) \\ u(t) \in U, \;\; t_f \geq t_i \end{cases} \end{aligned} \tag{6.13}$$

where $\phi(t, x)$ is assumed to be continuously differentiable with respect to both arguments, $f_0(t, x, u)$ and $f(t, x, u)$ are continuously differentiable with respect to t and x, and the terminal manifold may depend on time:

$$S_f(t) = \{x \in \mathbf{R}^n : G(t, x) = 0\} \quad \text{where} \quad G(t, x) = \begin{bmatrix} g_1(t, x) \\ \vdots \\ g_p(t, x) \end{bmatrix}$$

It is assumed that the functional matrix

$$\begin{bmatrix} \frac{\partial g_1(x)}{\partial x_1} & \cdots & \frac{\partial g_1(x)}{\partial x_n} & \frac{\partial g_1(x)}{\partial t} \\ \vdots & & \vdots & \\ \frac{\partial g_p(x)}{\partial x_1} & \cdots & \frac{\partial g_p(x)}{\partial x_n} & \frac{\partial g_p(x)}{\partial t} \end{bmatrix}$$

has full rank.

In order to derive necessary optimality conditions we transform the problem so that the previous results can be applied. The trick is now to introduce an additional state $x_{n+1}(t)$ that corresponds to the time variable. The following optimization problem is equivalent to (6.13)

$$\text{minimize}\, \phi(x_{n+1}(t_f), x(t_f)) + \int_{t_i}^{t_f} f_0(x_{n+1}(t), x(t), u(t))dt$$

$$\text{subj. to} \begin{cases} \dot{x}(t) = f(x_{n+1}(t), x(t), u(t)) \\ \dot{x}_{n+1}(t) = 1 \\ x(t_i) = x_i, \;\; G(x_{n+1}(t_f), x(t_f)) = 0 \\ x_{n+1}(t_i) = t_i, \;\; x_{n+1}(t_f) = t_f \\ u(t) \in U, \;\; t_f \geq 0 \end{cases}$$

If we use the result in the previous subsection then we get a result similar to the previous one except that the constancy property of the Hamiltonian now is replaced by a more complex condition. See, for example, [23, 13] for more details about the derivation.

We get the following optimality conditions for problem (6.13):
Note: We IGNORE the pathological case and USE $\lambda_0 = 1$.

PMP: Nonautonomous Systems: Define the Hamiltonian function

$$H(t, x, u, \lambda) = f_0(t, x, u) + \lambda^T f(t, x, u)$$

Assume that $(x^*(t), u^*(t), t_f^*)$ is an optimal solution to (6.13). Then there exists an adjoint function $\lambda(\cdot)$ that satisfies the following conditions

(*i*) $\dot{\lambda}(t) = -H_x(t, x^*(t), u^*(t), \lambda(t))$

(*ii*) $H^*(t) = \min_{v \in U} H(t, x^*(t), v, \lambda(t))$ satisfies

$$\begin{aligned} H^*(t) &= H^*(t_f^*) - \int_t^{t_f^*} \frac{\partial H}{\partial s}(s, x^*(s), u^*(s), \lambda(s))ds, \quad t \in [t_i, t_f^*] \\ H^*(t_f^*) &= -\sum_{k=1}^{p} \nu_k \frac{\partial g_k}{\partial t}(t_f^*, x^*(t_f^*)) - \frac{\partial \phi}{\partial t}(t_f^*, x^*(t_f^*)) \end{aligned} \tag{6.14}$$

(*iii*) $(\lambda(t_f^*) - \phi_x(t_f^*, x^*(t_f^*))) \perp S_f(t_f^*)$, which means that there must exist a vector $\nu = \begin{bmatrix}\nu_1 & \dots & \nu_p\end{bmatrix}^T$ such that

$$\lambda(t_f^*) = \sum_{k=1}^{p} \nu_k \frac{\partial g_k}{\partial x}(t_f^*, x^*(t_f^*)) + \frac{\partial \phi}{\partial x}(t_f^*, x^*(t_f^*))$$

Special Case: If the terminal time is fixed then we can remove the time dependence of ϕ and S_f, i.e., the g_k are now only functions of the state. Conditions (*ii*) and (*iii*) are then replaced by

(*ii*) $H^*(t) = \min_{v \in U} H(t, x^*(t), v, \lambda(t))$ satisfies

$$H^*(t) = H^*(t_f) - \int_t^{t_f} \frac{\partial H}{\partial t}(s, x^*(s), u^*(s), \lambda(s))ds, \quad t \in [t_i, t_f]$$

(*iii*) $\lambda(t_f) - \nabla\phi(x^*(t_f)) \perp S_f$ or equivalently

$$\lambda(t_f) = \sum_{k=1}^{p} \nu_k \frac{\partial g_k}{\partial x}(x^*(t_f)) + \frac{\partial \phi}{\partial x}(x^*(t_f))$$

for some suitable vector $\nu = \begin{bmatrix}\nu_1 & \dots & \nu_p\end{bmatrix}^T$.

6.4 How to Use PMP

A professional way to address optimal control problems is to start investigating the vector field and the cost function to determine if

- it is possible to conclude that there must exist an optimal solution,
- the optimal solution is unique.

The existence and uniqueness questions are addressed in more advanced books, like [12, 23].

The next step (in our case it would be the first) is to use PMP. We take the following steps (we consider problem (6.13) and assume $\lambda_0 = 1$)

1. Define the Hamiltonian: $H(t, x, u, \lambda) = f_0(t, x, u) + \lambda^T f(t, x, u)$
2. Perform pointwise minimization: $\tilde{\mu}(t, x, \lambda) = \text{argmin}_{u \in U} H(t, x, u, \lambda)$, which means that a candidate optimal control is $u^*(t) = \mu(t, x(t), \lambda(t))$.

3. Solve the Two Point Boundary Value Problem (TPBVP)

$$\dot{\lambda}(t) = -H_x(t, x(t), \widetilde{\mu}(t, x(t), \lambda(t)), \lambda(t)), \quad \lambda(t_f) - \frac{\partial \phi}{\partial x}(t_f, x(t_f)) \perp S_f(t_f)$$
$$\dot{x}(t) = H_\lambda(t, x(t), \widetilde{\mu}(t, x(t), \lambda(t)), \lambda(t)), \quad x(t_i) = x_i, \quad x(t_f) \in S_f(t_f)$$

   One of the difficulties when solving a TPBVP is to find appropriate boundary conditions for x and λ. In order to obtain conditions that help us find candidates for the optimal transition time we also use (6.14) above.

4. Compare the candidate solutions obtained using PMP.

The TPBVP must often be solved numerically using, for example, the shooting method explained in Chapter 10 or the book [4]

Optimal control SF2852

Chapter 7

Examples

We consider some applications of optimal control.

7.1 The Optimal Storage Problem

In this example we try to find an optimal storage strategy.

$$J^* = \text{minimize} \int_0^{t_f} (u(t)e^{rt} + cx(t))dt \quad \text{subj. to} \quad \begin{cases} \dot{x}(t) = u(t), & x(0) = 0 \\ x(t_f) = A, & 0 \leq u(t) \leq M \end{cases}$$

where the final time is assumed fixed. The variables denote (all are non-negative quantities)

- x stock size
- u production rate
- r production cost growth rate
- c storage cost

The Hamiltonian is

$$H(t, x, u, \tilde{\lambda}) = \lambda_0 f_0(t, x, u) + \lambda_1 f(t, x, u) = \lambda_0 cx + u(\lambda_0 e^{rt} + \lambda_1)$$

The adjoint equation for λ_1 becomes

$$\dot{\lambda}_1 = -\frac{\partial H(t, x, u, \tilde{\lambda})}{\partial x} = -\lambda_0 c$$

which gives the solution $\lambda_1(t) = -\lambda_0 ct + \lambda_1(0)$. Optimality condition (ii) requires that

$$\begin{aligned} H(t, x(t), u(t), \tilde{\lambda}(t)) &= \min_{0 \leq u \leq M} H(x(t), u, \tilde{\lambda}(t)) \\ &= \lambda_0 cx(t) + \min_{0 \leq u \leq M} u \cdot \sigma(t) \end{aligned} \tag{7.1}$$

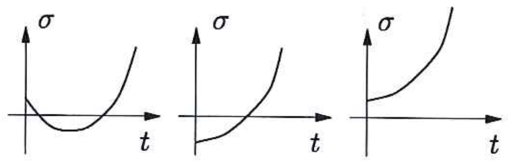

Figure 7.1: The figure shows the three possible shapes of the switching function $\sigma(t)$. We have three possible switching sequences for the control *a*) $\{0, M, 0\}$ or *b*) $\{M, 0\}$, or *c*) $\{0\}$, where the third is impossible by the endpoint condition.

where the switching function is $\sigma(t) = (\lambda_0 e^{rt} - \lambda_0 ct + \lambda_1(0))$. Consider first the case when $\lambda_0 = 0$. This means that $\lambda_1(t) = \lambda_1(0) \neq 0$ since otherwise $\tilde{\lambda}(t) = (\lambda_0(0), \lambda_1(0)) = 0$, which is not allowed. However, then (7.1) reduces to $H(x(t), u(t), \tilde{\lambda}(t)) = \min_{0 \leq u \leq M} \lambda_1(0)u$ and we must consider two cases

1. $\lambda_1(0) > 0$: Then $\min_{0 \leq u \leq M} \lambda_1(0)u = 0$ and $u(t) = 0$, which is impossible since then the end condition $x(t_f) = A$ will be violated.

2. $\lambda_1(0) < 0$: Then $\min_{0 \leq u \leq M} \lambda_1(0)u = \lambda_1(0)M$, and $u(t) = M$. This is only possible if $t_f = A/M$, because otherwise we have either $x(t_f) < A$ or $x(t_f) > A$.

We conclude that if $t_f = A/M$ then the unique control is $u(t) = M$ (it is the only control that brings us to the required end point $x(t_f) = A$). For all other t_f we must have $\lambda_0 > 0$ and then without loss of generality $\lambda_0 = 1$.

With $\lambda_0 = 1$ it is easy to see that the optimum in (7.1) is achieved by

$$u(t) = \begin{cases} M, & \sigma(t) < 0 \\ \text{arbitrary in } [0, t_f], & \sigma(t) = 0 \\ 0, & \sigma(t) > 0 \end{cases} \tag{7.2}$$

In our case $\sigma(t) = e^{rt} - ct + \lambda_1(0)$ becomes zero only at most at two distinctive points, where the value of u is irrelevant for the final cost. See Figure 7.1 for an illustration of the possible switching sequences for the control. The situation would be different if $\sigma(t)$ could have vanished on a nonzero time interval. This is called the singular case and then it requires more work to determine the value of the control.

We have thus seen that the optimal control must be on the form

$$u(t) = \begin{cases} 0, & 0 \leq t < t_1 \\ M, & t_1 \leq t < t_2 \\ 0, & t_2 \leq t \leq t_f \end{cases}$$

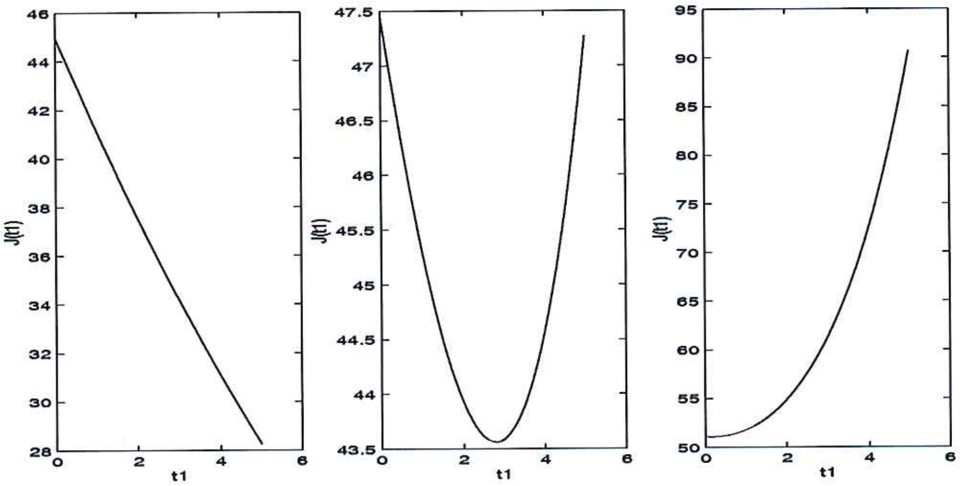

Figure 7.2: The cost $J(t_1)$ is plotted as a function of t_1 for the cases when $r = 0.15$, $r = 0.25$, and $r = 0.35$, respectively. We see that for low production cost growth it is optimal to produce as late as possible (left figure where $r = 0.15$) and for high production cost growth it is optimal to produce as early as possible (right figure where $r = 0.35$).

where t_1 and t_2 are variables to be determined. In order for $x(t_f) = A$, we need $t_2 - t_1 = A/M$. This shows that $t_f \geq A/M$ in order for the problem to be feasible.

The problem have now been reduced to the determination of a suitable $t_1 \in [0, t_f - A/M]$ such that the cost is minimized. For a given t_1 we have the cost

$$J(t_1) = \int_{t_1}^{t_1+A/M} [Me^{rt} + c(t - t_1)]dt + \int_{t_1+A/M}^{t_f} cAdt$$
$$= \frac{M}{r}(e^{r(t_1+A/M)} - e^{rt_1}) + \frac{cA^2}{2M^2} + cA(t_f - t_1 - A/M)$$

The optimal cost becomes

$$J^* = \min_{0 \leq t_1 \leq t_f - A/M} J(t_1)$$

which is a scalar optimization problem that can be addressed using basic methods from calculus. In Figure 7.2 we plot $J(t_1)$ when $t_f = 10$, $c = 1$, $M = 1$, and $A = 5$ for the case when $r = 0.15$, $r = 0.25$ and $r = 0.35$.

7.2 Dubins' Car

Consider the shortest path problem for Dubins' car, which was introduced in Example 2 in Chapter 1. We showed that the shortest path problem was equivalent to a minimum time problem. If we, without loss of generality, assume that the initial point is $(x(0), y(0), \theta(0)) = (0, 0, 0)$, then the minimum path problem corresponds to the optimal control problem

$$\text{minimize}\, T \quad \text{subj. to} \quad \begin{cases} \dot{x}(t) = v\cos(\theta(t)),\ x(0) = 0,\ x(T) = \bar{x} \\ \dot{y}(t) = v\sin(\theta(t)),\ y(0) = 0,\ y(T) = \bar{y} \\ \dot{\theta}(t) = \omega(t),\ \theta(0) = 0,\ \theta(T) = \bar{\theta} \\ |\omega(t)| \le v/R,\ T \ge 0 \end{cases} \tag{7.3}$$

where the terminal point is denoted $(\bar{x}, \bar{y}, \bar{\theta})$.

We use PMP and see how far this brings us. We first note that the cost function can be written $T = \int_0^T dt$, i.e., $f_0(x, y, \theta, \omega) = 1$. It is then easy to see that the Hamiltonian becomes

$$H((x, y, \theta), \omega, \tilde{\lambda}) = \lambda_0 + \lambda_1 v\cos(\theta) + \lambda_2 v\sin(\theta) + \lambda_3\omega$$

The adjoint equations become

$$\begin{aligned} &\dot{\lambda}_1(t) = 0 \\ &\dot{\lambda}_2(t) = 0 \\ &\dot{\lambda}_3(t) = \lambda_1(t)v\sin(\theta(t)) - \lambda_2(t)v\cos(\theta(t)) \end{aligned} \quad \Rightarrow \quad \begin{aligned} &\lambda_1(t) = \lambda_1^0 \\ &\lambda_2(t) = \lambda_2^0 \\ &\lambda_3(t) = \lambda_3(0) + \lambda_1^0 y(t) - \lambda_2^0 x(t) \end{aligned}$$

i.e., $\lambda_1(t)$ and $\lambda_2(t)$ are constant. The last equation (for $\lambda_3(t)$) follows by integrating the two equations in the constraint set of (7.3).

Our next step is to find a candidate controller by using *(ii)* in Theorem 7. We start with the pointwise minimization

$$\text{argmin}_{|\omega| \le v/R} H((x, y, \theta), \omega, \tilde{\lambda}) = \text{argmin}_{|\omega| \le v/R} \lambda_3\omega = \begin{cases} \frac{v}{R}, & \lambda_3 < 0 \\ ??, & \lambda_3 = 0 \\ -\frac{v}{R}, & \lambda_3 > 0 \end{cases}$$

We have a *Bang-Bang* solution unless $\lambda_3(t) = 0$ on a nonzero time interval. The case when $\lambda_3(t) = 0$ on a nonzero time interval is called *singular*. To understand what the control must be in the *singular case* we use *(ii)* in Theorem 7. We have

$$\min_{|\omega| \le v/R} H((x, y, \theta), \omega, \lambda) = \lambda_0 + \lambda_1^0 v\cos(\theta(t)) + \lambda_2^0 v\sin(\theta(t)) - \frac{v}{R}|\lambda_3(t)| = 0$$

for $t \in [0, T^*]$, where T^* denotes the minimum time. When $\lambda_3(t) = 0$ on a nonzero time interval, $I \subset [0, T^*]$, this becomes

$$\lambda_1^0 v\cos(\theta(t)) + \lambda_2^0 v\sin(\theta(t)) = -\lambda_0, \quad t \in I.$$

This can only hold if $\theta(t)$ is constant on I, which means that $\omega(t) = 0$ on I. We conclude that the optimal control must be

$$\omega(t) = \begin{cases} \frac{v}{R}, & \lambda_3(t) < 0 \\ 0, & \lambda_3(t) = 0 \text{ (on a nonzero time interval)} \\ -\frac{v}{R}, & \lambda_3(t) > 0 \end{cases}$$

and each of these cases corresponds to the following trajectory piece

1. $\omega(t) = \frac{v}{R}$ corresponds to a left turn with minimum turning radius R, i.e., we get a circular arc of radius R. We denote this trajectory segment l_L, where L is a number representing the time extent of the segment (or equivalently the length of the segment).

2. $\omega(t) = 0$ corresponds to a straight line segment. Indeed, $\omega(t) = 0$ when $\lambda_3(t) = 0$, so the expression for $\lambda_3(t)$ gives

$$\lambda_1^0 y(t) - \lambda_2^0 x(t) = -\lambda_3(0).$$

   This is a line with direction $(\lambda_1^0, \lambda_2^0)$. We denote this trajectory segment s_L, where again L is a number representing the time extent of the segment (or equivalently the length of the segment).

3. $\omega(t) = -\frac{v}{R}$ corresponds to a right turn with minimum turning radius R, i.e., we get a circular arc of radius R. We denote this trajectory segment r_L.

We have now learnt that the optimal path (if it exists) consists of a combination of at most three types of trajectory segments, "turn right" (r), "go straight" (s), and "turn left" (l). However, we still don't know how these pieces should be ordered, how many pieces there should be, and how "long" they should be. We would immediately get an answer to these questions if we knew the correct initial conditions $\lambda(0) = (\lambda_1^0, \lambda_2^0, \lambda_3(0))$. Numerical optimization can give candidate solutions by recursively updating $\lambda(0)$ until all the conditions of PMP hold. However, PMP generally gives several candidates as is illustrated in Figure 7.3. We can always compare the extremals provided in order to see which one is shortest. It would, however, be more efficient to have additional rules for what the optimal extremal must look like. Fortunately, Sussman and Tang has provided such rules in [27]. There they show that

1. There exists an optimal solution to (7.3).

2. The optimal solution must be either of the following types

   - $B_a S_b B_c$, where each B is either r or l, and $a, c \in [0, \frac{2\pi R}{v})$, $b > 0$.

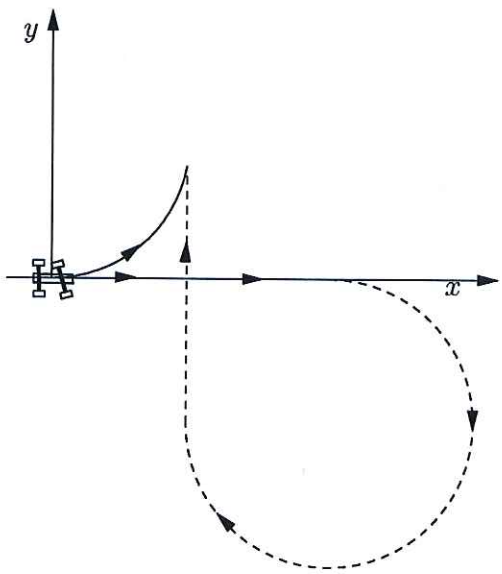

Figure 7.3: We want to find the shortest path for Dubins' car from the configuration $(0, 0, 0)$ to the configuration $(\bar{x}, \bar{y}, \bar{\theta}) = (R, R, \pi/2)$. In the figure we show two paths that both satisfy the conditions of PMP. The solid line $l_{\frac{\pi R}{2v}}$ is the shortest path. The dashed path $s_{\frac{2R}{v}} r_{\frac{3\pi R}{2v}} s_{\frac{2R}{v}}$ is just an extremal (stationary solution) for the optimization problem (7.3).

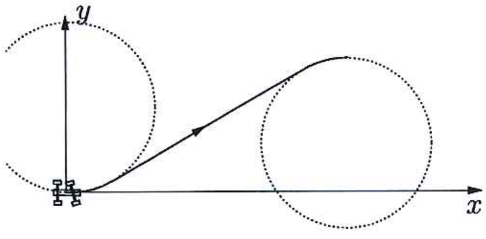

Figure 7.4: The shortest path for Dubins' car from the configuration $(0, 0, 0)$ to the configuration $(3.3R, 1.57R, 0)$. The optimal path is of the type lsr.

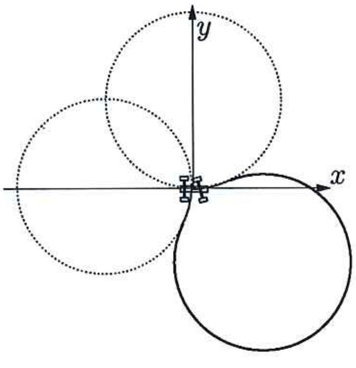

Figure 7.5: The shortest path for Dubins' car from the configuration $(0, 0, 0)$ to the configuration $(0, 0, \frac{\pi}{2})$. The path is of the type lrl.

- $B_aB_bB_c$, where each B is either r or l, i.e., this is either a $r_al_br_c$ sequence of trajectory pieces or a $l_ar_bl_c$ sequence of trajectory pieces. The time parameters are restricted by $b \in (\frac{\pi R}{v}, \frac{2\pi R}{v})$, $\min\{a, c\} < b - \frac{\pi R}{v}$, and $\max\{a, c\} < b$.

Note that $a, b, c \geq 0$, and that one or two of them can be zero. For example, s_b, r_a, and s_bl_c are admissible.

Two examples of minimal paths for Dubins' car are given in Figures 7.4 and 7.5.

Chapter 8

Infinite Horizon Optimal Control

It is common in applications that a system is designed to operate around a certain operating condition for long time periods. The transient behavior is then not the main design criterion but rather the ability of the system to maintain its position in the event of disturbances. We illustrate by an example

Example 21. Consider the inverted pendulum in Figure 8.1. The differential equation of this system is

$$ml\ddot{\theta} = mg\sin(\theta) + u$$

where θ is the angle between the rod and the vertical axis and u is the torque at the pivot point. Further, m is the mass of the bob and l denotes the length of the rod. We assume that $ml = 1$ and $mg = 1$. If we let $x_1 = \theta$ and $x_2 = \dot{\theta}$ then the state space representation of the inverted pendulum becomes

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ \sin(x_1) + u \end{bmatrix} =: f(x, u) \tag{8.1}$$

This system has an *equilibrium point* (*stationary point*) at $(x, u) = (0, 0)$. This means that if we bring the bob to rest at the upraised position $(\theta, \dot{\theta}) = (0, 0)$ then the pendulum stays in this upraised position. However, the slightest disturbance from this equilibrium position will make the pendulum fall down. We say that the equilibrium point is *unstable*.

From the basic course on automatic control we know that feedback control can help stabilize the pendulum in its upraised position. The design of such feedback laws are for simplicity often based on the linearized dynamics. Linearization of (8.1) around $x = 0$ gives

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u =: Ax + Bu \tag{8.2}$$

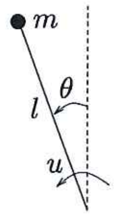

Figure 8.1: The inverted pendulum

which is a valid linear approximation of (8.1) in a neighborhood of $x = 0$. Note, in particular that it is an unstable system since the eigenvalues of A are ± 1.

If we apply the state feedback law $u = -Lx$, where $L = \begin{bmatrix} 2 & 2 \end{bmatrix}$ then the closed loop system becomes $\dot{x} = (A - BL)x$, where

$$A - BL = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}$$

This closed loop system matrix is stable since it has a double eigenvalue at -1. The practical implication is that this state feedback controller will stabilize the the unstable equilibrium $x = 0$ of the nonlinear system (8.1).

In this chapter we will learn that infinite horizon linear quadratic optimal control problems on the form

$$\text{minimize} \int_0^\infty (x^T Qx + ru^2)dt \quad \text{subject to} \quad \dot{x} = Ax + Bu, \; x(0) = x_0 \tag{8.3}$$

result in state feedback solutions that not only are stabilizing but also have good robustness properties.

The primary design criterion in control of a system over an infinite time horizon is stability. We will first formally state the stability definition used in this chapter before we discuss an optimal control problem that ensures stability as well as some additional performance.

Definition 3. Consider an autonomous system

$$\dot{x} = f(x). \tag{8.4}$$

A point x^* in the state space is called an equilibrium point of (8.4) if $f(x^*) = 0$. It has the property that if the system starts in x^* then it remains in x^* at all future time instances.

We will without any loss of generality only consider equilibrium points at the origin $x^* = 0$ (otherwise we make the change of variables $z = x - x^*$ and consider the system $\dot{z} = f(z + x^*)$).

Definition 4. Assume $f(0) = 0$. The equilibrium point $x = 0$ of $\dot{x} = f(x)$ is globally asymptotically stable if

(*i*) for any $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ such that

$$\|x(0)\| \leq \delta(\epsilon) \quad \Rightarrow \quad \|x(t)\| \leq \epsilon, \quad \forall t \geq 0$$

(*ii*) for any $x(0) \in \mathbf{R}^n$ the solution converges to zero, $\lim_{t\to\infty} x(t) = 0$.

Let us now consider the optimal control problem

$$\text{minimize} \int_0^\infty f_0(x,u)dt \quad \text{subject to} \quad \begin{cases} \dot{x} = f(x,u) \\ x(0) = x_0, \;\; u(t) \in U(x) \end{cases} \tag{8.5}$$

where f_0 and f are locally Lipschitz continuous function (this is the case if f_0 and f are continuously differentiable functions).

We assume without loss of generality that we want to control the system to an equilibrium point at $(x,u) = (0,0)$. This means that we assume $f(0,0) = 0$. In order to obtain a finite cost we further need to assume $f_0(0,0) = 0$. There is additional complication compared to the previous theory in that we now have to ensure that the optimal solution is asymptotically stable. To do this, we need to define positive definiteness of functions and state some additional assumptions on the vector field and the cost function in (8.5).

Definition 5. A function $V : \mathbf{R}^n \to \mathbf{R}$ is called *positive semi-definite* if $V(0) = 0$ and $V(x) \geq 0$ for all $x \in \mathbf{R}^n$. If it satisfies the stronger condition $V(x) > 0$ for all $x \neq 0$ then it is called *positive definite*. It is called radially unbounded if $V(x) \to \infty$ when $\|x\| \to \infty$.

Example 22. A quadratic form $V(x) = x^T Px$, where $P = P^T$, is positive definite (semi-definite) if $P > 0$ ($P \geq 0$), i.e., if all eigenvalues of P are positive (non-negative). It is radially unbounded if $P > 0$.

Assumption 3. We assume that f_0 is positive semi-definite and positive definite in u, i.e., $f_0(x,u) \geq 0$, $\forall (x,u) \in \mathbf{R}^{n \times m}$ and $f_0(x,u) > 0$ when $u \neq 0$.

Assumption 4. We will assume that the artificial output $h(x) = f_0(x,0)$ of the system $\dot{x} = f(x,0)$ is observable in the sense that $h(x(t)) = 0$ for all $t \geq 0$ implies that $x(t) = 0$ for all $t \geq 0$.

Before we consider an example that illustrates the last assumption we recall the definition of observability and controllability from the basic course.

Definition 6. Consider the linear system

$$\begin{aligned}\dot{x} &= Ax + Bu \\ y &= Cx\end{aligned} \tag{8.6}$$

The system is observable if (we sometimes say the pair (C, A) is observable)

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

has rank n, where n is the dimension of the state vector x. The system is controllable if ((B, A) is controllable)

$$\begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix}$$

has rank n. The system in (8.6) has a minimal state space realization if (C, A) is observable and (A, B) is controllable.

Example 23. Consider the linear quadratic control problem

$$\min \int_0^\infty (x^T Qx + u^T Ru)dt \quad \text{subject to} \quad \begin{cases} \dot{x} = Ax + Bu \\ x(0) = x_0 \end{cases}$$

i.e., $f_0(x, u) = x^T Qx + u^T Ru$ and $f(x, u) = Ax + Bu$. In order for Assumption 3 to hold we need $Q \geq 0$ and $R > 0$.

Let us assume we have the factorization $Q = C^T C$. Such a factorization is always possible to find when $Q = Q^T \geq 0$. We will now show that if (C, A) is observable then Assumption 4 holds.

First notice that the statement $f_0(x, 0) = 0$ for all solutions to $\dot{x} = f(x, 0)$ now means $Cx(t) = 0$ for all solutions to $\dot{x} = Ax$. Since $x(t) = e^{At}x_0$ this means that $y(t) = Ce^{At}x_0 = 0$ for all $t \geq 0$ and $x_0 \in \mathbf{R}^n$. Hence, $y(0) = \dot{y}(0) = \ldots = y^{n-1}(0) = 0$, which gives

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x_0 = 0, \quad \forall x_0 \in \mathbf{R}^n$$

Since (C, A) is observable this implies $x_0 = 0$, which in turn means that $x(t) = 0$ for all $t \geq 0$.

Let us now define the optimal cost-to-go function (value function) corresponding to (8.5)[1]

$$J^*(x_0) = \min_{u(\cdot)} \int_0^\infty f_0(x,u)dt.$$

The value function is independent of time since the dynamics and cost function of (8.5) both are independent of time.

Theorem 9. *Suppose Assumption 3 and Assumption 4 hold and*

(i) $V \in C^1$ *is positive definite, radially unbounded, and satisfies the (infinite horizon) HJBE*

$$\min_{u\in U}\left\{ f_0(x,u) + \frac{\partial V}{\partial x}(x)^T f(x,u)\right\} = 0 \tag{8.7}$$

(ii) $\mu(x) = \operatorname{argmin}_{u\in U}\left\{f_0(x,u) + \frac{\partial V}{\partial x}(x)^T f(x,u)\right\}$.

Then

(a) $V(x) = J^*(x)$

(b) $u = \mu(x)$ *is an optimal globally asymptotically stabilizing feedback control.*

Proof. Sketch: We first prove that $u(t) = \mu(x(t))$ is globally stabilizing [2]. Integration of HJBE gives

$$V(x_0) \le V(x(t)) + \int_0^t f_0(x(s),u(s))ds,$$

with equality if $u(t) = \mu(x(t))$. Since $V(x(t)) \ge 0$, we have

$$\int_0^t f_0(x(s),\mu(x(s)))ds = V(x_0) - V(x(t)) \le V(x_0). \tag{8.8}$$

Hence, since $f_0(x,\mu(x)) \ge 0$ it follows that $f_0(x(t),\mu(x(t))) \to 0$ as $t \to \infty$, because otherwise the integral would not be bounded, which violates (8.8). Moreover, since $f_0(x,u)$ is positive definite in u we get the stronger condition that $x(t) \to L := \{x \in \mathbf{R}^n : f_0(x,0) = 0\}$, i.e., $\mu(x(t)) \to 0$. However, this means that in the limit we have $h(x(t)) = f_0(x(t),0) = 0$, where $\dot{x}(t) = f(x(t),0)$. By the observability assumption it follows that $L = \{0\}$, i.e., $x(t) \to 0$ as $t \to \infty$.

[1]We assume there exists an optimal solution otherwise min should be replaced by inf. We make this assumption at several places in this chapter.

[2]In Remark 20 we discuss a more careful proof of stability. There we exploit that $V(x(t)) = \lim_{T\to\infty} \int_t^T f_0(x(s),\mu(x(s)))ds$ is well defined because the integral converges due to (8.8).

We have now proved that $u = \mu(x)$ is stabilizing in the sense that the closed loop state vector converges to zero. We will next see that it also gives the minimal cost. From (8.8) we have

$$\lim_{T\to\infty}\int_0^T f_0(x(t),u(t))dt \geq V(x_0) - \lim_{T\to\infty} V(x(T)) = V(x_0)$$

for any stabilizing control function $u(\cdot)$ (we used $x(T) \to 0 \Rightarrow V(x(T)) \to 0$). Since we have equality in (8.8) when $u = \mu(x)$, we get

$$V(x_0) = \int_0^\infty f_0(x(t),\mu(x(t)))dt \leq \int_0^\infty f_0(x(t),u(t))dt$$

which proves optimality. □

Remark 20. The above proof is not complete since we only proved global convergence of the state vector to zero, which is not as strong as global asymptotic stability. The reader who knows Lyapunov theory and in particular the LaSalle invariance principle can obtain the stability conclusions by exploiting the convergence we already proved. Indeed, we can use the value function to define a radially unbounded Lyapunov function that satisfies all the conditions of LaSalle's invariance principle [11]. We use that the system is time invariant, which means that if we start at time t at any state x then the closed loop solution still converges to zero. Consider

$$\dot{z} = f(z,\mu(z)), \quad z(t) = x.$$

and let

$$V(x) = \int_t^\infty f_0(z,\mu(z))ds$$

This Lyapunov function satisfies

(*a*) $V(x) > 0$ and $|V(x)| \to \infty$ when $\|x\| \to \infty$.

(*b*) $\dot{V}(x) = -f_0(x,\mu(x)) \leq 0$

(*c*) $S = \{x \in \mathbf{R}^n : \dot{V}(x) = 0\} = \{x \in \mathbf{R}^n : f_0(x,0) = 0\}$. Due to the observability assumption the only invariant subset of S is $\{0\}$.

Hence, by Corollary 3.2 in [11] it follows that the origin is globally asymptotically stable.

We next consider the special case of linear quadratic optimal control

Theorem 10. *Consider*

$$J^*(x_0) = \min \int_0^\infty [x^T Qx + u^T Ru]dt$$

$$\text{subject to} \quad \begin{cases} \dot{x} = Ax + Bu \\ x(0) = x_0 \end{cases}$$

where $Q = C^TC$ and $R > 0$. We assume that (C, A) is observable and that (A, B) is controllable. Then

(a) *$J^*(x_0) = x_0^T P x_0$, where P is the unique positive definite solution to the Algebraic Riccati Equation (ARE)*

$$A^T P + PA + Q = PBR^{-1}B^T P. \tag{8.9}$$

(b) *$\mu(x) = -R^{-1}B^T Px$ is the optimal, stabilizing, feedback control.*

Remark 21. Conclusion (*b*) in particular means that the closed loop system matrix $A - BR^{-1}B^T P$ has all eigenvalues in the open left half plane.

Proof. Let us try $V(x) = x^T Px$ in the HJBE (8.7). We get

$$\begin{aligned}\min_u \left\{x^T Qx + u^T Ru + 2x^T P(Ax + Bu)\right\} &= \left[\mu(x) = -R^{-1}B^T Px\right] \\ &= x^T Qx + x^T PBR^{-1}BPx + 2x^T P(Ax - BR^{-1}BP)x \\ &= x^T(A^T P + PA + Q - PBR^{-1}B^T P)x = 0\end{aligned}$$

where we used the ARE in the last equality. Hence, $V(x) = x^T Px$ is positive definite, radially unbounded, and satisfies the HJBE. Since the observability assumption (Assumption 4) holds it follows from Theorem 9 that $u = -R^{-1}B^T Px$ is the optimal stabilizing control.

The existence of a unique positive definite solution to the ARE can be proven as in [14] using controllability of (A, B) and observability of (C, A). □

We will next show that the linear quadratic regulator in Theorem 10 satisfies certain robustness properties. The following inequality derived from the ARE is of key importance

Proposition 7. *Let $L = R^{-1}B^T P$, where P is a solution to the ARE in (8.9). Then the transfer function*

$$G(s) = L(sI - A)^{-1}B$$

satisfies the inequality

$$(I + G(j\omega))^* R(I + G(j\omega)) \geq R \tag{8.10}$$

Proof. From the ARE we have

$$
\begin{aligned}
& A^TP + PA + Q - PBR^{-1}B^TP = 0 \\
\Leftrightarrow \quad & (-j\omega I - A^T)P + P(j\omega I - A) + PBR^{-1}B^TP = Q.
\end{aligned}
$$

If we multiply the last expression by $B^T(-j\omega I - A^T)^{-1}$ on the left and by $(j\omega I - A)^{-1}B$ on the right, then we get

$$
\begin{aligned}
& B^TP(j\omega I - A)^{-1}B + B^T(-j\omega I - A^T)^{-1}PB + \\
& B^T(-j\omega I - A^T)^{-1}PBR^{-1}B^TP(j\omega I - A)^{-1}B = B^T(-j\omega I - A^T)^{-1}Q(j\omega I - A)^{-1}B
\end{aligned}
$$

which is equivalent to

$$
RG(j\omega) + G(j\omega)^*R + G(j\omega)^*RG(j\omega) = B^T(-j\omega I - A^T)^{-1}Q(j\omega I - A)^{-1}B \geq 0
$$

since $B^TP = RL$. The right hand inequality follows since $Q \geq 0$. If we add the term R on both sides of the inequality then we get

$$
(I + G(j\omega))^*R(I + G(j\omega)) \geq R
$$

which is the desired inequality. □

We will next give an important interpretation of the inequality (8.10). Let us use the control $u = -Lx + v$, where $L = R^{-1}B^TP$, and where P satisfies the ARE in (8.9). The closed loop system

$$
\begin{aligned}
\dot{x} &= (A - BL)x + Bv \\
y &= Lx
\end{aligned}
\tag{8.11}
$$

has the block diagram representation in the upper diagram in Figure 8.2. Let us consider the scalar case when $R = 1$. The inequality in (8.10) then reduces to

$$
|1 + G(j\omega)| \geq 1.
$$

This means that the Nyquist curve of $G(s) = L(sI - A)^{-1}B$ stays outside the circle centered at $s = -1$ and with radius 1. The interpretation of this is that $G(s)$ always is far away from the critical point $s = -1$. This is illustrated for the case when A is stable in the middle part of Figure 8.2 (remember that for this case the closed loop system is stable as long as the Nyquist curve does not encircle[3] $s = -1$). Another interpretation of this robustness property is that

[3] If A is unstable then the Nyquist diagram interpretation must be slightly modified. If A has p eigenvalues in the right half plane then the Nyquist curve of G must encircle the critical point p times in anticlockwise direction. The robustness interpretation is still valid since $G(j\omega)$ stays outside the unit disc in the middle diagram of Figure 8.2, which implies that the encirclement condition remains satisfied when the transfer function is perturbed as long as $|\Delta(j\omega)G(j\omega)| < 1$.

the perturbed system in the lower part of Figure 8.2 remains stable as long as $|\Delta(j\omega)G(j\omega)| < 1$, where $\Delta(s)$ is an arbitrary stable transfer function. To see this, we need to show that $(1+\Delta(j\omega))G(j\omega)$ does not encircle $s=-1$ (again we assume that A is stable). We have

$$|1+(1+\Delta(j\omega))G(j\omega)| \geq |1+G(j\omega)| - |\Delta(j\omega)G(j\omega)| \geq 1 - |\Delta(j\omega)G(j\omega)| > 0,$$

which proves the claim.

Example 24. Now consider the inverted pendulum in the introductory example. If we solve the optimization problem (8.3) with $Q=I$, and $r=1$ then the ARE has the following positive definite solution

$$P = \begin{bmatrix} 2+\sqrt{2} & 1+\sqrt{2} \\ 1+\sqrt{2} & 1+\sqrt{2} \end{bmatrix}$$

and the optimal feedback control law is $u=-Lx$, where $L = B^T P = \begin{bmatrix} 1+\sqrt{2} & 1+\sqrt{2} \end{bmatrix}$. The closed loop system matrix

$$A - BL = \begin{bmatrix} 0 & 1 \\ -\sqrt{2} & -1-\sqrt{2} \end{bmatrix}$$

has eigenvalues in -1 and $-\sqrt{2}$. The Nyquist curve of $G(s) = L(sI-A)^{-1}B = (1+\sqrt{2})\dfrac{s+1}{s^2-1}$ is given in Figure 8.3. We see that it encircles the point $s=-1$ once, which is necessary in order to obtain closed loop stability since A has one unstable pole at $s=1$. Moreover, it lies outside the circle centered at $s=-1$ with radius 1, which gives the robustness property.

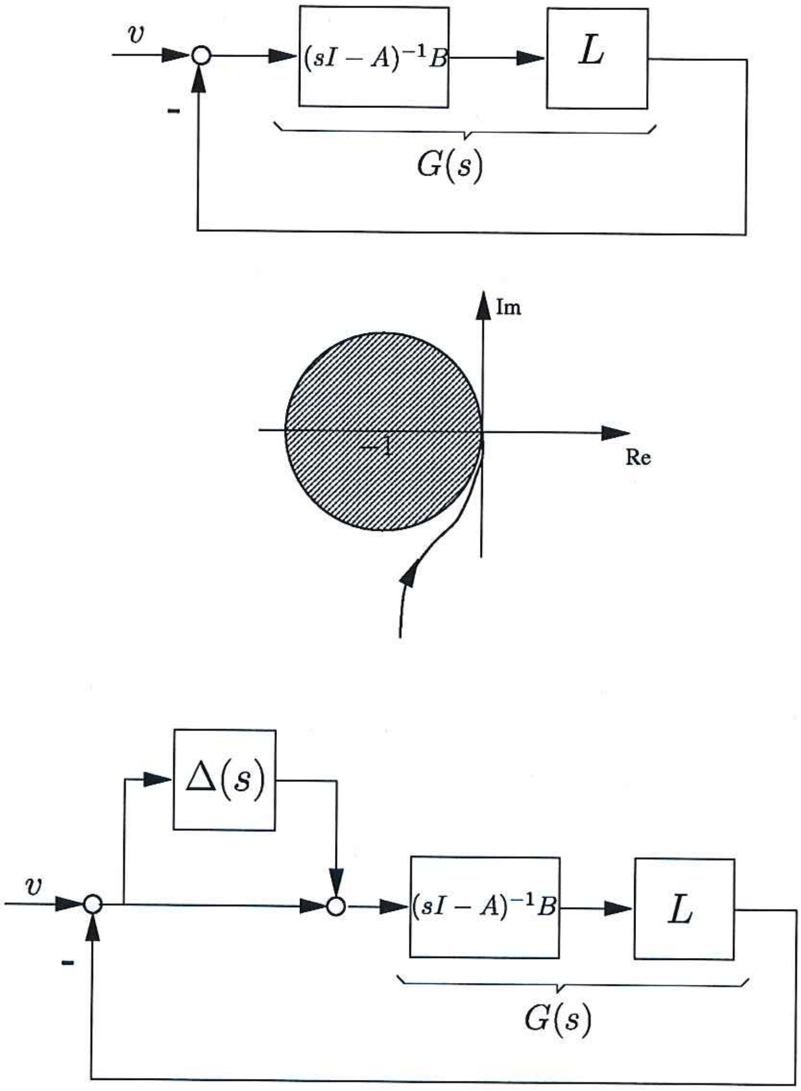

Figure 8.2: The upper figure shows a block diagram representation of the closed loop system in (8.11). The middle diagram illustrates that the Nyquist curve of $G(s) = L(sI - A)^{-1}B$ lies outside the unit ball centered at $s = -1$. This is due to the inequality (8.10). Another interpretation of inequality (8.10) is that the system in the lower diagram remains stable for all stable perturbations that satisfy $|\Delta(j\omega)G(j\omega)| < 1$.

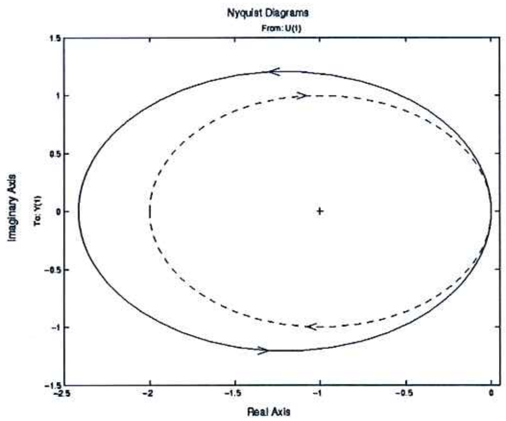

Figure 8.3: Nyquist curve corresponding to $G(s) = L(sI - A)^{-1}B = (1 + \sqrt{2})\dfrac{s+1}{s^2-1}$ is drawn in solid line. In dashed line we have the circle centered at $s = -1$ and with radius 1.

Chapter 9

Second Order Variations

We will in this chapter derive the second order variation of optimal control problems. This will be used to derive sufficient conditions for local minima. In the next chapter we use the second variation to derive Newton's algorithm for numerical computation of the optimal controller. For simplicity, we only consider optimization problems with fixed terminal time and no control constraint (we have for convenience assumed $t_0 = 0$):

$$\text{minimize}\,\phi(x(t_f)) + \int_0^{t_f} f_0(t, x(t), u(t))dt \quad \text{subj to} \quad \begin{cases} \dot{x} = f(t, x(t), u(t)) \\ x(0) = x_0, \end{cases} \tag{9.1}$$

where ϕ, f_0, and f are twice continuously differentiable with respect to x and u.

Define the Lagrange function (compare with Chapter 5)

$$l(u(\cdot), \lambda(\cdot)) = \phi(x(t_f)) + \int_0^{t_f} [H(t, x(t), u(t), \lambda(t)) - \lambda(t)^T \dot{x}(t)]dt \tag{9.2}$$

where

$$H(t, x, u, \lambda) = f_0(t, x, u) + \lambda^T f(t, x, u)$$

and where $\lambda(\cdot)$ is the adjoint vector.

Let us make a Taylor series expansion up to second order around an admissible solution $(x^0(\cdot), u^0(\cdot))$. Our experience from previous chapters suggests the following choice for the adjoint variable

$$\dot{\lambda}(t) = -H_x(t, x^0(t), u^0(t), \lambda(t)), \quad \lambda(t_f) = \phi_x(x^0(t_f))$$

Use of similar calculations as in Chapter 5 then gives the *second order variation*

$$l(u(\cdot), \lambda(\cdot)) = l(u^0(\cdot), \lambda(\cdot)) + \int_0^{t_f} H_u^0(t)^T \delta u(t)dt + \frac{1}{2}\Bigg(\delta x(t_f)^T \phi_{xx}(x^0(t_f))\delta x(t_f) + \\ \int_0^{t_f} \begin{bmatrix} \delta x(t) \\ \delta u(t) \end{bmatrix}^T \begin{bmatrix} H_{xx}^0(t) & H_{xu}^0(t) \\ H_{ux}^0(t) & H_{uu}^0(t) \end{bmatrix} \begin{bmatrix} \delta x(t) \\ \delta u(t) \end{bmatrix} dt\Bigg) + H.O.T. \tag{9.3}$$

where

$$H_u^0(t) = H_u(t, x^0(t), u^0(t), \lambda(t))$$
$$H_{xx}^0(t) = H_{xx}(t, x^0(t), u^0(t), \lambda(t))$$

and similarly for $H_{xu}^0(t)$, $H_{ux}^0(t)$, and H_{xx}^0(t). The second term in (9.3) corresponds to the "gradient of the Lagrangian". The gradient is a linear operator, which acts on a function $\delta u(\cdot)$ as follows

$$\nabla l(u^0(\cdot), \lambda(\cdot))(\delta u(\cdot)) = \int_0^{t_f} H_u^0(t)^T \delta u(t) dt$$

The third term corresponds to the "Hessian of the Lagrangian". The Hessian is a quadratic operator defined as follows

$$D^2 l(u^0(\cdot), \lambda(\cdot))(\delta x(\cdot), \delta u(\cdot)) = \delta x(t_f)^T \phi_{xx}(x^0(t_f))\delta x(t_f) + \int_0^{t_f} \begin{bmatrix} \delta x(t) \\ \delta u(t) \end{bmatrix}^T \begin{bmatrix} H_{xx}^0(t) & H_{xu}^0(t) \\ H_{ux}^0(t) & H_{uu}^0(t) \end{bmatrix} \begin{bmatrix} \delta x(t) \\ \delta u(t) \end{bmatrix} dt \quad (9.4)$$

It should not come as a surprise that we have the following result, which is analogous to the second order sufficient conditions for optimality in nonlinear programming[1], see Section 4.3

Proposition 8. *Suppose* $(x^*(\cdot), u^*(\cdot))$, *and* $\lambda(\cdot)$ *are such that*

(*i*) $\dot{x}^*(t) = f(t, x^*(t), u(t))$, $x^*(0) = x_0$,

(*iia*) $\dot{\lambda}(t) = -H_x(t, x^*(t), u^*(t), \lambda(t))$, $\lambda(t_f) = \phi_x(x^*(t_f))$

(*iib*) $H_u(t, x^*(t), u^*(t), \lambda(t)) = 0$

(*iiia*) $\phi_{xx}(x^*(t_f)) \geq 0$

(*iiib*) $H_{uu}^*(t) > 0$ *and* $\begin{bmatrix} H_{xx}^*(t) & H_{xu}^*(t) \\ H_{ux}^*(t) & H_{uu}^*(t) \end{bmatrix} \geq 0$, *where* $H_{uu}^*(t) = H_{uu}(t, x^*(t), u^*(t), \lambda(t))$ *and similarly for* H_{xx}^*, H_{xu}^* *and* H_{ux}^*.

Then $(x^*(\cdot), u^*(\cdot))$ *is a local minimum of (9.1).*

Remark 22. Conditions (*iia*) and (*iib*) correspond to $\nabla l(u^*(\cdot), \lambda(\cdot)) = 0$. Conditions (*iiia*) and (*iiib*) make the Hessian positive definite on the subspace of all solutions to the linearized dynamics

$$\delta \dot{x}(t) = f_x(t, x^*(t), u^*(t))\delta x(t) + f_u(t, x^*(t), u^*(t))\delta u(t), \quad \delta x(0) = 0. \quad (9.5)$$

This positive definiteness condition can be relaxed significantly.

[1]As in Proposition 6 in Chapter 5, we need the linearization of the dynamics at the optimal solution to be controllable.

Example 25. Consider the special case when the dynamics is linear

$$\min \phi(x(t_f)) + \int_0^{t_f} [g(t, x(t)) + h(t, u(t))]dt \quad \text{subj to} \quad \begin{cases} \dot{x} = Ax(t) + Bu(t) \\ x(0) = x_0, \end{cases} \tag{9.6}$$

Then condition *(iiib)* reduces to

$$Q(t) = \frac{\partial^2 g}{\partial x^2}(t, x^*(t)) \geq 0$$
$$R(t) = \frac{\partial^2 h}{\partial u^2}(t, u^*(t)) > 0$$

Hence, if the cost function in (9.6) locally around the extremal solution $(x^*(\cdot), u^*(\cdot))$ appears as an LQ optimal control problem with the usual assumptions on Q and R, then the extremal is a local minimum.

Chapter 10

Computational Algorithms

We will in this chapter present the five common methods for numerical solution of optimal control problems, viz.

Discretization: This is perhaps the most straightforward method. The idea is to discretize the continuous time dynamics and cost function. The optimal control problem then becomes a constrained nonlinear program.

Boundary condition iteration (Shooting): The idea behind the shooting method is to use numerical iteration to find the correct values of the unspecified initial/terminal conditions in the two point boundary value problem associated with PMP.

First order gradient methods: The idea behind these algorithms is to successively improve the control signal until the gradient of the cost function becomes sufficiently small.

Newton's method: The second order approximation of the Lagrangian of the cost function is minimized subject to the linearized dynamics and linearized boundary conditions. This approach is completely analogous to Newton's method in nonlinear programming.

Consistent approximations: The original optimal control problem is converted into a nonlinear program by using finite dimensional approximations of the control function (and sometimes also the state function). Different basis functions that are used to approximate the control give rise to different algorithms.

We will focus the discussion around the optimization problems

$$\text{minimize}\, \phi(x(t_f)) + \int_0^{t_f} f_0(t, x(t), u(t))dt \text{ subj. to } \begin{cases} \dot{x} = f(t, x(t), u(t)) \\ x(0) = x_0,\ x(t_f) \in S_f \\ u \in \mathbf{R}^m \end{cases} \tag{10.1}$$

where

$$S_f = \{x \in \mathbf{R}^n : G(x) = 0\}, \qquad G(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{bmatrix},$$

and all functions are twice continuously differentiable with respect to x and u.

Extension of the algorithms to the case of control variable constraints and free terminal time will be remarked on where we find it appropriate.

10.1 Discretization Methods

Discretization is perhaps the first approach one would think of to solve a continuous time optimal control problem. The most basic discretization method is to use a piecewise constant approximation of the control signal. If we use N samples and define $h = t_f/N$ then the discretized control becomes

$$u(t) = u_k, \quad kh \le t < (k+1)h, \quad k = 0, 1, \ldots, N-1$$

If we make a similar discretization of the state and use the forward Euler approximation of the differential operator, then we get the following approximation to (10.1)

$$\text{minimize}\, \phi(x_N) + \sum_{k=0}^{N-1} hf_0(kh, x_k, u_k) \quad \text{subj. to} \quad \begin{cases} x_{k+1} = x_k + hf(kh, x_k, u_k), \\ x_0 \text{ is given, } G(x_N) = 0 \end{cases} \tag{10.2}$$

This is a constrained nonlinear program. Indeed, if we define the parameter vector, objective function, and constraint function as

$$y = \begin{bmatrix} u_0^T & u_1^T & \ldots & u_{N-1}^T & x_1^T & \ldots & x_N^T \end{bmatrix}$$

$$\mathcal{F}(y) = \phi(x_N) + \sum_{k=0}^{N-1} hf_0(kh, x_k, u_k)$$

$$\mathcal{G}(y) = \begin{bmatrix} x_1 - x_0 - hf(0, x_0, u_0) \\ \vdots \\ x_N - x_{N-1} - hf((N-1)h, x_{N-1}, u_{N-1}) \\ G(x_N) \end{bmatrix}$$

then (10.2) becomes

$$\text{minimize}\, \mathcal{F}(y) \quad \text{subject to} \quad \mathcal{G}(y) = 0 \tag{10.3}$$

This optimization problem can be solved by any standard method for constrained nonlinear programming, e.g., Newton's method, quasi-Newton methods, penalty methods, or conjugate gradient methods. See, e.g., [15]. Some characteristics of and comments on the method are:

(+) There exist many good software packages for nonlinear optimization.

(+) There is a lot of structure in the optimization problem (10.2) and there are many ways to exploit it:

- A direct application of the first order necessary conditions for optimality (see Chapter 3) on (10.2) gives "Pontryagin minimum principle" for discrete time systems, see [4]. It is then often possible to eliminate the controller parameters from the optimization.
- The objective and constraint derivatives of (10.3) will be sparse and have a block structure. This can be exploited to make the solvers efficient.

(+) More sophisticated discretizations can be used. For example, Runge-Kutta formulas can be used to solve the differential equations.

(−) There are many variables and constraints in (10.3).

(−) The solution of the discretized problem may not converge to the solution of the original continuous time problem when the discretization is made finer and finer.

(−) We loose physical insight that we may have in the continuous time formulation.

10.2 Boundary condition iteration (Shooting)

This approach is based on successive improvements of the unspecified initial/terminal condition of the two point boundary value problem obtained from PMP. We assume for simplicity that G and ϕ depend on disjoint sets of state variables

- $G(x) = G(x_1, \ldots, x_p)$ (same p as number of components in g)
- $\phi(x) = \phi(x_{p+1}, \ldots, x_n)$.

This simplifies the transversality condition in PMP, see (6.12). With

$$H(t,x,u,\lambda) = f_0(t,x,u) + \lambda^T f(t,x,u)$$

the two point boundary value problem becomes

$$\dot{x} = H_\lambda(t,x,u,\lambda), \quad x(0) = x_0, \quad G(x(t_f)) = 0$$

$$\dot{\lambda} = -H_x(t,x,u,\lambda), \quad \begin{bmatrix} \lambda_{p+1}(t_f) \\ \vdots \\ \lambda_n(t_f) \end{bmatrix} = \begin{bmatrix} \frac{\partial \phi(x(t_f))}{\partial x_{p+1}} \\ \vdots \\ \frac{\partial \phi(x(t_f))}{\partial x_n} \end{bmatrix}$$

subject to $H_u(t,x,u,\lambda) = 0$, which at each time instant determines u as a function of (t,x,λ). The problem is to find an initial condition for $\lambda(\cdot)$ such that $\mu(x(t_f),\lambda(t_f)) = 0$, where

$$\mu(x(t_f),\lambda(t_f)) = \begin{bmatrix} g_1(x_1(t_f),\ldots,x_p(t_f)) \\ \vdots \\ g_p(x_1(t_f),\ldots x_p(t_f)) \\ \lambda_{p+1}(t_f) - \frac{\partial \phi(x(t_f))}{\partial x_{p+1}} \\ \vdots \\ \lambda_n(t_f) - \frac{\partial \phi(x(t_f))}{\partial x_n} \end{bmatrix} \tag{10.4}$$

The shooting method uses the following algorithm to update $\lambda(0)$.

Step 1 Make an initial guess $\lambda(0) = \lambda_0$

Step 2 Integrate the system

$$\begin{aligned} \dot{x}(t) &= H_\lambda(t,x(t),u(t),\lambda(t)), \quad x(0) = x_0 \\ \dot{\lambda}(t) &= -H_x(t,x(t),u(t),\lambda(t)), \quad \lambda(0) = \lambda_0 \end{aligned} \tag{10.5}$$

forward in time. Here $u(\cdot)$ is chosen such that $H_u(t,x(t),u(t),\lambda(t)) = 0$.

Step 3 Compute $\mu(t_f) = \mu(x(t_f),\lambda(t_f))$

Step 4 Update $\lambda_0 := \lambda_0 + \delta\lambda_0$, where $\delta\lambda_0 = -\alpha \left[\frac{\partial \mu(t_f)}{\partial \lambda(0)}\right]^{-1} \mu(t_f)$. Here the *transition matrix* $\frac{\partial \mu(t_f)}{\partial \lambda(0)}$ transfers a perturbation in $\lambda(0)$ into a perturbation in $\mu(t_f)$, i.e., $\delta\mu(t_f) = \frac{\partial \mu(t_f)}{\partial \lambda(0)} \delta\lambda(0)$. The parameter α is a step length.

Repeat steps 2 to 4 until $|\mu(t_f)|$ becomes sufficiently small.

The transition matrix $\frac{\partial \mu(t_f)}{\partial \lambda(0)}$ can be computed in either of the following ways

Numerical differentiation: Make n additional integrations, each with a slightly perturbed value of one of the components of $\lambda(0)$. The difference between the two values of $\mu(t_f)$ gives a column in the transition matrix.

Linearization: By linearizing the system (10.5) and the function $\mu(x, \lambda)$ in (10.4) we can compute the transition matrix after an integration of a linear Hamilton system, see the appendix to this chapter. This method gives better numerical accuracy than the numerical differentiation approach.

Some advantages and disadvantages of the shooting method are

(+) Conceptually simple. It was used to launch satellites in the 1950s.

(+) Control constraints are easy to deal with. The only modification is that in step 2 we must compute the control pointwise from the optimization problem $\min_{u \in U} H(t, x(t), u(t), \lambda(t))$.

(−) It can be crucial to find a good initial estimate of $\lambda(0)$.

(−) Integration of the system in (10.5) may be severely unstable. Indeed, the linearization of (10.5) in (10.12) is a linear Hamilton system. We know from Section 5.1 that such systems always have unstable modes. The transition matrix may therefore be ill conditioned.

Shooting methods are described in detail in [4].

10.3 Gradient Methods

The idea behind the gradient methods is to iteratively update the control signal in the direction of the negative gradient of the cost. We will only explain the gradient methods for the special case when there is no terminal constraint, i.e., the constraint $x(t_f) \in S_f$ is removed. More general cases are treated in [4].

As gradient we use

$$\nabla J(u(\cdot)) = H_u(t, x(\cdot), u(\cdot), \lambda(\cdot))$$

where H is the Hamiltonian, $x(\cdot)$ is the state trajectory, and $\lambda(\cdot)$ is the solution to the adjoint equation. From the previous chapter we know that this is the gradient of the Lagrangian $l(u(\cdot), \lambda(\cdot))$ introduced in (9.2). It can be proven that this is the correct gradient of the cost function as long as $\lambda(\cdot)$ satisfies the adjoint equation, [16].

The gradient algorithm becomes

Step 1 Guess $u(t)$, $t_0 \leq t \leq t_f$

Step 2 Integrate the system equation $\dot{x}(t) = f(t, x(t), u(t))$, $x(0) = x_0$ in the forward direction.

Step 3 Integrate the adjoint equation $\dot{\lambda}(t) = -H_x(t, x(t), u(t), \lambda(t))$, $\lambda(t_f) = \phi_x(x(t_f))$ in the backward direction.

Step 4 Update $u(t) := u(t) - \alpha H_u(t, x(t), u(t), \lambda(t))$, where α is the step length.

Repeat steps 2 to 4 until

$$\int_0^{t_f} |H_u(t, x(t), u(t), \lambda(t))|^2 dt$$

is sufficiently small. This means that the control is iterated until the conditions of PMP are satisfied.

Some characteristics of the gradient method

(+) It gives good improvement in the first iterations

(+) Stability is generally improved compared to the shooting method since the integration of $x(\cdot)$ and $\lambda(\cdot)$ is performed in the stable direction.

(+) Control constraints can be taken into account by projecting onto the control constraint set.

(+) It was used to solve a large number of aeronautical problems in the 1960s.

(−) Convergence tends to be slow.

10.4 Newton's Method

Newton's method in nonlinear programming is equivalent to solving a quadratic optimization problem at each iteration. The quadratic optimization problem is obtained by using the second order Taylor expansion of the Lagrangian as cost function and the first order Taylor expansion of the constraint function as a linear constraint in the optimization. The idea is completely analogous with Newton's method for nonlinear programming, which was presented in Section 4.3.

We will only consider the simplified case when the terminal condition $x(t_f) \in S_f$ is removed from (10.1). More general cases are treated in [20].

We use the Lagrangian defined in (9.2) when deriving the quadratic optimal control problem that must be solved in each iteration of Newton,s method. The second order Taylor expansion of the Lagrangian is given in (9.3) and the linearized constraint is given in (9.5). It follows that the quadratic optimal control problem becomes

$$\begin{aligned}&\text{minimize}\ \delta x(t_f)^T\psi_{xx}(x(t_f))\delta x(t_f) + \int_0^{t_f}\left(2H_u^T\delta u + \begin{bmatrix}\delta x\\ \delta u\end{bmatrix}^T\begin{bmatrix}H_{xx} & H_{xu}\\ H_{ux} & H_{uu}\end{bmatrix}\begin{bmatrix}\delta x\\ \delta u\end{bmatrix}\right)dt\\ &\text{subject to}\quad \delta\dot{x} = f_x\delta x + f_u\delta u,\quad \delta x(0) = 0\end{aligned} \tag{10.6}$$

where the functions are evaluated at the current iteration point

$$H_{xx}(t) = H_{xx}(t, x(t), u(t), \lambda(t))$$

and so on. The solution to this optimization problem can be obtained by solving one Riccati equation and one linear differential equation in a similar way as in the tracking example from the exercise.

$$\begin{aligned}&\dot{P} + PA + A^TP + Q = (PB+S)R^{-1}(PB+S)^T, \quad P(t_f) = \phi_{xx}(x(t_f))\\ &\dot{q} + A^Tq = (PB+S)R^{-1}(B^Tq + r), \quad q(t_f) = 0\end{aligned} \tag{10.7}$$

where

$$\begin{aligned}A(t) &= f_x(t, x(t), u(t)), \quad B(t) = f_u(t, x(t), u(t))\\ Q(t) &= H_{xx}(t, x(t), u(t), \lambda(t))\\ S(t) &= H_{xu}(t, x(t), u(t), \lambda(t))\\ R(t) &= H_{uu}(t, x(t), u(t), \lambda(t))\\ r(t) &= H_u(t, x(t), u(t), \lambda(t))\end{aligned}$$

Newton's algorithm becomes

Step 1 Guess $u(t)$, $0 \le t \le t_f$

Step 2 Integrate the system equation $\dot{x}(t) = f(t, x(t), u(t))$, $x(0) = x_0$ in the forward direction.

Step 3 Integrate the adjoint equation $\dot{\lambda}(t) = -H_x(t, x(t), u(t), \lambda(t))$, $\lambda(t_f) = \phi_x(x(t_f))$ in the backward direction.

Step 4 Solve for $P(\cdot)$ and $q(\cdot)$ in (10.7)

Step 5 Update, using the feedback law

$$u_{new} := u_{old} - \alpha R^{-1}[(B^TP + S^T)(x_{new} - x_{old}) + (r + B^Tq)]$$

where $\alpha \in (0, 1)$ is a step length parameter (should be equal to one in a proper Newton's method).

Repeat steps 2 to 5 until $\|H_u(t, x(t), u(t), \lambda(t))\|$ is sufficiently small. Some characteristics of the method are

(+) Fast convergence

(+) Solid theoretical justification

(−) Good initial guess is needed. Can be achieved by using a gradient method initially.

(−) Each iteration is computationally quite expensive.

10.5 Consistent Approximations

In consistent approximation methods the control and/or the state variables are approximated by a finite sum of basis functions. This makes the approximated optimal control problem a finite dimensional nonlinear optimization problem that can be solved using standard methods. The method has similarities with the discretization approach that was discussed previously but there are also significant differences. Particularly in the way the gradients are computed.

The following approximation methods can be used

Approximation by piecewise polynomials: The simplest case is when the control is approximated by a piecewise constant function. Then the control becomes

$$u(t|\mu) = \mu_k, \quad t \in [t_{k-1}, t_k), \quad t \in [0, t_f]$$

for $k = 1, \dots, N$, where the $\mu_k \in \mathbf{R}^m$ are parameters to be optimized and the t_k are the "sampling times". The notation $u(t|\mu)$ means that the control is parametrized by the vector $\mu = \begin{bmatrix}\mu_1^T & \mu_2^T & \dots & \mu_N^T\end{bmatrix}^T$. As N gets larger and larger we get better and better approximation capabilities.

More generally, we can use various spline functions to approximate the control function. For example, we could use

$$u(t|\mu) = \sum_{k=1}^{N} \mu_k B_k(t) = \sum_{k=1}^{N} \begin{bmatrix} B_k(t)\mu_{k_1} \\ \vdots \\ B_k(t)\mu_{k_n} \end{bmatrix}, \quad t \in [0, t_f]$$

where $\mu_k \in \mathbf{R}^m$ and $B_k(t)$ is a scalar spline function.

Approximation by orthogonal functions: It is very common to use an orthogonal basis for the approximation. In this case we let

$$u(t|\mu) = \sum_{k=1}^{N} \mu_k \varphi_k(t) = \sum_{k=1}^{N} \begin{bmatrix} \varphi_k(t)\mu_{k_1} \\ \vdots \\ \varphi_k(t)\mu_{k_n} \end{bmatrix}, \quad t \in [0, T] \qquad (10.8)$$

where $\mu_k \in \mathbf{R}^m$ and the $\phi_k(\cdot)$ are the (scalar) basis functions. The orthogonality of the basis means that we have the relations

$$\int_0^{t_f} \varphi_i(t)\varphi_j(t)dt = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

If we use the basis $\{\varphi_k\}_{k=0}^{\infty}$ to approximate a given control $u(t)$ then the N^{th} order approximation is obtained by letting the coefficients in (10.8) be obtained as

$$\mu_k = \begin{bmatrix} \int_0^T \varphi_k(t)u_1(t)dt \\ \vdots \\ \int_0^T \varphi_k(t)u_m(t)dt \end{bmatrix} \tag{10.9}$$

Common choices of basis are Fourier series, Chebyshev series, Legendre series, and Walsh series.

Let us now consider how the optimization problem will look like when we approximate the control with the expansion

$$u(t|\mu) = \sum_{k=1}^{N} \mu_k \varphi_k(t)$$

where

$$\mu = \begin{bmatrix} \mu_1^T & \mu_2^T & \dots & \mu_N^T \end{bmatrix}^T$$

are the parameters to be optimized. For each parameter vector μ, the state function $x(\cdot)$ will be completely defined by the differential equation

$$\dot{x}(t|\mu) = \hat{f}(t, x(t|\mu), \mu), \quad x(0) = x_0$$

where $\hat{f}(t, x, \mu) = f(t, x, u(t|\mu))$. If we define

$$\hat{f}_0(t, x, \mu) = f_0(t, x, u(t|\mu))$$

then the optimal control problem (10.1) can be formulated as

$$\text{minimize}\, \mathcal{F}(\mu) \quad \text{subject to} \quad \mathcal{G}(\mu) = 0 \tag{10.10}$$

where

$$\mathcal{F}(\mu) = \phi(x(t_f|\mu)) + \int_0^{t_f} \hat{f}_0(t, x(t|\mu), \mu)dt$$
$$\mathcal{G}(\mu) = G(x(t_f|\mu))$$

It can be solved using standard algorithms for nonlinear programming. One important aspect is that the gradients of $\mathcal{F}$ and $\mathcal{G}$ can be computed using the adjoint system as follows

$$\frac{\partial \mathcal{F}(\mu)}{\partial \mu} = \int_0^T \left(\frac{\partial \hat{f}_0}{\partial \mu}(t, x(t|\mu), \mu) + \frac{\partial \hat{f}}{\partial \mu}(t, x(t|\mu), \mu)^T \lambda_{\mathcal{F}}(t|\mu) \right) dt,$$
$$\frac{\partial \mathcal{G}(\mu)}{\partial \mu} = \int_0^T \left(\frac{\partial \hat{f}}{\partial \mu}(t, x(t|\mu), \mu)^T \lambda_{\mathcal{G}}(t|\mu) \right) dt$$

where

$$\dot{\lambda}_{\mathcal{F}}(t) = -\frac{\partial H_{\mathcal{F}}}{\partial x}(t, x(t|\mu), \mu, \lambda_{\mathcal{F}}(t)), \quad \lambda_{\mathcal{F}}(t_f) = \frac{\partial \phi}{\partial x}(x(t_f|\mu))$$
$$\dot{\lambda}_{\mathcal{G}}(t) = -\frac{\partial H_{\mathcal{G}}}{\partial x}(t, x(t|\mu), \mu, \lambda_{\mathcal{G}}(t)), \quad \lambda_{\mathcal{G}}(t_f) = \frac{\partial G}{\partial x}(x(t_f|\mu))$$

and where

$$H_{\mathcal{F}}(t, x, \mu, \lambda) = \hat{f}_0(t, x, \mu) + \lambda^T \hat{f}(t, x, \mu)$$
$$H_{\mathcal{G}}(t, x, \mu, \lambda) = \lambda^T \hat{f}(t, x, \mu)$$

The computation of the above integrals must be performed numerically. Some of the characteristics of the method are

(+) Solid theoretical justification

(+) There are several professional software packages developed based on the consistent approximation ideas:

- Riots, which was developed at UC Berkeley, see [24, 25].
- Miser3, which was developed at University of Western Australia, see [8].

A good survey of consistent approximation methods are given in [17].

10.6 Appendix

We will see how to compute the transition matrix $\frac{\partial \mu(t_f)}{\partial \lambda(0)}$ by linearizing the system (10.5). Linearization of the system gives (where we omit the arguments for brevity)

$$\begin{aligned} \delta\dot{x} &= H_{\lambda x}\delta x + H_{\lambda u}\delta u \\ \delta\dot{\lambda} &= -H_{xx}\delta x - H_{x\lambda}\delta\lambda - H_{xu}\delta u \end{aligned} \tag{10.11}$$

and a linearization of the constraint $H_u = 0$ gives $H_{ux}\delta x + H_{u\lambda}\delta\lambda + H_{uu}\delta u = 0$. If H_{uu} is invertible then we can solve for $\delta u = -H_{uu}^{-1}(H_{ux}\delta x + H_{u\lambda}\delta\lambda)$. If we plug this into (10.11) we get

$$\begin{bmatrix}\delta\dot{x}\\ \delta\dot{\lambda}\end{bmatrix} = \underbrace{\begin{bmatrix}\mathcal{A} & -\mathcal{R}\\ -\mathcal{Q} & -\mathcal{A}^T\end{bmatrix}}_{\mathcal{H}}\begin{bmatrix}\delta x\\ \delta\lambda\end{bmatrix} \tag{10.12}$$

where

$$\begin{aligned}\mathcal{A} &= H_{\lambda x} - H_{\lambda u}H_{uu}^{-1}H_{ux}\\ \mathcal{R} &= H_{\lambda u}H_{uu}^{-1}H_{u\lambda}\\ \mathcal{Q} &= H_{xx} - H_{xu}H_{uu}^{-1}H_{ux}\end{aligned}$$

Let Φ be the transition matrix corresponding to the time-varying linear system in (10.12). Then

$$\begin{bmatrix}\delta x(t_f)\\ \delta\lambda(t_f)\end{bmatrix} = \begin{bmatrix}\Phi_{11}(t_f,0) & \Phi_{12}(t_f,0)\\ \Phi_{21}(t_f,0) & \Phi_{22}(t_f,0)\end{bmatrix}\begin{bmatrix}0\\ \delta\lambda(0)\end{bmatrix}$$

where the zero in the right vector is due to the constraint $x(0) = x_0$ is fixed. We also have

$$\delta\mu(x(t_f),\lambda(t_f)) = \underbrace{\begin{bmatrix}G_x(x(t_f)) & 0_{p\times(n-p)} & 0_{p\times p} & 0_{p\times(n-p)}\\ 0_{(n-p)\times(n-p)} & -\phi_{xx}(x(t_f)) & 0_{(n-p)\times p} & I_{(n-p)\times(n-p)}\end{bmatrix}}_{\mathcal{M}}\begin{bmatrix}\delta x(t_f)\\ \delta\lambda(t_f)\end{bmatrix}$$

where the derivatives are evaluated as

$$G_x = \begin{bmatrix}\frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_p}\\ \vdots & & \vdots\\ \frac{\partial g_p}{\partial x_1} & \cdots & \frac{\partial g_p}{\partial x_p}\end{bmatrix} \qquad \phi_{xx} = \begin{bmatrix}\frac{\partial^2\phi}{\partial x_{p+1}\partial x_{p+1}} & \cdots & \frac{\partial^2\phi}{\partial x_{p+1}\partial x_n}\\ \vdots & & \vdots\\ \frac{\partial^2\phi}{\partial x_n\partial x_{p+1}} & \cdots & \frac{\partial^2\phi}{\partial x_n x_n}\end{bmatrix}$$

Hence, if we combine these two differentials we get

$$\frac{\partial\mu(t_f)}{\partial\lambda(0)} = \mathcal{M}\begin{bmatrix}\Phi_{12}(t_f,0)\\ \Phi_{22}(t_f,0)\end{bmatrix}.$$

Bibliography

[1] V. I. Arnold. *Mathematical Methods of Classical Mechanics.* Springer, 1989.

[2] V. I. Arnold. *Ordinary Differential Equations.* Springer-Verlag, Berlin, 1992.

[3] R.W. Brockett. *Finite Dimensional Linear Systems.* John Wiley and sons inc, New York, 1970.

[4] A. E. Bryson and Y. C. Ho. *Applied Optimal Control.* Blaisdell Waltham, 1969.

[5] Frank H. Clarke. *Optimization and Nonsmooth Analysis.* J. Wiley & Sons, New York, 1983.

[6] E. A. Coddington and N. Levinson. *Theory of Ordinary Differential Equations.* McGraw-Hill, New York, 1955.

[7] M. Egerstedt. *Motion Planning and Control of Mobile Robots.* PhD thesis, Report TRITA-MAT-2000-OS01, Division of Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden, 2000.

[8] L. S. Jennings, M. E. Fisher, K. L. Teo, and C. J. Goh. *MISER3, Optimal Control Software, Theory and User Manual*, 1991. Can be found at http://www.cado.uwa.edu.au/miser/.

[9] U. Jönsson, C. F. Martin, and Y. Zhou. Trajectory planning under a stochastic uncertainty. In *Fifteenth International Symposium on Mathematical Theory of Networks and Systems*, 2002.

[10] R. Kalman. Contributions to the theory of optimal control. *Boletin Sociedad Matematica Mexicana*, 5(1), 1960.

[11] H. K. Khalil. *Nonlinear Systems.* Macmillan, 1992.

[12] E. B. Lee and L. Marcus. *Foundations of Optimal Control.* Wiley, New York, 1967.

[13] George Leitmann. *The Calculus of Variations and Optimal Control.* Plenum Press, New York, 1981.

[14] A. Lindquist and J. Sand. *An Introduction to Mathematical System Theory.* Optimization and Systems Theory, KTH.

[15] D. G. Luenberger. *Linear and Nonlinear Programming.* Addison-Wesley Publishing Company, 1984.

[16] D.G. Luenberger. *Optimization by Vector Space Methods.* Wiley, New York, 1969.

[17] B. Ma. *An Improved Algorithm for Solving Constrained Optimal Control Problems.* PhD thesis, Institute for Systems Research, University of Maryland, 1994.

[18] J. Macki and A. Strauss. *Introduction to Optimal Control Theory.* Springer-Verlag, 1982.

[19] C. F. Martin, S. Sun, and M. Egerstedt. Optimal control, statistics, and path planning. To appear in Mathematical and Computer Modeling.

[20] S. K. Mitter. Successive approximation methods for the solution of optimal control problems. *Automatica*, 3:135–149, 1966.

[21] S. G. Nash and A. Sofer. *Linear and Nonlinear Programming.* Mc-Graw-Hill International Editions, 1996.

[22] E. R. Pinch. *Optimal Control and the Calculus of Variations.* Oxford Science Publications, 1993.

[23] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The Mathematical Theory of Optimal Processes.* Interscience Publishers, 1962.

[24] A. Schwartz. *Theory and Implementation of Numerical Methods Based on Runge-Kutta Integration for Solving Optimal Control Problems.* PhD thesis, U.C. Berkeley, 1996.

[25] A. Schwartz, E. Polak, and Y. Chen. *RIOTS, A Matlab Toolbox for Solving Optimal Control Problems*, 1997. Can be found at http://www.accesscom.com/~adam/RIOTS/.

[26] S. Sun, M. Egerstedt, and C. F. Martin. Control theoretic smothing splines. IEEE Transactions on Automatic Control. Accepted for publication 1999.

[27] H. J. Sussman and W. Tang. Shortest paths for the Reeds-Shepp car: a worked out example of the use of geometric techniques in nonlinear optimal control. Technical report, Report SYCON-91-10, Rutgers University, 1991.