Chapter (1)

Controllability, bang-bang principle

Section (1.1): Introduction to the basic problem

To discuss the dynamics, we open our discussion by considering an ordinary differential equation (ODE) having the form

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}\big(x(t)\big) \\ \mathbf{x}(0) = x^0. \end{cases} \qquad (t > 0) \tag{1.1}$$

We are here given the initial point $x^0 \in \mathbb{R}^n$ and the function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$. The unknown is the curve $\mathbf{x}: [0, \infty) \to \mathbb{R}^n$, which we interpret as the dynamical evolution of the state of some "system".

To discuss controlled dynamics, we generalize a bit and suppose now that $\mathbf{f}$ depends also upon some "control" parameters belonging to a set $A \subset \mathbb{R}^m$; so that $\mathbf{f}: \mathbb{R}^n \times A \to \mathbb{R}^n$. Then if we select some value $a \in A$ and consider the corresponding dynamics:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(x(t), a) \qquad (t > 0) \\ \mathbf{x}(0) = x^0, \end{cases}$$

we obtain the evolution of our system when the parameter is constantly set to the value a.

The next possibility is that we change the value of the parameter as the system evolves. For instance, suppose we define the function $\boldsymbol{\alpha}: [0, \infty) \to A$ this way:

$$\alpha(t) = \begin{cases} a_1 & 0 \le t \le t_1 \\ a_2 & t_1 < t \le t_2 \\ a_3 & t_2 < t \le t_3 \ \text{etc.} \end{cases}$$

for times $0 < t_1 < t_2 < t_3 \ldots$ and parameter values $a_1, a_2, a_3, \ldots \in A$; and we then solve the dynamical equation

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}\big(x(t), \alpha(t)\big) \quad (t > 0) \\ \mathbf{x}(0) = x^0, \end{cases}$$

The picture illustrates the resulting evolution. The point is that the system may behave quite differently as we change the control parameters.

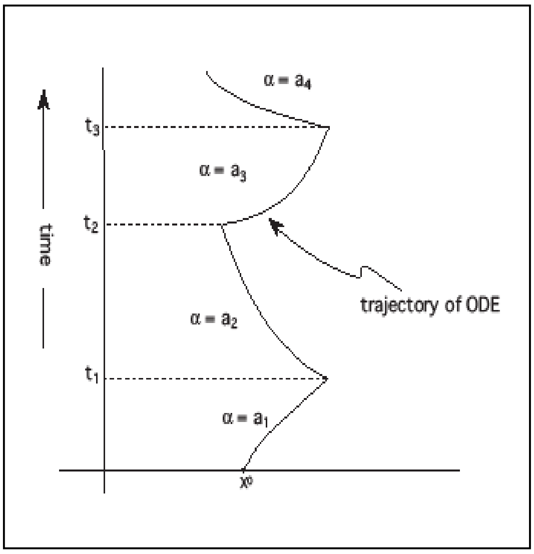

Figure (1.1) Controlled dynamics

More generally, we call a function $\alpha\colon [0, \infty) \to A$ a control. Corresponding to each control, we consider the ODE

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}\big(x(t), \alpha(T)\big) & (t > 0) \\ \mathbf{x}(0) = x^0, \end{cases}$$

and regard the trajectory $\mathbf{x}(\cdot)$ as the corresponding response of the system.

Notation (1.1.1):

(i) We will write

$$\mathbf{f}(x, a) = \begin{pmatrix} f^1(x, a) \\ \vdots \\ f^n(x, a) \end{pmatrix}$$

to display the components of $\mathbf{f}$, and similarly put

$$\mathbf{x}(t) = \begin{pmatrix} x^1(t) \\ \vdots \\ x^n(t) \end{pmatrix}.$$

We will therefore write vectors as columns in these notes and use boldface for vector-valued functions, the components of which have superscripts.

(ii) We also introduce

$$\mathcal{A} = \{\alpha: [0, \infty) \to A | \alpha(\cdot) \text{ measureable}\}$$

to denote the collection of all admissible controls, where

$$\boldsymbol{\alpha}(t) = \begin{pmatrix} \alpha^1(t) \\ \vdots \\ \alpha^m(t) \end{pmatrix}.$$

Note very carefully that our solution $\mathbf{x}(\cdot)$ of (ODE) depends upon $\alpha(\cdot)$ and the initial condition. Consequently our notation would be more precise, but more complicated, if we were to write

$$\mathbf{x}(\cdot) = \mathbf{x}(\cdot, \boldsymbol{\alpha}(\cdot), x^0),$$

displaying the dependence of the response $\mathbf{x}(\cdot)$ upon the control and the initial value.

Our overall task will be to determine what is the "best" control for our system. For this we need to specify a specific payoff (or reward) criterion. Let us define the payoff functional

$$\text{(P)} \qquad P[\boldsymbol{\alpha}(\cdot)] := \int_0^T r\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big)dt + g(\mathbf{x}(T)),$$

where $\mathbf{x}(\cdot)$ solves (ODE) for the control $\boldsymbol{\alpha}(\cdot)$. Here $r: \mathbb{R}^n \times A \to \mathbb{R}$ and $g: \mathbb{R}^n \to \mathbb{R}$ are given, and we call r the running payoff and g the terminal payoff. The terminal time $T > 0$ is given as well.

Our basic problem is to find a control $\boldsymbol{\alpha}^*(\cdot)$, which maximizes the payoff. In other words, we want

$$P[\boldsymbol{\alpha}^*(\cdot)] \geq P[\boldsymbol{\alpha}(\cdot)]$$

for all controls $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$. Such a control $\boldsymbol{\alpha}^*(\cdot)$ is called optimal.

This task presents us with these mathematical issues:

(i) Does an optimal control exist?
(ii) How can we characterize an optimal control mathematically?
(iii) How can we construct an optimal control?

These turn out to be sometimes subtle problems, as the following collection of examples illustrates.

Example (1.1.2): (Control of Production and Consumption)

Suppose we own, say, a factory whose output we can control. Let us begin to construct a mathematical model by setting

$$x(t) = \text{amount of output produced at time } t \geq 0.$$

We suppose that we consume some fraction of our output at each time, and likewise can reinvest the remaining fraction. Let us denote

$$\alpha(t) = \text{fraction of output reinvested at time } t \geq 0.$$

This will be our control, and is subject to the obvious constraint that

$$0 \leq \alpha(t) \leq 1 \text{ for each time } t \geq 0.$$

Given such a control, the corresponding dynamics are provided by the ODE

$$\begin{cases} \dot{x}(t) = \kappa\alpha(t)x(t) \\ x(0) = x^0. \end{cases}$$

the constant $\kappa > 0$ modelling the growth rate of our reinvestment. Let us take as a payoff functional

$$P[\boldsymbol{\alpha}(\cdot)] = \int_0^T \big(1 - \alpha(t)\big)x(t)\,dt.$$

The meaning is that we want to maximize our total consumption of the output, our consumption at a given time t being $\big(1 - \alpha(t)\big)x(t)$. This model fits into our general framework for $n = m = 1$, once we put

$$A = [0,1], f(x,a) = \kappa ax, \qquad r(x,a) = (1-a)x, \qquad g \equiv 0.$$

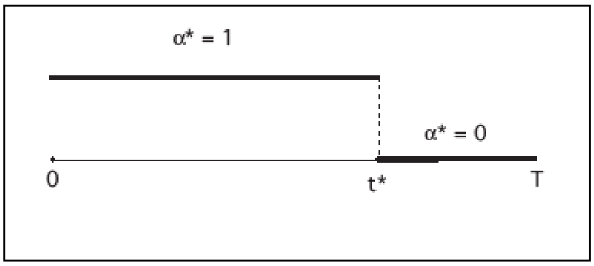

Figure (1.2) A bang-bang control

As we will see later in section (2.2), an optimal control $\alpha^*(\cdot)$ is given by

$$\alpha^*(t) = \begin{cases} 1 & \text{if } \ 0 \le t \le t^* \\ 0 & \text{if } \ t^* < t \le T \end{cases}$$

for an appropriate switching time $0 \le t \le T$. In other words, we should reinvest all the output (and therefore consume nothing) up until time t^*, and afterwards, we should consume everything (and therefore reinvest nothing). The switchover time t^* will have to be determined. We call $\alpha^*(\cdot)$ a bang.bang control.

Example (1.1.3): (Reproductive Stategies in Social Insects)

We attempt to model how social insects, say a population of bees, determine the makeup of their society. [7]

Let us write T for the length of the season, and introduce the variables

$w(t)$ = number of workers at time t

$q(t)$ = number of queens

$\alpha(t)$ = fraction of colony effort devoted to increasing work force

The control α is constrained by our requiring that

$$0 \le \alpha(t) \le 1.$$

We continue to model by introducing dynamics for the numbers of workers and the number of queens. The worker population evolves according to

$$\begin{cases} \dot{w}(t) = -\mu w(t) + bs(t)\alpha(t)w(t) \\ w(0) = w^0. \end{cases}$$

Here μ is a given constant (a death rate), b is another constant, and $s(t)$ is the known rate at which each worker contributes to the bee economy.

We suppose also that the population of queens changes according to

$$\begin{cases} \dot{q}(t) = -\nu q(t) + c\big(1 - \alpha(t)\big)s(t)w(t) \\ q(0) = q^0, \end{cases}$$

for constants ν and c.

Our goal, or rather the bees', is to maximize the number of queens at time T:

$$P[\alpha(\cdot)] = q(T).$$

So in terms of our general notation, we have $\mathbf{x}(t) = (w(t), q(t))^T$ and $x^0 = (w^0, q^0)^T$. We are taking the running payoff to be $r \equiv 0$, and the terminal payoff $g(w, q) = q$.

The answer will again turn out to be a bang-bang control, as we will explain later.

Example (1.1.4): (A pendulum)

We look next at a hanging pendulum, for which

$$\theta(t) = \text{angle at time } t.$$

If there is no external force, then we have the equation of motion

$$\begin{cases} \ddot{\theta}(t) + \lambda\dot{\theta}(t) + w^2\theta(t) = 0 \\ \theta(0) = \theta_1, \dot{\theta}(0) = \theta_2; \end{cases}$$

the solution of which is a damped oscillation, provided $\lambda > 0$.

Now let $\alpha(\cdot)$ denote an applied torque, subject to the physical constraint that

$$|\alpha| \leq 1.$$

Our dynamics now become

$$\begin{cases} \ddot{\theta}(t) + \lambda\dot{\theta}(t) + w^2\theta(t) = \alpha(t) \\ \theta(0) = \theta_1, \dot{\theta}(0) = \theta_2; \end{cases}$$

Define $x_1(t) = \theta(t), x_2(t) = \dot{\theta}(t)$, and $\mathbf{x}(t) = (x_1(t), x_2(t))$. Then we can write the evolution as the system

$$\dot{x}(t) = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} \dot{\theta} \\ \ddot{\theta} \end{pmatrix} = \begin{pmatrix} x_2 \\ -\lambda x_2 - w^2 x_1 + \alpha(t) \end{pmatrix} = \mathbf{f}(\mathbf{x}, \alpha).$$

We introduce as well

$$P[\alpha(\cdot)] = -\int_0^\tau 1 \, dt = -\tau,$$

for

$$\tau = \tau(\alpha(\cdot)) = \text{first time that } \mathbf{x}(\tau) = 0 \quad (\text{that is, } \theta(\tau) = \dot{\theta}(\tau) = 0)$$

We want to maximize $P[\cdot]$, meaning that we want to minimize the time it takes to bring the pendulum to rest.

Observe that this problem does not quite fall within a general framework, since the terminal time is not fixed, but rather depends upon the control. This is called a fixed endpoint, free time problem.

Example (1.1.5): (A moon lander)

This model asks us to bring a spacecraft to a soft landing on the lunar surface, using the least amount of fuel.

We introduce the notation

$h(t)$ = height at time t

$v(t)$ = velocity = $\dot{h}(t)$

$m(t)$ = mass of spacecraft (changing as fuel is burned)

$\alpha(t)$ = thrust at time t

We assume that

$$0 \leq \alpha(t) \leq 1,$$

and Newton's law tells us that

$$m\ddot{h} = -gm + \alpha,$$

the right hand side being the difference of the gravitational force and the thrust of the rocket. This system is modeled by the ODE

$$\begin{cases} \dot{v}(t) = -g + \dfrac{\alpha(t)}{m(t)} \\ \dot{h}(t) = v(t) \\ \dot{m}(t) = -\kappa\alpha(t). \end{cases}$$

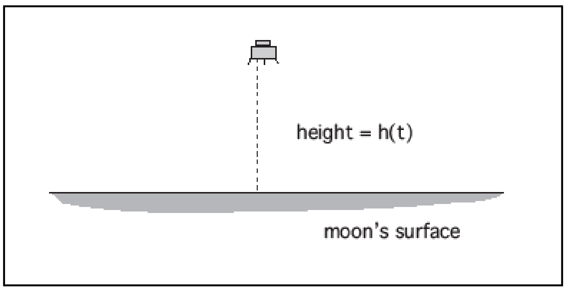

Figure (1.3) A spacecraft landing on the moon

We summarize these equations in the form

$$\dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t), \alpha(t)\big)$$

for $\mathbf{x}(t) = (v(t), h(t), m(t))$.

We want to minimize the amount of fuel used up, that is, to maximize the amount remaining once we have landed. Thus

$$P[\alpha(\cdot)] = m(\tau),$$

where τ denotes the first time that $h(\tau) = v(\tau) = 0$.

This is a variable endpoint problem, since the final time is not given in advance. We have also the extra constraints

$$h(t) \geq 0, \quad m(t) \geq 0.$$

Example (1.1.6): (Rocket Railroad Car)

Imagine a railroad car powered by rocket engines on each side. We introduce the variables

$q(t) =$ position at time t

$v(t) = \dot{q}(t) =$ velocity at time t

$\alpha(t) =$ thrust from rockets,

where

$$-1 \leq \alpha(t) \leq 1,$$

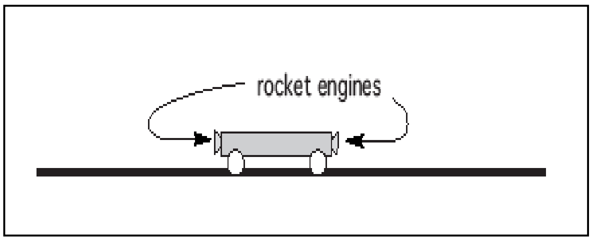

Figure (1.4) A rocket car on a train track

the sign depending upon which engine is firing.

We want to figure out how to fire the rockets, so as to arrive at the origin 0 with zero velocity in a minimum amount of time. Assuming the car has mass m, the law of motion is

$$m\ddot{q}(t) = \alpha(t).$$

We rewrite by setting $\mathbf{x}(t) = (q(t), v(t))^T$. Then

$$\begin{cases} \mathbf{x}(t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \mathbf{x}(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \alpha(t) \\ \mathbf{x}(0) = x^0 = (q_0, v_0)^T. \end{cases}$$

Since our goal is to steer to the origin $(0,0)$ in minimum time, we take

$$P[\alpha(\cdot)] = -\int_0^\tau 1\, dt = -\tau,$$

for

$$\tau = \text{first time that } q(\tau\ ) = v(\tau) = 0.$$

Now we discuss the Geometric solution:

To illustrate how actually to solve a control problem, in this last section we introduce some ad hoc calculus and geometry methods for the rocket car problem, Example (1.1.5) above.

First of all, let us guess that to find an optimal solution we will need only to consider the cases $a = 1$ or $a = -1$. In other words, we will focus our attention only upon those controls for which at each moment of time either the left or the right rocket engine is fired at full power. (We will later see in Chapter 2 some theoretical justification for looking only at such controls.)

Case 1:

Suppose first that $\alpha \equiv 1$ for some time interval, during which

$$\begin{cases} \dot{q} = v \\ \dot{v} = 1 \end{cases}$$

Then

$$v\dot{v} = \dot{q},$$

and so

$$\frac{1}{2}(v^2)^{\cdot} = \dot{q}.$$

Let t_0 belong to the time interval where $\alpha \equiv 1$ and integrate from t_0 to t:

$$\frac{v^2(t)}{2} - \frac{v^2(t_0)}{2} = q(t) - q(t_0).$$

Consequently

$$v^2(t) = 2q(t) + \underbrace{\left(v^2(t_0) - 2q(t_0)\right)}_{b}. \tag{1.3}$$

In other words, so long as the control is set for $\alpha \equiv 1$, the trajectory stays on the curve $v^2 = 2q + b$ for some constant b.

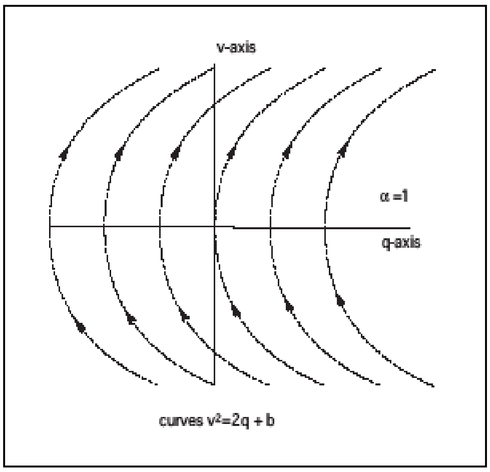

Figure (1.5)

Case 2:

Suppose now $\alpha \equiv 1$ on some time interval. Then as above

$$\begin{cases} \dot{q} = v \\ \dot{v} = -1, \end{cases}$$

and hence

$$\frac{1}{2}(v^2)^{\cdot} = -\dot{q}.$$

Let t_1 belong to an interval where $\alpha \equiv 1$ and integrate:

$$v^2(t) = -2q(t) + \underbrace{\left(2q(t_1) - v^2(t_1)\right)}_{c}. \tag{1.4}$$

Consequently, as long as the control is set for $\alpha \equiv 1$, the trajectory stays on the curve $v^2 = 2q + c$ for some constant c.

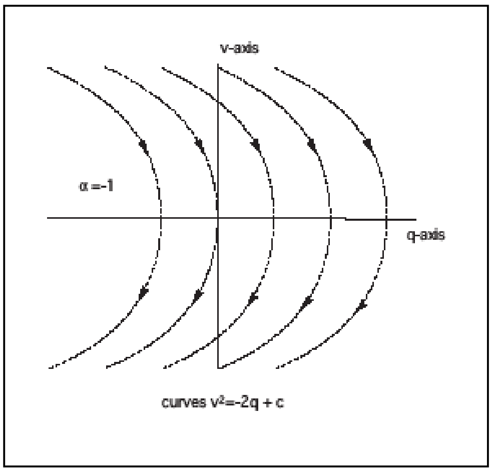

Figure (1.6)

Now to discuss the geometric interpretation, we have to know that formula (1.1) says if $\alpha \equiv 1$, then $(q(t), v(t))$ lies on a parabola of the form

$$v^2 = 2q + b.$$

Similarly, (1.2) says if $\alpha \equiv -1$, then $(q(t), v(t))$ lies on a parabola

$$v^2 = -2q + c.$$

Now we can design an optimal control $\alpha^*(\cdot)$, which causes the trajectory to jump between the families of right. and left-pointing parabolas, as drawn. Say we start at the black dot, and wish to steer to the origin. This we accomplish by first setting the control to the value $\alpha = -1$, causing us to move down along the second family of parabolas. We then switch to the control $\alpha = 1$, and thereupon move to a parabola from the first family, along which we move up and to the left, ending up at the origin. See the figure (1.5).

Section (1.2): Controllability, bang-bang principle

Definition (1.2.1):

We firstly recall from Chapter 1 the basic form of our controlled ODE:

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}\big(x(t), \boldsymbol{\alpha}(t)\big) \\ \mathbf{x}(0) = x^0, \end{cases}$$

Here $x^0 \in \mathbb{R}^n$, $\mathbf{f}: \mathbb{R}^n \times A \to \mathbb{R}^n$, $\boldsymbol{\alpha}: [0, \infty) \to A$, is the control, and $x: [0, \infty) \to \mathbb{R}^n$ is the response of the system.

This section addresses the following basic

(1.2.1): Controllability Question

Given the initial point x^0 and a "target" set $S \subset \mathbb{R}^n$, does there exist a control steering the system to S in finite time?

For the time being we will therefore not introduce any payoff criterion that would characterize an "optimal" control, but instead will focus on the question as to whether or not there exist controls that steer the system to a given goal. In this chapter we will mostly consider the problem of driving the system to the origin $S = \{0\}$.

Definition (1.2.2):

We define the reachable set for time t to be

$C(t) =$ set of initial points x^0 for which there exists a control such that $x(t) = 0$,

and the overall reachable set

$C =$ set of initial points x^0 for which there exists a control such that $\mathbf{x}(t) = 0$ for some finite time t.

Note that

$$C = \bigcup_{t \geq 0} C(t).$$

Hereafter, let $\mathbb{M}^{n\times m}$ denote the set of all $n \times m$ matrices. We assume for the rest of this and the next chapter that our ODE is linear in both the state $\mathbf{x}(\cdot)$ and the control $\alpha(\cdot)$, and consequently has the form

$$
\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) + N\boldsymbol{\alpha}(t) & (t > 0) \\ \mathbf{x}(0) = x^0, \end{cases}
$$

where $M \in \mathbb{M}^{n\times n}$ and $N \in \mathbb{M}^{n\times m}$. We assume the set A of control parameters is a cube in $\mathbb{R}^m$:

$$
A = [-1,1]^m = \{a \in \mathbb{R}^m |\ |a_i| \leq 1, i = 1, \dots, m\}.
$$

(1.2.2): Quick review of linear ODE

This section records for later reference some basic facts about linear systems of ordinary differential equations.

Definition (1.2.3):

Let $\mathbf{X}(\cdot) : \mathbb{R} \to \mathbb{M}^{n\times n}$ be the unique solution of the matrix

$$
\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{X}}(t) = M\mathbf{X}(t) & (t > 0) \\ \mathbf{x}(0) = I. \end{cases}
$$

We call $\mathbf{X}(\cdot)$ a fundamental solution, and sometimes write

$$
\mathbf{X}(t) = e^{tM} := \sum_{k=0}^{\infty} \frac{t^k M^k}{k!},
$$

the last formula being the definition of the exponential e^{tM}. Observe that

$$
\mathbf{X}^{-1}(t) = \mathbf{X}(-t).
$$

Theorem (1.2.4): (Solving linear systems of ODE)

(i) The unique solution of the homogeneous system of ODE

$$
\begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) \\ \mathbf{x}(0) = x^0 \end{cases}
$$

is

$$
\mathbf{X}(t) = \mathbf{X}(t)x^0 = e^{tM}x^0.
$$

(ii) The unique solution of the nonhomogeneous system

$$
\begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t)\mathbf{f}(t) \\ \mathbf{x}(0) = x^0. \end{cases}
$$

is

$$\mathbf{x}(t) = \mathbf{X}(t)x^0 + \mathbf{X}(t)\int_0^t \mathbf{X}^{-1}(s)\,\mathbf{f}(s)ds.$$

This expression is the variation of parameters formula.

(1.2.3): Controllability of Linear Equations

According to the variation of parameters formula, the solution of (ODE) for a given control $\boldsymbol{\alpha}(\cdot)$ is

$$\mathbf{x}(\cdot\ t) = \mathbf{X}(t)x^0 + \mathbf{X}(t)\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}(s)ds,$$

where $\mathbf{X}(t) = e^{tM}$. Furthermore, observe that

$$x^0 \in \mathcal{C}(t)$$

if and only if

$$\text{there exists a control } \alpha(\cdot) \in \mathcal{A} \text{ such that } \mathbf{x}(t) = 0 \tag{1.5}$$

if and only if

$$0 = \mathbf{X}(t)x^0 + \mathbf{X}(t)\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}(\cdot)ds \text{ for some control } \boldsymbol{\alpha}(\cdot) \in \mathcal{A} \tag{1.6}$$

if and only if

$$x^0 = -\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}(s)ds \text{ for some control } \boldsymbol{\alpha}(\cdot) \in \mathcal{A}. \tag{1.7}$$

We make use of these formulas to study the reachable set:

Theorem (1.2.5): (Structure of reachable set)

(i) The reachable set $\mathcal{C}$ is symmetric and convex.
(ii) Also, if $x^0 \in \mathcal{C}(t)$, then $x^0 \in \mathcal{C}(t)$ for all times $t \geq \bar{t}$.

Definition (1.2.6):

(i) We say a set S is symmetric if $x \in S$ implies $-x \in S$.
(ii) The set S is convex if $x, \hat{x} \in S$ and $0 \leq \lambda \leq 1$ imply $\lambda x + (1-\lambda)\hat{x} \in S$.

Proof:

1. (Symmetry) Let $t \geq 0$ and $x^0 \in \mathcal{C}(t)$. Then $x^0 = -\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}(s)ds$ for some admissible control $\boldsymbol{\alpha} \in \mathcal{A}$. Therefore $-x^0 = -\int_0^t \mathbf{X}^{-1}(s)N\big(-\boldsymbol{\alpha}(s)\big)ds$ and $-\boldsymbol{\alpha} \in \mathcal{A}$ since the set A is symmetric. Therefore $-x^0 \in \mathcal{C}(t)$, and so each set $\mathcal{C}(t)$ symmetric. It follows that $\mathcal{C}$ is symmetric.
2. (Convexity) Take $x^0, \hat{x}^0 \in \mathcal{C}$; so that $x^0 \in \mathcal{C}(t)$, $\hat{x}^0 \in \mathcal{C}(\hat{t})$ for appropriate times $t, \hat{t} \geq 0$. Assume $t \leq \hat{t}$. Then

$$x^0 = -\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}(s)ds \text{ for some control } \boldsymbol{\alpha} \in \mathcal{A},$$

$$\hat{x}^0 = -\int_0^{\hat{t}} \mathbf{X}^{-1}(s)N\widehat{\boldsymbol{\alpha}}(s)ds \text{ for some control } \widehat{\boldsymbol{\alpha}} \in \mathcal{A},$$

Define a new control

$$\widetilde{\boldsymbol{\alpha}}(s) := \begin{cases} \boldsymbol{\alpha}(s) & \text{if } 0 \leq s \leq t \\ 0 & \text{if } s > t. \end{cases}$$

Then

$$x^0 = -\int_0^{\hat{t}} \mathbf{X}^{-1}(s)N\widehat{\boldsymbol{\alpha}}(s)ds,$$

and hence $x^0 \in \mathcal{C}(t)$. Now let $0 \leq \lambda \leq 1$, and observe

$$\lambda x^0 + (1-\lambda)\hat{x}^0 = -\int_0^{\hat{t}} \mathbf{X}^{-1}(s)N\big(\lambda\widetilde{\boldsymbol{\alpha}}(s) + (1-\lambda)\,\widehat{\boldsymbol{\alpha}}(s)\big)ds.$$

Therefore $\lambda x^0 + (1-\lambda)\hat{x}^0 \in \mathcal{C}(\hat{t}) \subseteq \mathcal{C}$.

3. Assertion (ii) follows from the foregoing if we take $\bar{t} = \hat{t}$.

Example (1.2.7):

Let $n = 2$ and $m = 1$, $A = [-1,1]$, and write $\mathbf{x}(t) = \big(x^1(t), x^2(t)\big)^T$. Suppose

$$\begin{cases} \dot{x}^1 = 0 \\ \dot{x}^2 = \alpha(t). \end{cases}$$

This is a system of the form $\dot{\mathbf{x}} = M\mathbf{x} + N\alpha$, for

$$M = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, N = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Clearly $\mathcal{C} = \{(x_1, x_2) | x_1 = 0\}$, the x_2-axis.

We next wish to establish some general algebraic conditions ensuring that $\mathcal{C}$ contains a neighborhood of the origin.

Definition (1.2.8):

The controllability matrix is

$$G = G(M, N) := \underbrace{[N, MN, M^2N, \dots, M^{n-1}N]}_{n\times(mn) \text{ matrix}}.$$

Theorem (1.2.9): (Controllability Matrix)

We have

$$\operatorname{rank} G = n$$

if and only if

$$0 \in \mathcal{C}^0.$$

Notation (1.2.10):

We write $\mathcal{C}^0$ for the interior of the set $\mathcal{C}$. Remember that

$$\text{rank of } G = \text{number of linearly independent rows of } G$$

$$= \text{number of linearly independent columns of } G.$$

Clearly $\operatorname{rank} G \leq n$.

Proof:

1. Suppose first that $\operatorname{rank} G < n$. This means that the linear span of the columns of G has dimension less than or equal to $n - 1$. Thus there exists a vector $b \in \mathbb{R}^n, b \neq 0$, orthogonal to each column of G. This implies

$$b^T G = 0$$

and so

$$b^T N = b^T MN = \cdots = b^T M^{n-1} N = 0.$$

2. We claim next that in fact

$$b^T M^k N = 0 \text{ for all positive integers } k. \tag{1.8}$$

To confirm this, recall that

$$p(\lambda) \coloneqq \det(\lambda I - M)$$

is the characteristic polynomial of M. The Cayley-Hamilton Theorem states that

$$p(M) = 0.$$

So if we write

$$p(\lambda) = \lambda^n + \beta_{n-1}\lambda^{n-1} + \cdots + \beta_1\lambda^1 + \beta_0,$$

then

$$p(M) = M^n + \beta_{n-1}M^{n-1} + \cdots + \beta_1 M + \beta_0 I = 0.$$

Therefore

$$M^n = -\beta_{n-1}M^{n-1} - \beta_{n-2}M^{n-2} - \cdots - \beta_1 M - \beta_0 I,$$

and so

$$b^T M^n N = b^T(-\beta_{n-1}M^{n-1} - \cdots)N = 0.$$

Similarly, $b^T M^{n+1} N = b^T(-\beta_{n-1}M^n - \cdots)N = 0$, etc. The claim (1.8) is proved.

Now notice that

$$b^T \mathbf{X}^{-1}(s)N = b^T e^{-sM}N = b^T \sum_{k=0}^{\infty} \frac{(-s)^k M^k N}{k!} = \sum_{k=0}^{\infty} \frac{(-s)^k}{k!} b^T M^k N = 0,$$

according to (1.8).

3. Assume next that $x^0 \in \mathcal{C}(t)$. This is equivalent to having

$$x^0 = -\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}(s)ds \text{ for some control } \boldsymbol{\alpha}(\cdot) \in \mathcal{A}.$$

Then

$$b \cdot x^0 = -\int_0^t b^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s) ds = 0.$$

This says that b is orthogonal x^0. In other words, $\mathcal{C}$ must lie in the hyperplane orthogonal to $b \neq 0$. Consequently $\mathcal{C}^0 = \emptyset$.

4. Conversely, assume $0 \notin \mathcal{C}^0$. Thus $0 \notin \mathcal{C}^0(t)$ for all $t > 0$. Since $\mathcal{C}(t)$ is convex, there exists a supporting hyperplane to $\mathcal{C}(t)$ through 0. This means that there exists $b \neq 0$ such that $b \cdot x^0 \leq 0$ for all $x^0 \in \mathcal{C}(t)$.

Choose any $x^0 \in \mathcal{C}(t)$. Then

$$x^0 = -\int_0^t \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s) ds$$

for some control $\boldsymbol{\alpha}$, and therefore

$$0 \geq b \cdot x^0 = -\int_0^t b^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s) ds.$$

Thus

$$\int_0^t b^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s) ds \geq 0 \quad \text{for all control } \boldsymbol{\alpha}(\cdot).$$

We assert that therefore

$$b^T \mathbf{X}^{-1}(s) N \equiv 0, \tag{1.9}$$

a proof of which follows as a lemma below. We rewrite (1.9) as

$$b^T e^{-sM} N \equiv 0. \tag{1.10}$$

Let $s = 0$ to see that $b^T N = 0$. Next differentiate (1.10) with respect to s, to find that

$$b^T (-M) e^{-sM} N \equiv 0.$$

For $s = 0$ this says

$$b^T M N = 0.$$

We repeatedly differentiate, to deduce

$$b^T M^k N = 0 \quad \text{for all} \quad k = 0,1, \ldots,$$

and so $b^T G = 0$. This implies rank $G < n$, since $b \neq 0$.

Lemma (1.2.11): (Integral Inequalities)

Assume that

$$\int_0^t b^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s) ds \geq 0 \tag{1.11}$$

for all $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$. Then

$$b^T \mathbf{X}^{-1}(s) N \equiv 0.$$

Proof:

Replacing α by α in (1.11), we see that

$$x^0 = -\int_0^t b^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s) ds = 0$$

for all $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$. Define

$$\mathbf{v}(s) := b^T \mathbf{X}^{-1}(s) N.$$

If $\mathbf{v} \neq 0$, then $\mathbf{v}(s_0) \neq 0$ for some s_0. Then there exists an interval I such that $s_0 \in I$ and $\mathbf{v} \neq 0$ on I. Now define $\alpha(\cdot) \in \mathcal{A}$ this way:

$$\begin{cases} \boldsymbol{\alpha}(s) = 0 & (s \notin I) \\ \boldsymbol{\alpha}(s) = \dfrac{\mathbf{v}(s)}{|\mathbf{v}(s)|} \dfrac{1}{\sqrt{n}} & (s \in I), \end{cases}$$

where $|v| := \left(\sum_{i=1}^n |v_i|^2\right)^{\frac{1}{2}}$. Then

$$0 = \int_0^t \mathbf{v}(s) \cdot \boldsymbol{\alpha}(s) ds = \int_I \frac{\mathbf{v}(s)}{\sqrt{n}} \cdot \frac{\mathbf{v}(s)}{|\mathbf{v}(s)|} ds = \frac{1}{\sqrt{n}} \int_I |\mathbf{v}(s)| ds$$

This implies the contradiction that $\mathbf{v} \equiv 0$ in I.

Definition (1.2.12):

We say the linear system (ODE) is controllable if $\mathcal{C} = \mathbb{R}^n$.

Theorem (1.2.13): (criterion for controllability)

Let A be the cube $[-1,1]^n$ in $\mathbb{R}^n$. Suppose as well that rank $G = n$, and Re $\lambda < 0$ for each eigenvalue λ of the matrix M.

Then the system (ODE) is controllable.

Proof:

Since rank $G = n$, Theorem (1.2.9) tells us that $\mathcal{C}$ contains some ball B centered at 0. Now take any $x^0 \in \mathbb{R}^n$ and consider the evolution

$$\begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) \\ \mathbf{x}(0) = x^0; \end{cases}$$

in other words, take the control $\boldsymbol{\alpha}(\cdot) \equiv 0$. Since Re $\lambda < 0$ for each eigenvalue λ of M, then the origin is asymptotically stable. So there exists a time T such that $\mathbf{x}(t) \in B$. Thus $\mathbf{x}(t) \in B \subset \mathcal{C}$; and hence there exists a control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$ steering $\mathbf{x}(T)$ into 0 in finite time.

Example (1.2.14):

We once again consider the rocket railroad car, from section (1.2), for which $n = 2, m = 1, A = [-1, 1]$, and

$$\dot{\mathbf{x}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \alpha.$$

Then

$$G = [N, MN] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Therefore

$$\operatorname{rank} G = 2 = n.$$

Also, the characteristic polynomial of the matrix M is

$$p(\lambda) = \det(\lambda I - M) = \det \begin{pmatrix} \lambda & -1 \\ 0 & \lambda \end{pmatrix} = \lambda^2.$$

Since the eigenvalues are both 0, we fail to satisfy the hypotheses of Theorem (1.2.13).

This example motivates the following extension of the previous theorem:

Theorem (1.2.15): (Improved Criterion for Controllability)

Assume $\operatorname{rank} G = n$ and $\operatorname{Re} \lambda \leq 0$ for each eigenvalue λ of M. Then the system (ODE) is controllable.

Proof:

1. If $\mathcal{C} \neq \mathbb{R}^n$, then the convexity of $\mathcal{C}$ implies that there exist a vector $b \neq= 0$ and a real number μ such that

$$b \cdot x^0 \leq \mu \tag{1.12}$$

for all $x^0 \in C$. Indeed, in the picture we see that $b \cdot (x^0 - z^0) \leq 0$; and this implies (1.12) for $\mu \coloneqq b \cdot z^0$.

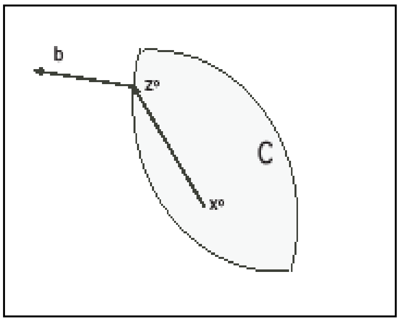

Figure (1.7)

We will derive a contradiction.

2. Given $b \neq 0$, our intention is to find $x^0 \in \mathcal{C}$ so that (1.12) fails. Recall $x^0 \in \mathcal{C}$ if and only if there exist a time $t > 0$ and a control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$ such that

$$x^0 = -\int_0^t \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s) ds.$$

Then

$$b \cdot x^0 = -\int_0^t b^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s) ds$$

Define

$$\mathbf{v}(s) \coloneqq b^T \mathbf{X}^{-1}(s) N$$

3. We assert that

$$\mathbf{v} \neq 0. \tag{1.13}$$

To see this, suppose instead that $\mathbf{v} \equiv 0$. Then k times differentiate the expression $b^T\mathbf{X}^{-1}(s)N$ with respect to s and set $s = 0$, to discover

$$b^T M^k N = 0$$

for $= 0,1,2,\dots$. This implies b is orthogonal to the columns of G, and so $\operatorname{rank} G < n$. This is a contradiction to our hypothesis, and therefore (1.13) holds.

4. Next, define $\boldsymbol{\alpha}(\cdot)$ this way:

$$\boldsymbol{\alpha}(s) := \begin{cases} -\dfrac{\mathbf{v}(s)}{|\mathbf{v}(s)|} & \text{if } \mathbf{v}(s) \neq 0 \\ 0 & \text{if } \mathbf{v}(s) = 0. \end{cases}$$

Then

$$b \cdot x^0 = -\int_0^t \mathbf{v}(s)\boldsymbol{\alpha}(s)ds = \int_0^t |\mathbf{v}(s)|ds$$

We want to find a time $t > 0$ so that $\int_0^t |\mathbf{v}(s)|ds > \mu$. In fact, we assert that

$$\int_0^\infty |\mathbf{v}(s)|ds = +\infty. \tag{1.14}$$

To begin the proof of (1.14), introduce the function

$$\phi(t) := \int_t^\infty \mathbf{v}(s)ds.$$

We will find an ODE ϕ satisfies. Take $p(\cdot)$ to be the characteristic polynomial of M. Then

$$\begin{aligned} p\left(-\frac{d}{dt}\right)\mathbf{v}(t) = p\left(-\frac{d}{dt}\right)[b^T e^{-tM}N] = b^T\left(p\left(-\frac{d}{dt}\right)e^{-tM}\right)N \\ = b^T(p(M)e^{-tM})N \equiv 0, \end{aligned}$$

since $p(M) = 0$, according to the Cayley-Hamilton Theorem. But since $p\left(-\frac{d}{dt}\right)\mathbf{v}(t) \equiv 0$, it follows that

$$-\frac{d}{dt}p\left(-\frac{d}{dt}\right)\phi(t) = p\left(-\frac{d}{dt}\right)\left(-\frac{d}{dt}\phi\right) = p\left(-\frac{d}{dt}\right)\mathbf{v}(t) = 0.$$

Hence - solves the $(n+1)^{th}$ order ODE

$$\frac{d}{dt}p\left(-\frac{d}{dt}\right)\phi(t) = 0.$$

We also know $\phi(\cdot) \not\equiv 0$. Let $\mu_1, \dots, \mu_{n+1}$ be the solutions of $\mu p(-\mu) = 0$. According to ODE theory, we can write

$$\phi(t) = \text{sum of terms of the form } p_i(t)e^{\mu_i t}$$

for appropriate polynomials $p_i(\cdot)$.

Furthermore, we see that $\mu_{n+1} = 0$ and $\mu_k = -\lambda_k$, where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of M. By assumption Re $\mu_k \geq 0$, for $k = 1, \dots, n$. If $\int_0^\infty |\mathbf{v}(s)| ds < \infty$, then

$$|\phi(t)| \leq \int_0^\infty |\mathbf{v}(s)| ds \to 0 \quad \text{as } t \to \infty;$$

that is, $\phi(t) \to 0$ as $t \to \infty$. This is a contradiction to the representation formula of $\phi(t) = \sum p_i(t)\, e^{\mu_i t}$, with Re $\mu_i \geq 0$. Assertion (2.10) is proved.

5. Consequently given any μ, there exists $t > 0$ such that

$$b \cdot x^0 = \int_0^\infty |\mathbf{v}(s)| ds > \mu,$$

a contradiction to (1.12). Therefore $\mathcal{C} = \mathbb{R}^n$.

(1.2.4): Observability:

We again consider the linear system of ODE

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) & (t > 0) \\ \mathbf{x}(0) = x^0 \end{cases}$$

where $M \in \mathbb{M}^{n\times n}$.

In this subsection we address the observability problem, modeled as follows: We suppose that we can observe

$$\text{(O)} \qquad \mathbf{y}(t) := N\mathbf{x}(t) \qquad (t \geq 0),$$

for a given matrix $N \in \mathbb{M}^{m\times n}$. Consequently, $\mathbf{y}(t) \in \mathbb{R}^n$. The interesting situation is when $m << n$ and we interpret $\mathbf{y}(\cdot)$ as low-dimensional "observations" or "measurements" of the high-dimensional dynamics $\mathbf{x}(\cdot)$.

(1.2.5): OBSERVABILITY QUESTION:

Given the observations $\mathbf{y}(\cdot)$, can we in principle reconstruct $\mathbf{x}(\cdot)$? In particular, do observations of $\mathbf{y}(\cdot)$ provide enough information for us to deduce the initial value x^0 for (ODE)?

Definition (1.2.16):

The pair (ODE), (O) is called observable if the knowledge of $\mathbf{y}(\cdot)$ on any time interval $[0, t]$ allows us to compute x^0.

More precisely, (ODE), (O) is observable if for all solutions $\mathbf{x}_1(\cdot), \mathbf{x}_2(\cdot), N\mathbf{x}_1(\cdot) \equiv N\mathbf{x}_2(\cdot)$ on a time interval $[0, t]$ implies $\mathbf{x}_1(0) = \mathbf{x}_2(0)$.

Example (1.2.17):

(i) If $N \equiv 0$, then clearly the system is not observable.
(ii) On the other hand, if $m = n$ and N is invertible, then clearly $\mathbf{x}(t) = N^{-1}\mathbf{y}(t)$ is observable.

The interesting cases lie between these extremes.

Theorem (1.2.18): (Observability and Controllability)

The system

$$\begin{cases}\dot{\mathbf{x}}(t) = M\mathbf{x}(t) \\ \mathbf{y}(t) = N\mathbf{x}(t)\end{cases} \tag{1.15}$$

is observable if and only if the system

$$\dot{\mathbf{z}}(t) = M^T\mathbf{z}(t) + N^T\boldsymbol{\alpha}(t), \quad A = \mathbb{R}^m \tag{1.16}$$

is controllable, meaning that $\mathcal{C} = \mathbb{R}^m$.

INTERPRETATION

This theorem asserts that somehow "observability and controllability are dual concepts" for linear systems.

Proof:

Suppose (1.15) is not observable. Then there exist points $x^1 \neq x^2 \in \mathbb{R}^n$, such that

$$\begin{cases} \dot{\mathbf{x}}_1(t) = M\mathbf{x}_1(t), & \mathbf{x}_1(0) = x^1 \\ \dot{\mathbf{x}}_2(t) = M\mathbf{x}_2(t), & \mathbf{x}_2(0) = x^2 \end{cases}$$

but $\mathbf{y}(t) \coloneqq N\mathbf{x}_1(t) \equiv N\mathbf{x}_2(t)$ for all times $t \geq 0$. Let

$$\mathbf{x}(t) \coloneqq \mathbf{x}_1(t) - \mathbf{x}_2(t), \qquad x^0 \coloneqq x^1 - x^2.$$

Then

$$\dot{\mathbf{x}}t = M\mathbf{x}(t), \qquad \mathbf{x}(0) = x^0 \neq 0,$$

but

$$Nx(t) = 0 \qquad (t \geq 0).$$

Now

$$\mathbf{x}(t) \;=\; \mathbf{X}(t)x^0 \;=\; e^{tM}x^0.$$

Thus

$$Ne^{tM}x^0 = 0 \qquad (t \geq 0).$$

Let $t = 0$, to find $Nx^0 = 0$. Then differentiate this expression k times in t and let $t =$ 0, to discover as well that

$$NM^k x^0 = 0$$

for $= 0,1,2,\ldots$. Hence $(x^0)^T(M^k)^T N^T = 0$, and hence $(x^0)^T(M^T)^k N^T = 0$. This implies

$$(x^0)^T[N^T, M^T N^T, \ldots, (M^T)^{n-1}N^T\,] = 0.$$

Since $x^0 \neq 0$, $\operatorname{rank}[N^T, \ldots, (M^T)^{n-1}N^T\,] < n$. Thus problem (1.16) is not controllable. Consequently, (1.16) controllable implies (1.15) is observable.

2. Assume now (1.16) not controllable. Then $\operatorname{rank}[N^T, \ldots, (M^T)^{n-1}N^T\,] < n$, and consequently according to Theorem (1.2.9) there exists $x^0 \neq 0$ such that

$$(x^0)^T\,[N^T, \ldots, (M^T)^{n-1}N^T\,] \;=\; 0.$$

That is, $NM^k x^0 = 0$ for all $k = 0,1,2,\dots,n-1$.

We want to show that $\mathbf{y}(t) = N\mathbf{x}(t)0$, where

$$\begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) \\ \mathbf{x}(0) = x^0. \end{cases}$$

According to the Cayley-Hamilton Theorem, we can write

$$M^n = -\beta_{n-1}M^{n-1} - \cdots - \beta_0 I.$$

for appropriate constants. Consequently $NM^n x^0 = 0$. Likewise,

$$M^{n+1} = M(-\beta_{n-1}M^{n-1} - \cdots - \beta_0 I) = -\beta_{n-1}M^n - \cdots - \beta_0 M;$$

and so $NM^{n+1}x^0 = 0$. Similarly, $NM^k x^0 = 0$ for all k.

Now

$$\mathbf{x}(t) = \mathbf{X}(t)x^0 = e^{Mt}x^0 = \sum_{k=0}^{\infty} \frac{t^k M^k}{k!} x^0\,;$$

and therefore $N\mathbf{x}(t) = N\sum_{k=0}^{\infty}\frac{t^k M^k}{k!}x^0 = 0$.

We have shown that if (1.16) is not controllable, then (1.15) is not observable.

(1.2.6): Bang-Bang Principle

For this section, we will again take A to be the cube $[-1,1]^m$ in $\mathbb{R}^m$.

Definition (1.2.19):

A control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$ is called bang-bang if for each time $t \geq 0$ and each index $i = 1,\dots,m$, we have $|\alpha^i(t)| = 1$, where

$$\boldsymbol{\alpha}(t) = \begin{pmatrix} \alpha^1(t) \\ \vdots \\ \alpha^m(t) \end{pmatrix}.$$

Theorem (1.2.20): (Bang-Bang Principle)

Let $t > 0$ and suppose $x^0 \in \mathcal{C}(t)$, for the system

$$\dot{\mathbf{x}}(t) = M\mathbf{x}(t) + N\boldsymbol{\alpha}(t)$$

Then there exists a bang-bang control $\boldsymbol{\alpha}(\cdot)$ which steers x^0 to 0 at time t.

To prove the theorem we need some tools from functional analysis, among them the Krein-Milman Theorem, expressing the geometric fact that every bounded convex set has an extreme point.

(1.2.7): Some functional analysis

We will study the "geometry" of certain infinite dimensional spaces of functions.

To establish our discussing we have to know that:

$$L^\infty = L^\infty(0,t;\mathbb{R}^m) = \left\{\boldsymbol{\alpha}(\cdot):(0,t)\to\mathbb{R}^m \middle| \sup_{0\le s\le t}|\boldsymbol{\alpha}(s)|<\infty\right\}.$$

$$\|\boldsymbol{\alpha}\|_{L^\infty} = \sup_{0\le s\le t}|\boldsymbol{\alpha}(t)|.$$

Definition (1.2.21):

Let $\alpha_n \in L^\infty$ for $n = 1, \ldots$ and $\alpha \in L^\infty$. We say $\boldsymbol{\alpha}_n$ converges to α in the weak* sense, written

$$\alpha_n \overset{*}{\longrightarrow} \alpha,$$

provided

$$\int_0^\infty \boldsymbol{\alpha}_n(s)\cdot\mathbf{v}(s)ds\, t\int_0^\infty \boldsymbol{\alpha}(s)\cdot\mathbf{v}(s)ds$$

as $n\to\infty$, for all $\mathbf{v}(\cdot):[0,t]\to\mathbb{R}^m$ satisfying $\int_0^t|\mathbf{v}(s)|ds<\infty$.

We will need the following useful weak* compactness theorem for L^∞:

Theorem (1.2.22): (Alaoglu's Theorem)

Let $\boldsymbol{\alpha}(s) \in \mathcal{A}, n = 1, \ldots$. Then there exists a subsequence $\boldsymbol{\alpha}_{n_k}$ and $\in \boldsymbol{\mathcal{A}}$, such that

$$\alpha_{n_k} \overset{*}{\longrightarrow} \alpha.$$

Definitions (1.2.23):

(i) The set $\mathbb{K}$ is convex if for all $x, \hat{x} \in \mathbb{K}$ and all real numbers $0 \le \lambda \le 1$,

$$\lambda x + (1-\lambda)\hat{x} \in \mathbb{K}.$$

(ii) A point $z \in \mathbb{K}$ is called extreme provided there do not exist points $x, \hat{x} \in \mathbb{K}$ and $0 < \lambda < 1$ such that

$$z = \lambda x + (1 - \lambda)\hat{x}.$$

Theorem (1.2.24): (Krein-Milman Theorem)

Let $\mathbb{K}$ be a convex, nonempty subset of L^{∞}, which is compact in the weak * topology.

Then $\mathbb{K}$ has at least one extreme point.

(1.2.8): Application to Bang-Bang Controls:

The foregoing abstract theory will be useful for us in the following setting. We will take $\mathbb{K}$ to be the set of controls which steer x^0 to 0 at time t, prove it satisfies the hypotheses of Krein-Milman Theorem and finally show that an extreme point is a bang-bang control.

So consider again the linear dynamics

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) + N\boldsymbol{\alpha}(t) \\ \mathbf{x}(0) = x^0. \end{cases}$$

Take $x^0 \in \mathcal{C}(t)$ and write

$$\mathbb{K} = \{\boldsymbol{\alpha}(\cdot) \in \mathcal{A} | \boldsymbol{\alpha}(\cdot) \text{ steers } x^0 \text{ to } 0 \text{ at time } t\}.$$

Lemma (1.2.25): (Geometry of set of controls)

The collection $\mathbb{K}$ of admissible controls satisfies the hypotheses of the Krein-Milman Theorem.

Proof:

Since $x^0 \in \mathcal{C}(t)$, we see that $\mathbb{K} \neq \emptyset$.

Next we show that $\mathbb{K}$ is convex. For this, recall that $\boldsymbol{\alpha}(\cdot) \in \mathbb{K}$if and only if

$$x^0 = -\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}(s)ds.$$

Now take also $\hat{\boldsymbol{\alpha}} \in \mathbb{K}$ and $0 \leq \lambda \leq 1$. Then

$$x^0 = -\int_0^t \mathbf{X}^{-1}(s)N\hat{\boldsymbol{\alpha}}(s)ds\,;$$

and so

$$x^0 = -\int_0^t \mathbf{X}^{-1}(s)N\big(\lambda\boldsymbol{\alpha}(s) + (1-\lambda)\hat{\boldsymbol{\alpha}}(s)\big)ds$$

Hence $\lambda\boldsymbol{\alpha} + (1-\lambda)\hat{\boldsymbol{\alpha}} \in \mathbb{K}$.

Lastly, we confirm the compactness. Let $\boldsymbol{\alpha}_n \in \mathbb{K}$ for $n = 1, \dots$. According to Alaoglu fs Theorem there exist $n_k \to \infty$ and $\boldsymbol{\alpha} \in \mathcal{A}$ such that $\alpha_{n_k} \xrightarrow{*} \alpha$. We need to show that $\boldsymbol{\alpha} \in \mathbb{K}$.

Now $\boldsymbol{\alpha}_{n_k} \in \mathbb{K}$ implies

$$x^0 = -\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}_{n_k}(s)ds \to -\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}(s)ds$$

by definition of weak-* convergence. Hence $\boldsymbol{\alpha} \in \mathbb{K}$.

We can now apply the Krein-Milman Theorem to deduce that there exists an extreme point $\boldsymbol{\alpha}^* \in \mathbb{K}$. What is interesting is that such an extreme point corresponds to a bang-bang control.

Theorem (1.2.26): (Extremality and Bang-Bang Principle)

The control $\boldsymbol{\alpha}^*(\cdot)$ is bang-bang.

Proof:

1. We must show that for almost all times $0 \le s \le t$ and for each $i = 1, \dots, m$, we have

$$\left|\boldsymbol{\alpha}^{i*}(s)\right| = 1.$$

Suppose not. Then there exists an index $i \in \{1, \dots, m\}$ and a subset $E \subset [0, t]$ of positive measure such that $\left|\alpha^{i*}(s)\right| < 1$ for $s \in E$. In fact, there exist a number $\varepsilon > 0$ and a subset $F \subseteq E$ such that

$$|F| > 0 \;\text{ and }\; \left|\boldsymbol{\alpha}^{i*}(s)\right| \le 1 - \varepsilon \;\text{ for }\; s \in F.$$

Define

$$I_F(\beta(\cdot)) := \int_F \mathbf{X}^{-1}(s)N\boldsymbol{\beta}(s)ds,$$

for

$$\boldsymbol{\beta}(\cdot) := (0, \dots, \beta(\cdot), \dots, 0)^T,$$

the function β in the ith slot. Choose any real-valued function $\beta(\cdot) \not\equiv 0$, such that

$$I_F(\beta(\cdot)) = 0$$

and $|\beta(\cdot)| \le 1$. Define

$$\begin{aligned} \boldsymbol{\alpha}_1(\cdot) &:= \boldsymbol{\alpha}^*(\cdot) + \varepsilon\beta(\cdot) \\ \boldsymbol{\alpha}_2(\cdot) &:= \boldsymbol{\alpha}^*(\cdot) + \varepsilon\beta(\cdot), \end{aligned}$$

where we redefine β to be zero off the set F

2. We claim that

$$\boldsymbol{\alpha}_1(\cdot), \boldsymbol{\alpha}_2(\cdot) \in \mathbb{K}.$$

To see this, observe that

$$\begin{aligned} -\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}_1(s)ds &= -\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}^*(s)ds - \varepsilon\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\beta}(s)ds \\ &= x^0 - \varepsilon \underbrace{\int_F \mathbf{X}^{-1}(s)N\boldsymbol{\beta}(s)ds}_{I_F(\beta(\cdot))=0} = x^0. \end{aligned}$$

Note also $\boldsymbol{\alpha}_1(\cdot) := \boldsymbol{\alpha}^* \in \mathcal{A}$. Indeed,

$$\begin{cases} \boldsymbol{\alpha}_1(s) = \boldsymbol{\alpha}^*(s) & (s \notin F) \\ \boldsymbol{\alpha}_1(s) = \boldsymbol{\alpha}^*(s) + \varepsilon\beta(s) & (s \in F). \end{cases}$$

But on the set F, we have $|\boldsymbol{\alpha}_i^*(s)| \le 1 - \varepsilon$, and therefore

$$|\boldsymbol{\alpha}_1(s)| \le |\boldsymbol{\alpha}^*(s)| + \varepsilon|\beta(s)| \le 1 - \varepsilon + \varepsilon = 1.$$

Similar considerations apply for $\boldsymbol{\alpha}_2$. Hence $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \mathbb{K}$**,** as claimed above.

3. Finally, observe that

$$\begin{cases} \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}^* + \varepsilon\beta, & \boldsymbol{\alpha}_1 \neq \boldsymbol{\alpha}^* \\ \boldsymbol{\alpha}_2 = \boldsymbol{\alpha}^* - \varepsilon\beta, & \boldsymbol{\alpha}_2 \neq \boldsymbol{\alpha}^*. \end{cases}$$

But

$$\frac{1}{2}\boldsymbol{\alpha}_1 + \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{\alpha}_2 = \boldsymbol{\alpha}^*;$$

and this is a contradiction, since $\boldsymbol{\alpha}^*$ is an extreme point of $\mathbb{K}$.

Chapter (2)

Optimal Control and Pontryagin Maximum

Section (2.1): Linear time-optimal Control

To discuss the existence of time-optimal controls, consider the linear system of ODE:

$$(\mathrm{ODE}) \qquad \begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) + N\boldsymbol{\alpha}(t) \\ \mathbf{x}(0) = x^0, \end{cases}$$

for given matrices $M \in \mathbb{M}^{n\times n}$ and $N \in \mathbb{M}^{n\times n}$. We will again take A to be the cube $[-1,1]^m \subset \mathbb{R}^m$.

Define next

$$(\mathrm{P}) \qquad \boldsymbol{P}[\boldsymbol{\alpha}(\cdot)] := -\int_0^\tau 1\, ds = -\tau,$$

where $\tau = \tau(\boldsymbol{\alpha}(\cdot))$ denotes the first time the solution of our ODE (2.1) hits the origin 0. (If the trajectory never hits 0, we set $\tau = \infty$).

And to study the optimal time problem, we are given the starting point $x^0 \in \mathbb{R}^n$, and want to find an optimal control $\boldsymbol{\alpha}^*(\cdot)$ such that

$$\boldsymbol{P}[\boldsymbol{\alpha}(\cdot)] = \max_{\boldsymbol{\alpha}(\cdot)\in\mathcal{A}} \boldsymbol{P}[\boldsymbol{\alpha}(\cdot)].$$

Then

$$\tau^* = -\mathcal{P}[\boldsymbol{\alpha}^*(\cdot)] \quad \text{is the minimum time to steer to the origin.}$$

Theorem (2.1.1): (Existence of time-optimal Control)

Let $x^0 \in \mathbb{R}^n$. Then there exists an optimal bang-bang control $\boldsymbol{\alpha}^*(\cdot)$.

Proof:

Let $\tau^* := \inf\{t \mid x^0 \in C(t)\}$. We want to show that $x^0 \in C(\tau^*)$; that is, there exists an optimal control $\boldsymbol{\alpha}^*(\cdot)$ steering x^0 to 0 at time τ^*.

Choose $t_1 \geq t_2 \geq t_3 \geq \cdots$ so that $x^0 \in C(t_n)$ and $t_n \to \tau^*$. Since $x^0 \in C(t_n)$, there exists a control $\boldsymbol{\alpha}_n(\cdot) \in \mathcal{A}$ such that

$$x^0 = -\int_0^{t_n} \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}_n(s)\, ds.$$

If necessary, redefine $\boldsymbol{\alpha}_n(s)$ to be 0 for $t_n \leq s$. By Alaoglu's Theorem, there exists a subsequence $n_k \to \infty$ and a control $\boldsymbol{\alpha}^*(\cdot)$ so that

$$\boldsymbol{\alpha}_n \overset{*}{\to} \boldsymbol{\alpha}^*.$$

We assert that $\boldsymbol{\alpha}^*(\cdot)$ is an optimal control. It is easy to check that $\boldsymbol{\alpha}^*(s) = 0, s \geq \tau^*$. Also

$$x^0 = -\int_0^{t_{n_k}} \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}_{n_k}(s)\, ds = -\int_0^{t_1} \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}_{n_k}(s)\, ds,$$

since $\boldsymbol{\alpha}_{n_k} = 0$ for $s \geq t_{n_k}$. Let $n_k \to \infty$:

$$x^0 = -\int_0^{t_1} \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}^*(s)\, ds = \int_0^{\tau^*} \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}^*(s)\, ds$$

because $\boldsymbol{\alpha}^*(s) = 0$ for $s \geq \tau^*$. Hence $x^0 \in C(\tau^*)$, and therefore $\boldsymbol{\alpha}^*(\cdot)$ is optimal.

According to Theorem (1.2.26) there in fact exists an optimal bang-bang control.

The really interesting practical issue now is understanding how to compute an optimal control $\boldsymbol{\alpha}^*(\cdot)$.

Definition (2.1.2):

We define $K(t, x^0)$ to be the reachable set for time t. That is,

$K(t, x^0) = \{x^1 \mid \text{there exists } \boldsymbol{\alpha}(\cdot) \in \mathcal{A} \text{ which steers from } x^0 \text{ to } x^1 \text{ at time } t\}$.

Since $\mathbf{x}(\cdot)$ solves (ODE), we have $x^1 \in K(t, x^0)$ if and only if

$$x^1 = \mathbf{X}(t) x^0 + \mathbf{X}(t) \int_0^t \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s)\, ds = \mathbf{x}(t)$$

for some control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$.

Theorem (2.1.3): (Geometry of the Set K)

The set $K(t, x^0)$ is convex and closed.

Proof:

1. (Convexity) Let $x^1, x^2 \in K(t, x^0)$. Then there exists $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \mathcal{A}$ such that

$$x^1 = \mathbf{X}(t)x^0 + \mathbf{X}(t)\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}_1(s)\,ds$$

$$x^2 = \mathbf{X}(t)x^0 + \mathbf{X}(t)\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}_2(s)\,ds.$$

Let $0 \leq \lambda \leq 1$. Then

$$\lambda x^1 + (1-\lambda)x^2 = \mathbf{X}(t)x^0 + \mathbf{X}(t)\int_0^t \mathbf{X}^{-1}(s)N\underbrace{\left(\boldsymbol{\alpha}_1(s) + (1-\lambda)\boldsymbol{\alpha}_2(s)\right)}_{\in\mathcal{A}}\,ds,$$

and hence $\lambda x^1 + (1-\lambda)x^2 \in K(t, x^0)$.

2. (Closedness) Assume $x^k \in K(t, x^0)$ for $(k = 1,2,...)$ and $x^k \to y$. We must show $y \in K(t, x^0)$. As $x^k \in K(t, x^0)$, there exists $\boldsymbol{\alpha}_k(\cdot) \in \mathcal{A}$ such that

$$x^k = \mathbf{X}(t)x^0 + \mathbf{X}(t)\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}_k(s)\,ds.$$

According to Alaoglu's Theorem, there exist a subsequence $k_j \to \infty$ and $\boldsymbol{\alpha} \in \mathcal{A}$ such that $\boldsymbol{\alpha}_k \overset{*}{\to} \boldsymbol{\alpha}$. Let $k = k_j \to \infty$ in the expression above, to find

$$y = \mathbf{X}(t)x^0 + \mathbf{X}(t)\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}(s)\,ds.$$

Thus $y \in K(t, x^0)$, and hence $K(t, x^0)$ is closed.

If S is a set, we write ∂S to denote the boundary of S.

Recall that τ^*. denotes the minimum time it takes to steer to 0, using the optimal control $\boldsymbol{\alpha}^*$. Note that then $0 \in \partial K(\tau^*, x^0)$.

Theorem (2.1.4): (Pontryagin Maximum Principle for Linear time-optimal Control)

There exists a nonzero vector h such that

(M) $$h^T\mathbf{X}^{-1}(t)N\boldsymbol{\alpha}^*(t) = \max_{a\in A}\{h^T\mathbf{X}^{-1}(t)Na\}$$

for each time $0 \leq t \leq \tau^*$.

The significance of this assertion is that if we know h then the maximization principle (M) provides us with a formula for computing $\boldsymbol{\alpha}^*(\cdot)$, or at least extracting useful information.

We will see in the next section that assertion (M) is a special case of the general Pontryagin Maximum Principle.

Proof:

1. We know $0 \in \partial K(\tau^*, x^0)$. Since $K(\tau^*, x^0)$ is convex, there exists a supporting plane to $K(\tau^*, x^0)$ at 0; this means that for some $g \neq 0$, we have

$$g \cdot x_1 \leq 0 \quad \text{for all} \quad x_1 \in K(\tau^*, x^0).$$

2. Now $x^1 \in K(\tau^*, x^0)$ if and only if there exists $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$ such that

$$x^1 = \mathbf{X}(\tau^*)x^0 + \mathbf{X}(\tau^*) \int_0^{\tau^*} \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s)\, ds.$$

Also

$$0 = \mathbf{X}(\tau^*)x^0 + \mathbf{X}(\tau^*) \int_0^{\tau^*} \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}^*(s)\, ds.$$

Since $g \cdot x^1 \leq 0$, we deduce that

$$g^T \left(\mathbf{X}(\tau^*)x^0 + \mathbf{X}(\tau^*) \int_0^{\tau^*} \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s)\, ds \right) \leq 0$$

$$= g^T \left(\mathbf{X}(\tau^*)x^0 + \mathbf{X}(\tau^*) \int_0^{\tau^*} \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}^*(s)\, ds \right).$$

Define $h^T := g^T \mathbf{X}(\tau^*)$. Then

$$\int_0^{\tau^*} h^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s)\, ds \leq \int_0^{\tau^*} h^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}^*(s)\, ds;$$

and therefore

$$\int_0^{\tau^*} h^T \mathbf{X}^{-1}(s) N\big(\boldsymbol{\alpha}^*(s) - \boldsymbol{\alpha}(s)\big)\, ds \geq 0$$

for all controls $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$.

3. We claim now that the foregoing implies

$$h^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}^*(s) = \max_{a \in A}\{h^T \mathbf{X}^{-1}(s) N a\}$$

for almost every time s.

For suppose not; then there would exist a subset $E \subset [0, \tau^*]$ of positive measure, such that

$$h^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}^*(s) < \max_{a \in A}\{h^T \mathbf{X}^{-1}(s) N a\}$$

for $s \in E$. Design a new control $\widehat{\boldsymbol{\alpha}}(\cdot)$ as follows:

$$\boldsymbol{\alpha}^*(s) = \begin{cases} \boldsymbol{\alpha}^*(s) & (s \notin E) \\ \boldsymbol{\alpha}(s) & (s \in E) \end{cases}$$

where $\boldsymbol{\alpha}(s)$ is selected so that

$$\max_{a \in A}\{h^T \mathbf{X}^{-1}(s) N a\} = h^T \mathbf{X}^{-1}(s) N \boldsymbol{\alpha}(s).$$

Then

$$\int_E \underbrace{h^T \mathbf{X}^{-1}(s) N\big(\boldsymbol{\alpha}^*(s) - \widehat{\boldsymbol{\alpha}}(s)\big)}_{<0}\, ds \geq 0.$$

This contradicts Step 2 above.

For later reference, we pause here to rewrite the foregoing into different notation; this will turn out to be a special case of the general theory developed later in section (2.2). First of all, define the Hamiltonian

$$H(x, p, a) := (Mx + Na) \cdot p \quad (x, p \in \mathbb{R}^n, a \in A).$$

Theorem (2.1.5): (Another way to write Pontryagin Maximum Principle for time-optimal Control)

Let $\mathbf{x}^*(\cdot)$ be a time optimal control and $\mathbf{x}^*(\cdot)$ the corresponding response.

Then there exists a function $\mathbf{p}^*(\cdot) : [0, \tau^*] \to \mathbb{R}^n$, such that

(ODE) $$\dot{\mathbf{x}}^*(t) = \nabla_p H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big),$$

(ADJ) $$\dot{\mathbf{p}}^*(t) = -\nabla_x H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big),$$

and

(M) $$H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big) = \max_{a\in A} H\big(\mathbf{x}^*(t), \mathbf{p}^*(t)\big).$$

We call (ADJ) the adjoint equations and (M) the maximization principle. The function $\mathbf{p}^*(\cdot)$ is the costate.

Proof:

1. Select the vector h as in Theorem (2.1.4), and consider the system

$$\begin{cases} \mathbf{p}^*(t) = -M^T\mathbf{p}^*(t) \\ \mathbf{p}^*(0) = h. \end{cases}$$

The solution is $\mathbf{p}^*(t) = e^{-tM^T}h$; and hence

$$\mathbf{p}^*(t)^T = h^T\mathbf{X}^{-1}(t),$$

since $\left(e^{-tM^T}\right)^T = e^{-tM} = \mathbf{X}^{-1}(t)$.

2. We know from condition (M) in Theorem (2.1.4) that

$$h^T\mathbf{X}^{-1}(t)N\boldsymbol{\alpha}^*(t) = \max_{a\in A}\{h^T\mathbf{X}^{-1}(t)Na\}$$

Since $\mathbf{p}^*(t)^T = h^T\mathbf{X}^{-1}(t)$, this means that

$$\mathbf{p}^*(t)^T\big(M\mathbf{x}^*(t) + N\boldsymbol{\alpha}^*(t)\big) = \max_{a\in A}\{\mathbf{p}^*(t)^T(M\mathbf{x}^*(t) + Na)\}.$$

3. Finally, we observe that according to the definition of the Hamiltonian H, the dynamical equations for $\mathbf{x}^*(\cdot), \mathbf{p}^*(\cdot)$ take the form (ODE) and (ADJ), as stated in the Theorem.

Example (2.1.6): (Rocket Railroad Car)

We recall this example, introduced in section (1.1). We have

(ODE) $$\dot{\mathbf{x}}(t) = \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{M} \mathbf{x}(t) + \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{N} \alpha(t)$$

for

$$\mathbf{x}(t) = \begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix}, \quad \mathrm{A} = [-1,1].$$

According to the Pontryagin Maximum Principle, there exists $h \neq 0$ such that

(M) $$h^T \mathbf{X}^{-1}(t) N \alpha^*(t) = \max_{|a| \le 1} \{h^T \mathbf{X}^{-1}(t) N a\}.$$

We will extract the interesting fact that an optimal control α^*. switches at most one time.

We must compute e^{tM}. To do so, we observe

$$M^0 = I, \quad M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad M^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 0;$$

and therefore $M^k = 0$ for all $k \ge 2$. Consequently,

$$e^{tM} = I + tM = \begin{pmatrix} 0 & t \\ 0 & 1 \end{pmatrix}.$$

Then

$$\mathbf{X}^{-1}(t) = \begin{pmatrix} 0 & -t \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{X}^{-1}(t) N = \begin{pmatrix} 0 & -t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -t \\ 1 \end{pmatrix}$$

$$h^T \mathbf{X}^{-1}(t) N = (h_1 \quad h_2) \begin{pmatrix} -t \\ 1 \end{pmatrix} = -th_1 + h_2.$$

The Maximum Principle asserts

$$(-th_1 + h_2)\alpha^*(t) = \max_{|a| \le 1} \{(-th_1 + h_2)a\};$$

and this implies that

$$\alpha^*(t) = \operatorname{sgn}(-th_1 + h_2)$$

for the sign function

$$\operatorname{sgn} x = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0. \end{cases}$$

Therefore the optimal control α^* switches at most once; and if $h_1 = 0$, then α^* is constant.

Since the optimal control switches at most once, then the control we constructed by a geometric method in section (1.1) must have been optimal.

Example (2.1.7): (Control of A vibrating Spring)

Consider next the simple dynamics

$\ddot{x} + x = \alpha,$

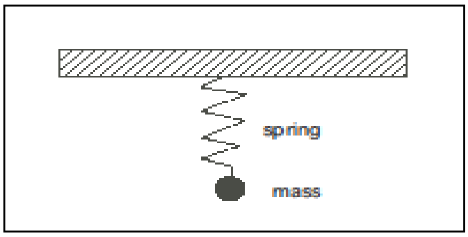

Figure (2.1)

where we interpret the control as an exterior force acting on an oscillating weight (of unit mass) hanging from a spring. Our goal is to design an optimal exterior forcing $\alpha^*(\cdot)$ that brings the motion to a stop in minimum time.

We have $n = 2, m = 1$. The individual dynamical equations read:

$$\begin{cases} \dot{x}^1(t) = x^2(t) \\ \dot{x}^2(t) = -x^1(t) + \alpha(t); \end{cases}$$

which in vector notation become

(ODE) $$\dot{\mathbf{x}}(t) = \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{M} \mathbf{x}(t) + \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{N} \alpha(t)$$

for $|\alpha(t)| \leq 1$. That is, $A = [-1,1]$.

We employ the Pontryagin Maximum Principle, which asserts that there exists $h \neq 0$ such that

(M) $$h^T \mathbf{X}^*(t) N \alpha^*(t) = \max_{a \in A} \{h^T \mathbf{X}^{-1}(t) N a\}.$$

To extract useful information from (M) we must compute $\mathbf{X}(\cdot)$. To do so, we observe that the matrix M is skew symmetric, and thus

$$M^0 = I, \quad M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad M^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I$$

Therefore

$$M^k = I \quad \text{if } k = 0,4,8, \dots$$

$$M^k = M \quad \text{if } k = 1,5,9, \dots$$

$$M^k = -I \quad \text{if } k = 2,6, \dots$$

$$M^k = -M \quad \text{if } k = 3,7, \dots;$$

and consequently

$$\begin{aligned} e^{tM} &= I + tM + \frac{t^2}{2!}M^2 + \cdots \\ &= I + tM - \frac{t^2}{2!}I - \frac{t^3}{3!}M + \frac{t^4}{4!}I + \cdots \\ &= \left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \cdots\right)I + \left(t - \frac{t^3}{3!} + \frac{t^5}{5!} - \cdots\right)M \\ &= \cos t\, I + \sin t\, M = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}. \end{aligned}$$

So we have

$$\mathbf{X}^{-1}(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

and

$$\mathbf{X}^{-1}(t)N = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix};$$

whence

$$h^T \mathbf{X}^{-1}(t)N = (h_1, h_2)\begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} = -h_1 \sin t + h_2 \cos t.$$

According to condition (M), for each time t we have

$$(-h_1 \sin t + h_2 \cos t)\alpha^*(t) = \max_{|a|\leq 1}\{(-h_1 \sin t + h_2 \cos t)a\}.$$

Therefore

$$\alpha^*(t) = \operatorname{sgn}(-h_1 \sin t + h_2 \cos t).$$

To simplify further, we may assume $h_1^2 + h_2^2 = 1$. Recall the trig identity $\sin(x+y) = \sin x \cos y + \cos x \sin y$, and choose δ such that $-h_1 = \cos\delta, h_2 = \sin\delta$. Then

$$\alpha^*(t) = \operatorname{sgn}(\cos\delta \sin t + \sin\delta \cos t) = \operatorname{sgn}(\sin(t+\delta)).$$

We deduce therefore that α^* switches from $+1$ to -1, and vice versa, every π units of time.

Next, we figure out the geometric consequences. When $\alpha \equiv 1$, our (ODE) becomes

$$\begin{cases} \dot{x}^1 = x^2 \\ \dot{x}^2 = -x^{-1} + 1. \end{cases}$$

In this case, we can calculate that

$$\frac{d}{dt}\left((x^1(t)-1)^2 + (x^2)^2(t)\right) = 2(x^1(t)-1)\dot{x}^1(t) + 2x^2(t)\dot{x}^2(t)$$

$$= 2(x^1(t)-1)x^2(t) + 2x^2(t)(-x^1(t)+1) = 0.$$

Consequently, the motion satisfies $(x^1(t)-1)^2 + (x^2)^2(t) \equiv r_1^2$, for some radius r_1, and therefore the trajectory lies on a circle with center $(0,1)$, as illustrated.

If $\alpha \equiv 1$, then (ODE) instead becomes

$$\begin{cases} \dot{x}^1 = x^2 \\ \dot{x}^2 = -x^{-1} - 1; \end{cases}$$

in which case

$$\frac{d}{dt}\left((x^1(t)+1)^2 + (x^2)^2(t)\right) = 0.$$

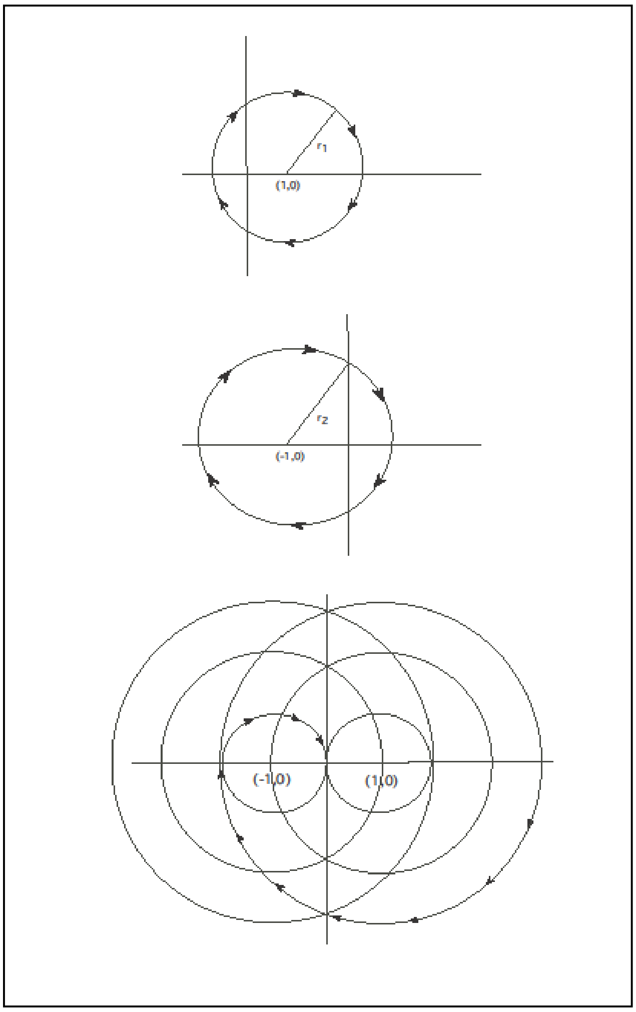

Figure (2.2)

Thus $(x^1(t)+1)^2+(x^2)^2(t)=r_2^2$ for some radius r_2, and the motion lies on a circle with center $(-1,0)$.

In summary, to get to the origin we must switch our control $\alpha(\cdot)$ back and forth between the values ± 1, causing the trajectory to switch between lying on circles centered at $(\pm 1,0)$. The switches occur each π units of time.

Section (2.2): The Pontryagin Maximum Principle

This important chapter moves us beyond the linear dynamics assumed in Chapter (1), to consider much wider classes of optimal control problems, to introduce the fundamental Pontryagin Maximum Principle, and to illustrate its uses in a variety of examples.

We begin in this section with a quick introduction to some variational methods. These ideas will later serve as motivation for the Pontryagin Maximum Principle.

Assume we are given a smooth function $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} = L(x, v)$; L is called the Lagrangian. Let $T > 0, x^0, x^1 \in \mathbb{R}^n$ be given.

We have to note that the basic problem of the calculus of variations is to find a curve $x^*(\cdot) : [0, T] \to \mathbb{R}^n$ that minimizes the functional

$$I[\mathbf{x}(\cdot)] := \int_0^T L\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big)\, dt \tag{2.1}$$

among all functions $\mathbf{x}(\cdot)$ satisfying $\mathbf{x}(0) = x^0$ and $\mathbf{x}(T) = x^1$.

Now assume $\mathbf{x}^*(\cdot)$ solves our variational problem. The fundamental question is this: how can we characterize $\mathbf{x}^*(\cdot)$?

Now to the discuss derivation of Euler-Lagrange equations, we have to note that we write $L = L(x, v)$, and regard the variable x as denoting position, the variable v as denoting velocity. The partial derivatives of L are

$$\frac{\partial L}{\partial x_i} = L_{x_i}, \qquad \frac{\partial L}{\partial v_i} = L_{v_i} \qquad (1 \le i \le n),$$

and we write

$$\nabla_x L := \big(L_{x_1}, \dots, L_{x_n}\big), \quad \nabla_v L := \big(L_{v_1}, \dots, L_{v_n}\big).$$

Theorem (2.2.1): (Euler-Lagrange Equations)

Let $\mathbf{x}^*(\cdot)$ solve the calculus of variations problem. Then $\mathbf{x}^*(\cdot)$ solves the Euler.Lagrange differential equations:

$$\text{(E} - \text{L)} \qquad \frac{d}{dt}\big[\nabla_v L\big(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t)\big)\big] = \nabla_x L\big(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t)\big).$$

The significance of preceding theorem is that if we can solve the Euler-Lagrange equations (E-L), then the solution of our original calculus of variations problem (assuming it exists) will be among the solutions.

Note that (E-L) is a quasilinear system of n second-order ODE. The i^{th} component of the system reads

$$\frac{d}{dt}\left[L_{v_i}\big(\mathbf{x}^*(t),\dot{\mathbf{x}}^*(t)\big)\right] = L_{x_i}\big(\mathbf{x}^*(t),\dot{\mathbf{x}}^*(t)\big).$$

Proof:

(i) Select any smooth curve $\mathbf{y}[0,T] \to \mathbb{R}^n$, satisfying $\mathbf{y}(0) = \mathbf{y}(T) = 0$. Define

$$i(\tau) := I[\mathbf{x}(\cdot) + \mathcal{T}\mathbf{y}(\cdot)]$$

for $\mathcal{T} \in \mathbb{R}$ and $\mathbf{x}(\cdot) = \mathbf{x}^*(\cdot)$. (To simplify we omit the superscript $*$). Notice that $\mathbf{x}(\cdot) + \mathcal{T}\mathbf{y}(\cdot)$ takes on the proper values at the endpoints. Hence, since $\mathbf{x}(\cdot)$ is minimizer, we have

$$i(\mathcal{T}) \geq I\big(\mathbf{x}(\cdot)\big) = i(0).$$

Consequently $i(\cdot)$ has a minimum at $\mathcal{T} = 0$, and so

$$i'(0) = 0.$$

(ii) We must compute $i'(\mathcal{T})$. Note first that

$$i(\mathcal{T}) = \int_0^T L\big(\mathbf{x}(t) + \mathcal{T}\mathbf{y}(t), \dot{\mathbf{x}}(t) + \mathcal{T}\mathbf{y}(t)\big)\, dt;$$

and hence

$$i'(\mathcal{T}) = \int_0^T \left(\sum_{i=1}^{n} L_{x_i}\big(\mathbf{x}(t) + \mathcal{T}\mathbf{y}(t) + \dot{\mathbf{x}}(t) + \mathcal{T}\dot{\mathbf{y}}(t)\big)\, \mathbf{y}_i(t) + \sum_{i=1}^{n} L_{v_i}(\cdots)\dot{\mathbf{y}}_i(t) \right) dt.$$

Let $\mathcal{T} = 0$. Then

$$0 = i'(0) = \sum_{i=1}^{n} \int_0^T L_{x_i}\big(\mathbf{x}(t), \dot{\mathbf{x}}(t) y_i(t)\big) + L_{x_i}\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big)\dot{y}(t)\, dt.$$

This equality holds for all choices of $\mathbf{y} :\ [0,T] \to \mathbb{R}^n$, with $\mathbf{y}(0) = \mathbf{y}(T) = 0$.

(iii) Fix any $1 \leq j \leq n$. Choose $\mathbf{y}(\cdot)$ so that

$$y_i(t) \equiv 0 \quad i \neq j, \;\; y_j(t) = \psi(t),$$

where ψ is an arbitary function. Use this choice of $\mathbf{y}(\cdot)$ above:

$$0 = \int_0^T L_{x_j}\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big)\psi(t) + L_{v_j}\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big)\psi(t)\, dt.$$

Integrate by parts, recalling that $\psi(0) = \psi(T) = 0$:

$$0 = \int_0^T \left[L_{x_j}\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big)\psi(t) - \frac{d}{dt}\Big(L_{v_j}\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big)\Big)\right]\psi(t)\, dt.$$

This holds for all $\psi :\ [0, T] \to \mathbb{R}\ \psi(0) = \psi(T) = 0$ and therefore

$$L_{x_j}\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big) - \frac{d}{dt}\Big(L_{v_j}\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big)\Big) = 0$$

for all times $0 \leq t \leq T$. To see this, observe that otherwise $L_{x_j} - \frac{d}{dt}\big(L_{v_j}\big)$ would be, say, positive on some subinterval on $I \subseteq [0, T]$. Choose $\psi \equiv 0$ off $I, \psi > 0$ on I. Then

$$\int_0^T \left(L_{x_j} - \frac{d}{dt}\big(L_{v_j}\big)\right)\psi\, dt > 0,$$

a contradiction.

Definition (2.2.2):

For the given curve $\mathbf{x}(\cdot)$, define

$$\mathbf{p}(t) := \nabla_v L\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big) \quad (0 \leq t \leq T).$$

We call $\mathbf{p}(\cdot)$ the generalized momentum.

Our intention now is to rewrite the Euler-Lagrange equations as a system of first-order ODE for $\mathbf{x}(\cdot), \mathbf{p}(\cdot)$.

Assume that for all $x, p \in \mathbb{R}^n$, we can solve the equation

$$p = \nabla_v L(x, v) \tag{2.2}$$

for v in terms of x and p. That is, we suppose we can solve the identity (2.2) for $v = v(x, p)$.

Definition (2.2.3):

Define the dynamical systems Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by the formula

$$H(x, p) = p \cdot \boldsymbol{v}(x, p) - L\big(x, \boldsymbol{v}(x, p)\big),$$

where $\boldsymbol{v}$ is defined above.

Remember that, the partial derivatives of H are

$$\frac{\partial H}{\partial x_i} = H_{x_i}, \frac{\partial H}{\partial p_i} = H_{p_i} \quad (1 \le i \le n),$$

and we write

$$\nabla_x H := \big(H_{x_1}, \dots, H_{x_n}\big), \quad \nabla_p H := \big(H_{p_1}, \dots, H_{p_n}\big).$$

Theorem (2.2.4): (Hamiltonian Dynamics)

Let $\mathbf{x}(\cdot)$ solve the Euler-Lagrange equations (E-L) and define $\mathbf{p}(\cdot)$ as above. Then the pair $\big(\mathbf{x}(\cdot), \mathbf{p}(\cdot)\big)$ solves Hamilton's equations:

$$\text{(H)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = \nabla_p H\big(\mathbf{x}(t), \mathbf{p}(t)\big) \\ \dot{\mathbf{p}}(t) = -\nabla_x H\big(\mathbf{x}(t), \mathbf{p}(t)\big) \end{cases}$$

Furthermore, the mapping $t \mapsto H\big(\mathbf{x}(t), \mathbf{p}(t)\big)$ is constant.

Proof:

Recall that $H(x, p) = p \cdot \mathbf{v}(x, p) - L\big(x, \mathbf{v}(x, p)\big)$, where $v = \mathbf{v}(x, p)$ or, equivalently, $p = \nabla_v L(x, v)$. Then

$$\begin{aligned} \nabla_x H(x, p) &= p \cdot \nabla_x \mathbf{v} - \nabla_x L\big(x, \mathbf{v}(x, p)\big) - \nabla_v L\big(x, \mathbf{v}(x, p)\big) \cdot \nabla_x \mathbf{v} \\ &= -\nabla_x L\big(x, \mathbf{v}(x, p)\big) \end{aligned}$$

because $p = \nabla_v L$. Now $\mathbf{p}(t) = \nabla_v L\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big)$ if and only if $\dot{\mathbf{x}}(t) = \mathbf{v}\big(\mathbf{x}(t), \mathbf{p}(t)\big)$. Therefore (E-L) implies

$$\dot{\mathbf{p}}(t) = \nabla_x L\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big)$$

$$= \nabla_x L\left(\mathbf{x}(t), \mathbf{v}(\mathbf{x}(t), \mathbf{p}(t))\right) = -\nabla_x H(\mathbf{x}(t), \mathbf{p}(t)).$$

Also

$$\nabla_p H(x,p) = \mathbf{v}(x,p) + p \cdot \nabla_p - \mathbf{v}\nabla_v L \cdot \nabla_p \mathbf{v} = \mathbf{v}(x,p)$$

since $p = \nabla_v L(x, \mathbf{v}(x,p))$. This implies

$$\nabla_p H(\mathbf{x}(t), \mathbf{p}(t)) = \mathbf{v}(\mathbf{x}(t), \mathbf{p}(t)).$$

But

$$\mathbf{p}(t) = \nabla_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t))$$

and so $\dot{\mathbf{x}}(t) = \mathbf{v}(\mathbf{x}(t), \mathbf{p}(t))$. Therefore

$$\dot{\mathbf{x}}(t) = \nabla_p H(\mathbf{x}(t), \mathbf{p}(t)).$$

Finally note that

$$\frac{d}{dt} H(\mathbf{x}(t), \mathbf{p}(t)) = \nabla_x H \cdot \dot{\mathbf{x}}(t) + \nabla_p H \cdot \dot{\mathbf{p}}(t) = \nabla_x H \cdot \nabla_p H + \nabla_p H \cdot (-\nabla_x H) = 0.$$

Now let us discuss a physical example:

We define the Lagrangian

$$L(x,v) = \frac{m|v|^2}{2} - V(x),$$

which we interpret as the kinetic energy minus the potential energy V. Then

$$\nabla_x L = -\nabla V(x), \quad \nabla_v L = mv.$$

Therefore the Euler-Lagrange equation is

$$m\ddot{\mathbf{x}}(t) = -\nabla V(\mathbf{x}(t)),$$

which is Newton's law. Furthermore

$$p = \nabla_v L(x,v) = mv$$

is the momentum, and the Hamiltonian is

$$H(x,p) = p \cdot \frac{p}{m} - L\left(x, \frac{p}{m}\right) = \frac{|p|^2}{m} - \frac{m}{2}\left|\frac{p}{m}\right|^2 + V(x) = \frac{|p|^2}{2m} + V(x),$$

the sum of the kinetic and potential energies. For this example, Hamilton's equations read

$$\begin{cases} \dot{\mathbf{x}}(t) = \dfrac{\mathbf{p}(t)}{m} \\ \dot{\mathbf{p}}(t) = -\nabla V(\mathbf{x}(t)). \end{cases}$$

What first strikes us about general optimal control problems is the occurence of many constraints, most notably that the dynamics be governed by the differential equation

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \boldsymbol{\alpha}(t)) \quad (t > 0) \\ \mathbf{x}(0) = x^0. \end{cases}$$

This is in contrast to standard calculus of variations problems, as discussed in section (2.1), where we could take any curve $\mathbf{x}(\cdot)$ as a candidate for a minimizer.

Now it is a general principle of variational and optimization theory that "constraints create Lagrange multipliers" and furthermore that these Lagrange multipliers often "contain valuable information''. This section provides a quick review of the standard method of Lagrange multipliers in solving multivariable constrained optimization problems.

Suppose first that we wish to find a maximum point for a given smooth function $f : \mathbb{R}^n \to \mathbb{R}$. In this case there is no constraint, and therefore if $f(x^*) = \max_{x \in \mathbb{R}^n} f(x)$, then x^* is a critical point of f:

$$\nabla f(x^*) = 0.$$

Hence to discuss the constrained optimization, we modify the problem above by introducing the region

$$R := \{x \in \mathbb{R}^n |\, g(x) \leq 0\},$$

determined by some given function $g : \mathbb{R}^n \to \mathbb{R}$. Suppose $x^* \in R$ and $f(x^*) = \max_{x \in R} f(x)$. We would like a characterization of x^* in terms of the gradients of f and g.

Case (1):x^* lies in the interior of R

Then the constraint is inactive, and so

$$\nabla f(x^*) = 0. \tag{2.3}$$

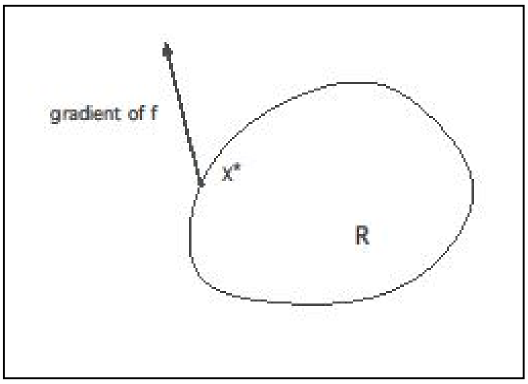

Figure (2.3)

Case (2): x^* lies on ∂R

We look at the direction of the vector $\nabla f(x^*)$. A geometric picture like Figure (2.3) is impossible; for if it were so, then $f(y^*)$ would be greater that $f(x^*)$ for some other point $y^* \in \partial R$. So it must be $\nabla f(x^*)$ is perpendicular to ∂R at x^*, as shown in Figure (2.4).

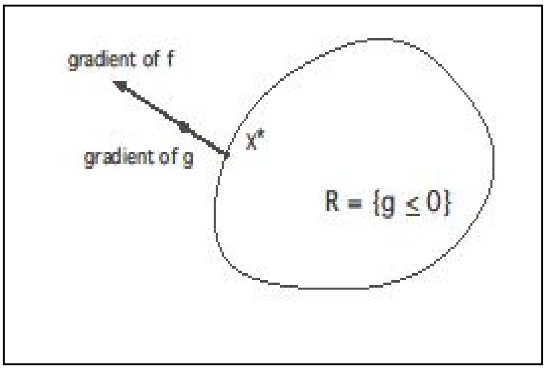

Figure (2.4)

Since ∇g is perpendicular to $\partial R = \{g = 0\}$, it follows that $\nabla f(x^*)$ is parallel to $\nabla g(x^*)$. Therefore

$$\nabla f(x^*) = \lambda \nabla g(x^*) \tag{2.4}$$

for some real number λ, called a Lagrange multiplier.

The foregoing argument is in fact incomplete, since we implicitly assumed that $\nabla g(x^*) \neq 0$, in which case the Implicit Function Theorem implies that the set $\{g = 0\}$ is an $(n-1)$-dimensional surface near x^* (as illustrated).

If instead $\nabla g(x^*) = 0$, the set $\{g = 0\}$ need not have this simple form near x^*; and the reasoning discussed as Case (2) above is not complete.

The correct statement is this:

$$\begin{cases} \text{There exist real numbers } \lambda \text{ and } \mu, \text{not both equal to } 0, \text{such that} \\ \qquad\qquad \mu \nabla f(x^*) = \lambda \nabla g(x^*). \end{cases} \tag{2.5}$$

If $\mu \neq 0$, we can divide by μ and convert to the formulation (2.4). And if $\nabla g(x^*) = 0$,we can take $\lambda = 1, \mu = 0$, making assertion (2.5) correct (if not particularly useful).

We come now to the key assertion of this section, the theoretically interesting and practically useful theorem that if $\boldsymbol{\alpha}^*(\cdot)$ is an optimal control, then there exists a function $\mathbf{p}^*(\cdot)$, called the costate, that satisfies a certain maximization principle. We should think of the function $\mathbf{p}^*(\cdot)$ as a sort of Lagrange multiplier, which appears owing to the constraint that the optimal curve $\mathbf{x}^*(\cdot)$ must satisfy (ODE). And just as conventional Lagrange multipliers are useful for actual calculations, so also will be the costate.

We quote Francis Clarke [5]: "The maximum principle was, in fact, the culmination of a long search in the calculus of variations for a comprehensive multiplier rule, which is the correct way to view it: $p(t)$ is a " Lagrange multiplier " ... It makes optimal control a design tool, whereas the calculus of variations was a way to study nature ".

Now Let us review the basic set-up for our control problem.

We are given $A \subseteq \mathbb{R}^m$ and also $\mathbf{f} : \mathbb{R}^n \times A \to \mathbb{R}^n, x^0 \in \mathbb{R}^n$. We as before denote the set of admissible controls by

$$\mathcal{A} = \{\boldsymbol{\alpha}(\cdot) : [0, \infty) \to A \mid \boldsymbol{\alpha}(\cdot) \text{ is measurable}\}.$$

Then given $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$, we solve for the corresponding evolution of our system:

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\alpha}(t)) & (t \geq 0) \\ \mathbf{x}(t) = x^0. \end{cases}$$

We also introduce the payoff functional

(P) $$[\boldsymbol{\alpha}(\cdot)] = \int_0^T r\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big)\, dt + g\big(\mathbf{x}(T)\big),$$

where the terminal time $T > 0$, running payoff $r : \mathbb{R}^n \times A \to \mathbb{R}$ and terminal payoff $g : \mathbb{R}^n \to \mathbb{R}$ are given.

Our basic problem, is to find a control $\alpha^*(\cdot)$ such that

$$[\boldsymbol{\alpha}(\cdot)] = \max_{\boldsymbol{\alpha}(\cdot)\in\mathcal{A}} P[\boldsymbol{\alpha}(\cdot)].$$

The Pontryagin Maximum Principle, stated below, asserts the existence of a function $\boldsymbol{P}^*(\cdot)$, which together with the optimal trajectory $\mathbf{x}^*(\cdot)$ satisfies an analog of Hamilton's ODE from section (2.1). For this, we will need an appropriate Hamiltonian:

Definition (2.2.5):

The control theory Hamiltonian is the function

$$H(x, p, a) := \mathbf{f}(x, a) \cdot p + r(x, a) \quad (x, p \in \mathbb{R}^n, a \in A).$$

Theorem (2.2.6): (Pontryagin Maximum Principle)

Assume $\boldsymbol{\alpha}^*(\cdot)$ is optimal for (ODE), (P) and $\mathbf{x}^*(\cdot)$ is the corresponding trajectory. Then there exists a function $\mathbf{p}^* : [0, T] \to \mathbb{R}^n$ such that

(ODE) $$\dot{\mathbf{x}}^*(t) = \nabla_p H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big),$$

(ADJ) $$\mathbf{p}^*(t) = -\nabla_x H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big),$$

and

(M) $$H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big) = \max_{a\in A} H(\mathbf{x}^*(t), \mathbf{p}^*(t), a) \quad (0 \le t \le T)$$

In addition,

$$\text{the mapping} \quad t \mapsto H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big) \quad \text{is constant.}$$

Finally, we have the terminal condition

(T) $$\mathbf{p}^*(T) = \nabla g\big(\mathbf{x}^*(T)\big)$$

Remarks and Interpretations (2.2.7):

(i) The identities (ADJ) are the adjoint equations and (M) the maximization principle. Notice that (ODE) and (ADJ) resemble the structure of Hamilton's equations, discussed in section (2.1).

We also call (T) the transversality condition and will discuss its significance later.

(ii) More precisely, formula (ODE) says that for $1 \leq i \leq n$, we have

$$\dot{x}^{i_*}(t) = H_{p_i}\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big) = f^i\big(\mathbf{x}^*(t), \boldsymbol{\alpha}^*(t)\big),$$

which is just the original equation of motion. Likewise, (ADJ) says

$$\begin{aligned} \dot{p}^{i_*}(t) &= -H_{x_i}\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big) \\ &= -\sum_{j=1}^{n} p^{j^*}(t) f^j_{x_i}\big(\mathbf{x}^*(t), \boldsymbol{\alpha}^*(t)\big) - r_{x_i}\big(\mathbf{x}^*(t), \boldsymbol{\alpha}^*(t)\big). \end{aligned}$$

Let us next record the appropriate form of the Maximum Principle for a fixed endpoint problem.

As before, given a control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$, we solve for the corresponding evolution of our system:

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big) & (t \geq 0) \\ \mathbf{x}(0) = x^0. \end{cases}$$

Assume now that a target point $x^1 \in \mathbb{R}^n$ is given. We introduce then the payoff functional

$$\text{(P)} \qquad P[\boldsymbol{\alpha}(\cdot)] = \int_0^T r\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big)\, dt$$

Here $r : \mathbb{R}^n \times A \to \mathbb{R}$ is the given running payoff, and $\mathcal{T} = \mathcal{T}[\boldsymbol{\alpha}(\cdot)] \leq \infty$ denotes the first time the solution of (ODE) hits the target point x^1.

As before, the basic problem is to find an optimal control $\boldsymbol{\alpha}^*(\cdot)$ such that

$$P[\boldsymbol{\alpha}^*(\cdot)] = \max_{\boldsymbol{\alpha}(\cdot) \in \mathcal{A}} P[\boldsymbol{\alpha}(\cdot)].$$

Define the Hamiltonian H as in section (2.2).

Theorem (2.2.8): (Pontryagin Maximum Principle)

Assume $\boldsymbol{\alpha}^*(\cdot)$ is optimal for (ODE), (P) and $\mathbf{x}^*(\cdot)$ is the corresponding trajectory. Then there exists a function $\mathbf{p}^* : [0,\mathcal{T}^*] \to \mathbb{R}$ such that

$$\text{(ODE)} \qquad \dot{\mathbf{x}}^*(t) = \nabla_p H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big),$$

$$\text{(ADJ)} \qquad \dot{\mathbf{p}}^*(t) = -\nabla_x H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big),$$

and

$$\text{(M)} \qquad H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big) = \max_{a\in A} H(\mathbf{x}^*(t), \mathbf{p}^*(t), a) \quad (0 \le t \le \mathcal{T}^*).$$

Also,

$$H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big) \equiv 0 \quad (0 \le t \le \mathcal{T}^*).$$

Here $\mathcal{T}^*$ denotes the first time the trajectory $x^*(\cdot)$ hits the target point x^1. We call $\mathbf{x}^*(\cdot)$ the state of the optimally controlled system and $\mathbf{p}^*(\cdot)$ the costate.

More precisely, we should define

$$H(x,p,q,a) = \mathbf{f}(x,a) \cdot p + r(x,a)q \quad (q \in \mathbb{R}).$$

A more careful statement of the Maximum Principle says "there exists a constant $q \ge 0$ and a function $\mathbf{p}^* : [0,t^*] \to \mathbb{R}^n$ such that (ODE), (ADJ), and (M) hold ".

If $q > 0$, we can renormalize to get $q = 1$, as we have done above. If $q = 0$, then H does not depend on running payoff r and in this case the Pontryagin Maximum Principle is not useful. This is a so-called " abnormal problem ".

Recall our discussion in section (2.2) about finding a point x^* that maximizes a function f, subject to the requirement that $g \le 0$. Now $x^* = (x_1^*, \dots, x_n^*)^T$ has n unknown components we must find. Somewhat unexpectedly, it turns out in practice to be easier to solve (2.2) for the $n+1$ unknowns $x_1^*, \dots, x_n^*$ and λ. We repeat this key insight: it is actually easier to solve the problem if we add a new unknown, namely the Lagrange multiplier.

This same principle is valid for our much more complicated control theory problems: it is usually best not just to look for an optimal control $\boldsymbol{\alpha}^*(\cdot)$ and an optimal trajectory $\mathbf{x}^*(\cdot)$ alone, but also to look as well for the costate $\mathbf{p}^*(\cdot)$. In practice, we add the equations (ADJ) and (M) to (ODE) and then try to solve for $\boldsymbol{\alpha}^*(\cdot)\mathbf{x}^*(\cdot)$ and for $\mathbf{p}^*(\cdot)$.

The following examples show how this works in practice, in certain cases for which we can actually solve everything explicitly or, failing that, at least deduce some useful information.

Example (2.2.9): (Linear Time-optimal Control)

For this example, let A denote the cube $[-1,1]^n$ in $\mathbb{R}^n$. We consider again the linear dynamics:

(ODE) $$\begin{cases}\dot{\mathbf{x}}(t) = M\mathbf{x}(t) + N\boldsymbol{\alpha}(t)\\ \mathbf{x}(0) = x^0.\end{cases}$$

for the payoff functional

(P) $$\boldsymbol{P}[\boldsymbol{\alpha}(\cdot)] = -\int_0^{\mathcal{T}} 1\, dt = -\mathcal{T},$$

where $\mathcal{T}$ denotes the first time the trajectory hits the target point $x^1 = 0$. We have $\mathcal{T} \equiv -1$, and so

$$H(x,p,a) = \mathbf{f}\cdot p + r = (Mx + Na)\cdot p - 1.$$

In section (2.1) we introduced the Hamiltonian $H = (Mx + Na)\cdot p$, which differs by a constant from the present H. We can redefine H in section(2.1) to match the present theory: compare then Theorems (2.1.5) and (2.2.8).

Example (2.2.10): (Control of Production and Consumption)

We return to Example (1.1.2) in Chapter (1), a model for optimal consumption in a simple economy. Recall that

$$x(t) = \text{output of economy at time } t,$$

$$\alpha(t) = \text{fraction of output reinvested at time } t$$

We have the constraint $0 \le \alpha(t) \le 1$; that is, $A = [0,1] \subset \mathbb{R}$. The economy evolves according to the dynamics

(ODE) $$\begin{cases}\dot{x}(t) = \alpha(t)x(t) \quad (0 \le t \le T)\\ x(0) = x^0.\end{cases}$$

where $x^0 > 0$ and we have set the growth factor $k = 1$. We want to maximize the total consumption

(P) $$P[\alpha(\cdot)] := \int_0^T \big(1-\alpha(t)\big)x(t)\,dt$$

How can we characterize an optimal control $x^*(\cdot)$?

We apply Pontryagin Maximum Principle, and to simplify notation we will not write the superscripts $*$ for the optimal control, trajectory, etc. We have $n = m = 1$,

$$f(x,a) = xa, \quad g \equiv 0, \quad r(x,a) = (1-a)x;$$

and therefore

$$H(x,p,a) = f(x,a)p + r(x,a) = pxa + (1-a)x = x + ax(p-1).$$

The dynamical equation is

(ODE) $$\dot{x}(t) = H_p = \alpha(t)x(t),$$

and the adjoint equation is

(ADJ) $$\dot{p}(t) = -H_x = -1 - \alpha(t)(p(t)-1).$$

The terminal condition reads

(T) $$p(T) = g_x = 0.$$

Lastly, the maximality principle asserts

(M) $$H\big(x(t),p(t),\alpha(t)\big) = \max_{0\le a\le 1}\{x(t) + ax(t)(p(t)-1)\}.$$

We now deduce useful information from (ODE), (ADJ), (M) and (T).

According to (M), at each time t the control value $\alpha(t)$ must be selected to maximize $a(p(t)-1)$ for $0 \le a \le 1$. This is so, since $x(t) > 0$. Thus

$$\alpha(t) = \begin{cases} 1 & \text{if } p(t) > 1 \\ 0 & \text{if } p(t) \le 1. \end{cases}$$

Hence if we know $p(\cdot)$, we can design the optimal control $\alpha(\cdot)$.

So next we must solve for the costate $p(\cdot)$. We know from (ADJ) and (T) that

$$\begin{cases} \dot{p}(t) = -1 - \alpha(t)[p(t)-1] \quad (0 \le t \le T) \\ p(T) = 0. \end{cases}$$

Since $p(T) = 0$, we deduce by continuity that $p(t) \leq 1$ for t close to $T, t < T$. Thus $\alpha(t) = 0$ for such values of t. Therefore $\dot{p}(t) = -1$, and consequently $p(t) = T - t$ for times t in this interval. So we have that $p(t) = T - t$ so long as $p(t) \leq 1$. And this holds for $T - 1 \leq t \leq T$.

But for times $t \leq T - 1$, with t near $T - 1$, we have $\alpha(t) = 1$; and so (ADJ) becomes

$$\dot{p}(t) = -1 - (p(t) - 1) = -p(t).$$

Since $p(T - 1) = 1$, we see that $p(t) = e^{T-1-t} > 1$ for all times $0 \leq t \leq T - 1$. In particular there are no switches in the control over this time interval.

Restoring the superscript $*$ to our notation, we consequently deduce that an optimal control is

$$\alpha^*(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq t^* \\ 0 & \text{if } t^* \leq t \leq T \end{cases}$$

for the optimal switching time $t^* = T - 1$.

Example (2.2.11): (A Simple linear-Quadratic Regulator)

We take $n = m = 1$ for this example, and consider the simple linear dynamics

$$\text{(ODE)} \qquad \begin{cases} \dot{x}(t) = x(t) + \alpha(t) \\ x(0) = x^0, \end{cases}$$

with the quadratic cost functional

$$\int_0^T x(t)^2 + \alpha(t)^2 \, dt,$$

which we want to minimize. So we want to maximize the payoff functional

$$\text{(P)} \qquad P[\alpha(\cdot)] = -\int_0^T x(t)^2 + \alpha(t)^2 \, dt.$$

For this problem, the values of the controls are not constrained; that is, $A = \mathbb{R}$.

To simplify notation further we again drop the superscripts $*$. We have $n = m = 1$,

$$f(x, a) = x + a, \quad g \equiv 0, \quad r(x, a) = -x^2 - a^2;$$

and hence

$$H(x,p,a) = fp + r = (x+a)p - (x^2 + a^2)$$

The maximality condition becomes

$$\text{(M)} \qquad H(x(t), p(t), \alpha(t)) = \max_{a \in \mathbb{R}} \{-(x(t)^2 + a^2) + p(t)(x(t) + a)\}$$

We calculate the maximum on the right hand side by setting $H_a = -2a + p = 0$. Thus $a = \frac{p}{2}$, and so

$$\alpha(t) = \frac{p(t)}{2}.$$

The dynamical equations are therefore

$$\text{(ODE)} \qquad \dot{x}(t) = x(t) + \frac{p(t)}{2}$$

and

$$\text{(ADJ)} \qquad \dot{p}(t) = -H_x = 2x(t) - p(t).$$

Moreover $x(0) = x^0$, and the terminal condition is

$$\text{(T)} \qquad p(T) = 0.$$

So we must look at the system of equations

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1/2 \\ 2 & -1 \end{pmatrix}}_{M} \begin{pmatrix} x \\ p \end{pmatrix},$$

the general solution of which is

$$\begin{pmatrix} x(t) \\ p(t) \end{pmatrix} = e^{tM} \begin{pmatrix} x^0 \\ p^0 \end{pmatrix}.$$

Since we know x^0, the task is to choose p^0 so that $p(T) = 0$.

An elegant way to do so is to try to find optimal control in linear feedback form; that is, to look for a function $c(\cdot) : [0,T] \to \mathbb{R}$ for which

$$\alpha(t) = c(t)\, x(t).$$

We henceforth suppose that an optimal feedback control of this form exists, and attempt to calculate $c(\cdot)$. Now

$$\frac{p(t)}{2} = \alpha(t) = c(t)x(t);$$

whence $c(t) = \frac{p(t)}{2x(t)}$. Define now

$$d(t) := \frac{p(t)}{x(t)};$$

so that $c(t) = \frac{d(t)}{2}$.

We will next discover a differential equation that $d(\cdot)$ satisfies. Compute

$$\dot{d} = \frac{\dot{p}}{x} - \frac{p\dot{x}}{x^2},$$

and recall that

$$\begin{cases} \dot{x} = x + \frac{p}{2} \\ \dot{p} = 2x - p. \end{cases}$$

Therefore

$$\dot{d} = \frac{2x - p}{x} - \frac{p}{x^2}\left(x + \frac{p}{2}\right) = 2 - d - d\left(1 + \frac{d}{2}\right) = 2 - 2d - \frac{d^2}{2}.$$

Since $p(T) = 0$, the terminal condition is $d(T) = 0$.

So we have obtained a nonlinear first-order ODE for $d(\cdot)$ with a terminal boundary condition:

(R)
$$\begin{cases} \dot{d} = 2 - 2d - \frac{1}{2}d^2 \quad (0 \le t \le T) \\ d(T) = 0. \end{cases}$$

This is called the Riccati equation.

In summary so far, to solve our linear-quadratic regulator problem, we need to first solve the Riccati equation (R) and then set

$$\alpha(t) = \frac{1}{2}d(t)x(t).$$

How to solve the Riccati equation. It turns out that we can convert (R) it into a second-order, linear ODE. To accomplish this, write

$$d(t) = \frac{2\dot{b}(t)}{b(t)}$$

for a function $b(\cdot)$ to be found. What equation does $b(\cdot)$ solve? We compute

$$\dot{d} = \frac{2\ddot{b}}{b} - \frac{2(\dot{b})^2}{b^2} = \frac{2\ddot{b}}{b} - \frac{d^2}{2}.$$

Hence (R) gives

$$\frac{2\ddot{b}}{b} = \dot{d} + \frac{d^2}{2} = 2 - 2d = 2 - 2\frac{2\dot{b}}{b};$$

and consequently

$$\begin{cases} \ddot{b} = b - 2\dot{b} \quad (0 \leq t < T) \\ \dot{b}(T) = b\,, b(T) = 1. \end{cases}$$

This is a terminal-value problem for second-order linear ODE, which we can solve by standard techniques. We then set $d = \frac{2\dot{b}}{b}$, to derive the solution of the Riccati equation (R).

Example (2.2.12): (Moon Lander)

This is a much more elaborate and interesting example, already introduced in Chapter (1). We follow the discussion of [17].

Introduce the notation

$h(t) = \text{height at time } t$

$v(t) = \text{velocity} = \dot{h}(t)$

$m(t) = \text{mass of spacecraft (changing as fuel is used up)}$

$\alpha(t) = \text{thrust at time } t.$

The thrust is constrained so that $0 \leq \alpha(t) \leq 1$; that is, $A = [0,1]$. There are also the constraints that the height and mass be nonnegative: $h(t) \geq 0, m(t) \geq 0$.

The dynamics are

$$
\text{(ODE)} \qquad \begin{cases} \dot{h}(t) = v(t) \\ \dot{v}(t) = -g + \dfrac{\alpha(t)}{m(t)} \\ \dot{m}(t) = -k\alpha(t), \end{cases}
$$

with initial conditions

$$
\begin{cases} h(0) = h_0 > 0 \\ v(0) = v_0 \\ m(0) = m_0 > 0. \end{cases}
$$

The goal is to land on the moon safely, maximizing the remaining fuel $m(\mathcal{T})$, where $\mathcal{T} = \mathcal{T}[\alpha(\cdot)]$ is the first time $h(\mathcal{T}) = v(\mathcal{T}) = 0$. Since $\alpha = -\frac{\dot{m}}{k}$, our intention is equivalently to minimize the total applied thrust before landing; so that

$$
\text{(P)} \qquad P[\alpha(\cdot)] = -\int_0^{\mathcal{T}} \alpha(t)\, dt.
$$

This is so since

$$
\int_0^{\mathcal{T}} \alpha(t)\, dt = \frac{m_0 - m(\mathcal{T})}{k}.
$$

In terms of the general notation, we have

$$
\mathbf{x}(t) = \begin{pmatrix} h(t) \\ v(t) \\ m(t) \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} v \\ -g + a/m \\ -ka \end{pmatrix}.
$$

Hence the Hamiltonian is

$$
\begin{aligned}
H(x, p, a) &= \mathbf{f} \cdot p + r \\
&= (v, -g + a/m\,, -ka) \cdot (p_1, p_2, p_3) - a \\
&= -a + p_1 v + p_2\left(-g + \frac{a}{m}\right) + p_3(-ka).
\end{aligned}
$$

We next have to figure out the adjoint dynamics (ADJ). For our particular Hamiltonian,

$$H_{x_1} = H_h = 0, \;\; H_{x_2} = H_v = p_1, \;\; H_{x_3} = H_m = -\frac{p_2 a}{m^2}.$$

Therefore

$$\text{(ADJ)} \qquad \begin{cases} \dot{p}^1(t) = 0 \\ \dot{p}^2(t) = -p^1(t) \\ \dot{p}^3(t) = \dfrac{p^2(t)\alpha(t)}{m(t)^2}. \end{cases}$$

The maximization condition (M) reads

$$\text{(M)} \quad H(\mathbf{x}(t), \mathbf{p}(t), \alpha(t)) = \max_{0 \le a \le 1} H(\mathbf{x}(t), \mathbf{p}(t), a)$$

$$= \max_{0 \le a \le 1} \left\{ -a + p^1(t)v(t) + p^2(t)\left[-g + \frac{a}{m(t)}\right] + p^3(t)(-ka) \right\}$$

$$= p^1(t)v(t) - p^2(t) + \max_{0 \le a \le 1} \left\{ a\left(-1 + \frac{p^2(t)}{m(t)} - kp^3(t)\right) \right\}.$$

Thus the optimal control law is given by the rule:

$$\alpha(t) = \begin{cases} 1 & \text{if } 1 - \dfrac{p^2(t)}{m(t)} + kp^3(t) < 0 \\ 0 & \text{if } 1 - \dfrac{p^2(t)}{m(t)} + kp^3(t) > 0. \end{cases}$$

Now we will attempt to figure out the form of the solution, and check it accords with the Maximum Principle.

Let us start by guessing that we first leave rocket engine of (i.e., set $\alpha \equiv 0$) and turn the engine on only at the end. Denote by $\mathcal{T}$ the first time that $h(\mathcal{T}) = v(\mathcal{T}) = 0$, meaning that we have landed. We guess that there exists a switching time $t^* < T$ when we turned engines on at full power (i.e., set $\alpha \equiv 1$).Consequently,

$$\alpha(t) = \begin{cases} 0 & \text{for} \quad 0 \le t \le t^* \\ 1 & \text{for} \quad t^* \le t \le \mathcal{T}. \end{cases}$$

Therefore, for times $t^* \le t \le \mathcal{T}$ our ODE becomes

$$\begin{cases} \dot{h}(t) = v(t) \\ \dot{v}(t) = -g + \dfrac{1}{m(t)} \quad (t^* \leq t \leq \mathcal{T}) \\ \dot{m}(t) = -k \end{cases}$$

with $h(\mathcal{T}) = 0, v(\mathcal{T}) = 0, m(t^*) = m_0$. We solve these dynamics:

$$\begin{cases} m(t) = m_0 + k(t^* - t) \\ v(t) = g(\mathcal{T} - t) + \dfrac{1}{k}\log\left[\dfrac{m_0 + k(t^* - \mathcal{T})}{m_0 + k(t^* - t)}\right] \\ h(t) = \text{complicated formua.} \end{cases}$$

Now put $t = t^*$:

$$\begin{cases} m(t^*) = m_0 \\ v(t^*) = g(\mathcal{T} - t^*) + \dfrac{1}{k}\log\left[\dfrac{m_0 + k(t^* - \mathcal{T})}{m_0}\right] \\ h(t^*) = -\dfrac{g(t^* - \mathcal{T})^2}{2} - \dfrac{m_0}{k^2}\log\left[\dfrac{m_0 + k(t^* - \mathcal{T})}{m_0}\right] + \dfrac{t^* - \mathcal{T}}{k}. \end{cases}$$

Suppose the total amount of fuel to start with was m_1; so that $m_0 - m_1$ is the weight of the empty spacecraft. When $\alpha \equiv 1$, the fuel is used up at rate k. Hence

$$k(\mathcal{T} - t^*) \leq m_1,$$

and so $0 \leq \mathcal{T} - t^* \leq \frac{m_1}{k}$.

Before time t^*, we set $\alpha \equiv 0$. Then (ODE) reads

$$\begin{cases} \dot{h} = v \\ \dot{v} = -g \\ \dot{m} = 0; \end{cases}$$

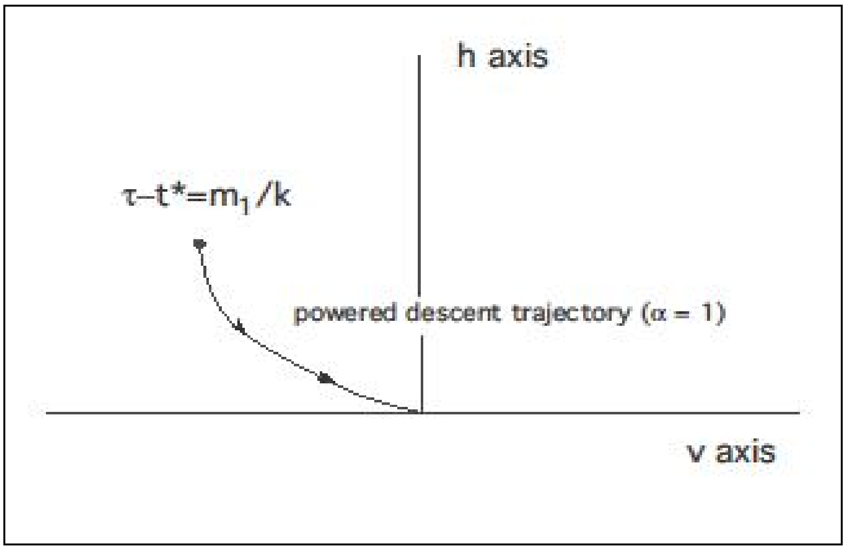

Figure (2.5)

and thus

$$\begin{cases} m(t) = m_0 \\ v(t) = -gt + v_0 \\ h(t) = -\dfrac{1}{2}gt^2 + tv_0 + h_0. \end{cases}$$

We combine the formulas for $v(t)$ and $h(t)$, to discover

$$h(t) = h_0 - \frac{1}{2g}(v^2(t) - v_0^2) \quad (0 \leq t \leq t^*).$$

We deduce that the freefall trajectory $\big(v(t), h(t)\big)$ therefore lies on a parabola

$$h = h_0 - \frac{1}{2g}(v^2 - v_0^2).$$

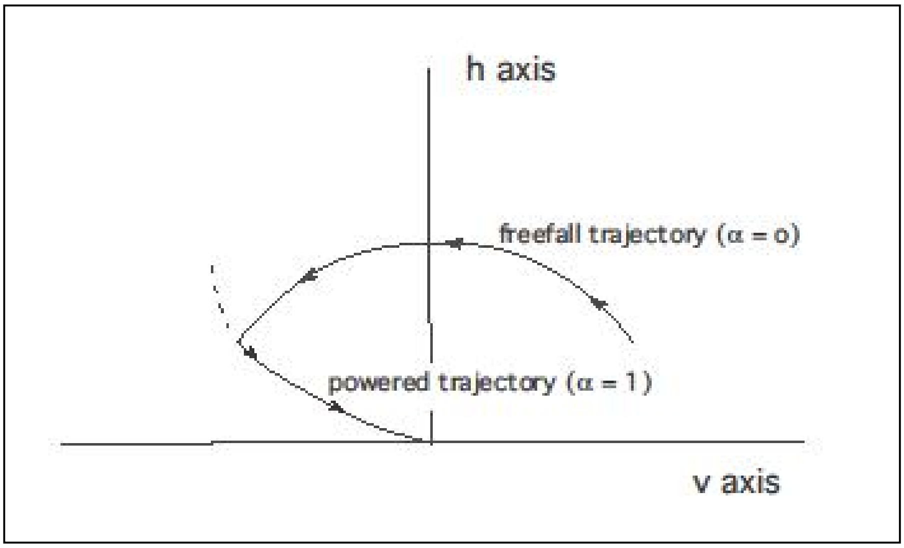

Figure (2.6)

If we then move along this parabola until we hit the soft-landing curve from the previous picture, we can then turn on the rocket engine and land safely.

In the second case illustrated, we miss switching curve, and hence cannot land safely on the moon switching once.

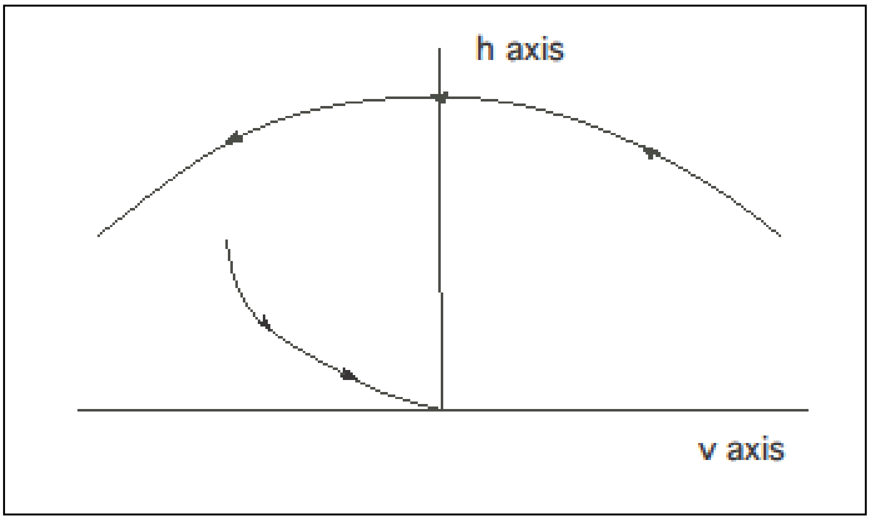

Figure (2.7)

To justify our guess about the structure of the optimal control, let us now find the costate $\mathbf{p}(\cdot)$ so that $\alpha(\cdot)$ and $\mathbf{x}(\cdot)$ described above satisfy (ODE), (ADJ), (M).To do this, we will have to figure out appropriate initial conditions

$$p^1(0) = \lambda_1, \quad p^2(0) = \lambda_2, \quad p^3(0) = \lambda_3.$$

We solve (ADJ) for $\alpha(\cdot)$ as above, and find

$$\begin{cases} p^1(t) \equiv \lambda_1 & (0 \le t \le \mathcal{T}) \\ p^2(t) = \lambda_2 - \lambda_1 t & (0 \le t \le \mathcal{T}) \\ p^3(t) = \begin{cases} \lambda_3 & (0 \le t \le t^*) \\ \lambda_3 + \int_{t^*}^{t} \dfrac{\lambda_2 - \lambda_1 s}{\big(m_0 + k(t^* - s)\big)^2} ds & (t^* \le t \le \mathcal{T}). \end{cases} \end{cases}$$

Define

$$r(t) := 1 - \frac{p^2(t)}{m(t)} + p^3(t)k;$$

then

$$\dot{r} = -\frac{\dot{p}^2}{m} + \frac{p^2\dot{m}}{m^2} + \dot{p}^3 k = \frac{\lambda_1}{m} + \frac{p^2}{m^2}(-k\alpha) + \left(\frac{p^2\alpha}{m^2}\right)k = \frac{\lambda_1}{m(t)}.$$

Choose $\lambda_1 < 0$, so that r is decreasing. We calculate

$$r(t^*) = 1 - \frac{(\lambda_2 - \lambda_1 t^*)}{m_0} + \lambda_3 k$$

and then adjust λ_2, λ_3 so that $r(t^*) = 0$.

Then r is nonincreasing, $r(t^*) = 0$, and consequently $r > 0$ on $[0, t^*), r < 0$ on $(t^*, \mathcal{T}]$. But (M) says

$$\alpha(t) = \begin{cases} 1 & \text{if } r(t) < 0 \\ 0 & \text{if } r(t) > 0. \end{cases}$$

Thus an optimal control changes just once from 0 to 1; and so our earlier guess of $\alpha(\cdot)$ does indeed satisfy the Pontryagin Maximum Principle.

Now to discuss the Maximum Principle with transversality conditions, consider again the dynamics

$$\text{(ODE)} \qquad \dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big) \quad (t \ge 0)$$

In this section we discuss another variant problem, one for which the initial position is constrained to lie in a given set $X_0 \subset \mathbb{R}^n$ and the final position is also constrained to lie within a given set $X_1 \subset \mathbb{R}^n$.

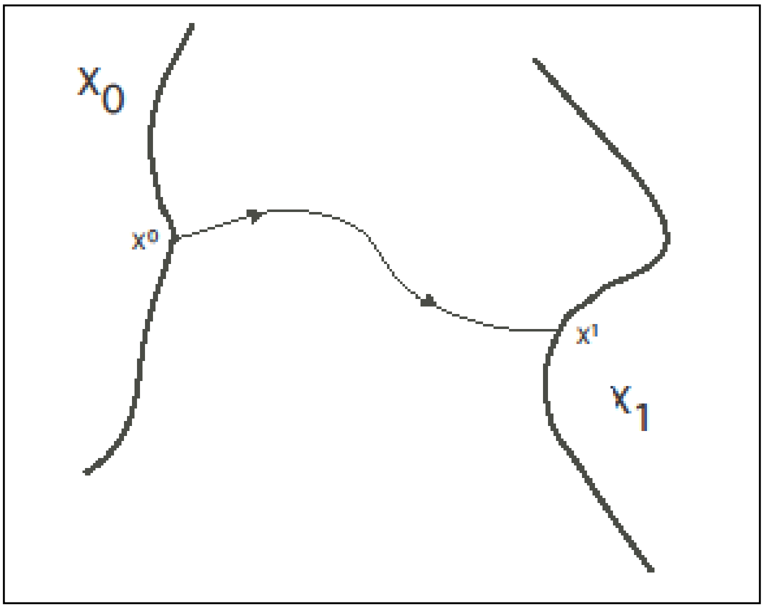

Figure (2.8)

So in this model we get to choose the starting point $x^0 \in X_0$ in order to maximize

(P) $$P[\boldsymbol{\alpha}(\cdot)] = \int_0^{\mathcal{T}} r\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big)\, dt,$$

where $\mathcal{T} = \mathcal{T}[\boldsymbol{\alpha}(\cdot)]$ is the first time we hit X_1.

In the following we will assume that X_0, X_1 are in fact smooth surfaces in $\mathbb{R}^n$.We let T_0 denote the tangent plane to X_0 at x^0, and T_1 the tangent plane to X_1 at x^1.

Theorem (2.2.13): (More Transversality Conditions)

Let $\boldsymbol{\alpha}^*(\cdot)$ and $\mathbf{x}^*(\cdot)$ solve the problem above, with

$$x^0 = \mathbf{x}^*(0), \quad x^1 = \mathbf{x}^*(\mathcal{T}^*).$$

Then there exists a function $\mathbf{p}^*(\cdot) : [0, \mathcal{T}^*] \to \mathbb{R}^n$, such that (ODE), (ADJ) and (M) hold for $0 \le t \le \mathcal{T}^*$. In addition,

(T) $$\begin{cases} \mathbf{p}^*(\mathcal{T}^*) \text{ is perpendicular to } T_1, \\ \mathbf{p}^*(0) \quad \text{is perpendicular to } T_0. \end{cases}$$

We call (T) the transversality conditions.

Remarks and Interpretations (2.2.14):

(i) If we have $T > 0$ fixed and

$$P[\boldsymbol{\alpha}(\cdot)] = \int_0^T r\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big)\, dt + g\big(x(T)\big),$$

then (T) says

$$\mathbf{p}^*(t) = \nabla g\big(\mathbf{x}^*(T)\big),$$

in agreement with our earlier form of the terminal / transversality condition.

(ii) Suppose that the surface X_1 is the graph $X_1 = \{x \mid g_k(x) = 0, k = 1, \dots, l\}$.Then (T) says that $\mathbf{p}^*(\mathcal{T}^*)$ belongs to the " orthogonal complement " of the subspace T_1. But orthogonal complement of T_1 is the span of $g_k(x^1)$ $(k = 1, \dots, l)$. Thus

$$\mathbf{p}^*(\mathcal{T}^*) = \sum_{k=1}^{l} \lambda_k \nabla g_k(x^1)$$

for some unknown constants $\lambda_1, \dots, \lambda_l$.

Example (2.2.15): (Distance between two Sets)

As a first and simple example, let

$$\text{(ODE)} \qquad \dot{\mathbf{x}}(t) = \boldsymbol{\alpha}(t)$$

for $A = S^1$, the unit sphere in $\mathbb{R}^2$: $a \in S^1$ if and only if $|a|^2 = a_1^2 + a_2^2 = 1$. In other words, we are considering only curves that move with unit speed.

We take

$$\text{(P)} \qquad \begin{aligned} P[\boldsymbol{\alpha}(\cdot)] &= -\int_0^{\mathcal{T}} |\dot{\mathbf{x}}(t)|\, dt = -\text{the length of the curve} \\ &= -\int_0^{\mathcal{T}} dt = -\text{time it takes to reach } X_1. \end{aligned}$$

We want to minimize the length of the curve and, as a check on our general theory, will prove that the minimum is of course a straight line.

Now we study how can using the maximum principle:

We have

$$H(x,p,a) = \mathbf{f}(x,a) \cdot p + r(x,a)$$
$$= a \cdot p - 1 = p_1 a_1 + p_2 a_2 - 1.$$

The adjoint dynamics equation (ADJ) says

$$\dot{\mathbf{p}}(t) = -\nabla_x H\big(\mathbf{x}(t), \mathbf{p}(t), \boldsymbol{\alpha}(t)\big) = 0,$$

and therefore

$$\mathbf{p}(t) \equiv \text{constant} = p^0 \neq 0.$$

The maximization principle (M) tells us that

$$H\big(\mathbf{x}(t), \mathbf{p}(t), \boldsymbol{\alpha}(t)\big) = \max_{a \in S^1} [-1 + p_1^0 a_1 + p_2^0 a_2].$$

The right hand side is maximized by $a^0 = \frac{p^0}{|p^0|}$, a unit vector that points in the same direction of p^0. Thus $\boldsymbol{\alpha}(\cdot) \equiv a^0$ is constant in time. According then to (ODE) we have $\dot{\mathbf{x}} = a^0$, and consequently $\mathbf{x}(\cdot)$ is a straight line.

Finally, the transversality conditions say that

(T) $$\mathbf{p}(0) \perp T_0, \mathbf{p}(t_1) \perp T_1.$$

In other words, $p^0 \perp T_0$ and $p^0 \perp T_1$; and this means that the tangent planes T_0 and T_1 are parallel.

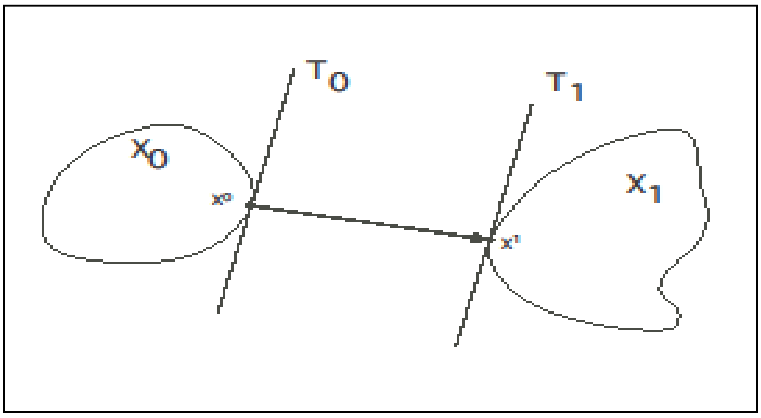

Figure (2.9)

Now all of this is pretty obvious from the picture, but it is reassuring that the general theory predicts the proper answer.

Example (2.2.16): (Commodity Trading)

Next is a simple model for the trading of a commodity, say wheat. We let T be the fixed length of trading period, and introduce the variables

$$x^1(t) = \text{money on hand at time } t$$

$$x^2(t) = \text{ amount of wheat owned at time } t$$

$$\alpha(t) = \text{rate of buying or selling of wheat}$$

$$q(t) = \text{price of wheat at time } t \text{ (known)}$$

$$\lambda = \text{cost of storing a unit amount of wheat for a unit of time.}$$

We suppose that the price of wheat $q(t)$ is known for the entire trading period $0 \leq t \leq T$ (although this is probably unrealistic in practice). We assume also that the rate of selling and buying is constrained:

$$|\alpha(t)| \leq M,$$

where $\alpha(t) > 0$ means buying wheat, and $\alpha(t) < 0$ means selling.

Our intention is to maximize our holdings at the end time T, namely the sum of the cash on hand and the value of the wheat we then own:

(P) $$P[\boldsymbol{\alpha}(\cdot)] = x^1(T) + q(T)x^2(T).$$

The evolution is

(ODE) $$\begin{cases} \dot{x}^1(t) = -\lambda x^2(t) - q(t)\alpha(t) \\ \dot{x}^2(t) = \alpha(t), \end{cases}$$

This is a nonautonomous (= time dependent) case, but it turns out that the Pontryagin Maximum Principle still applies.

Now we discuss how can using the maximum principle, or what is our optimal buying and selling strategy? First, we compute the Hamiltonian

$$H(x,p,t,a) = \mathbf{f} \cdot p + r = p_1(-\lambda x_2 - q(t)a) + p_2 a,$$

since $r \equiv 0$. The adjoint dynamics read

$$\text{(ADJ)} \qquad \begin{cases} \dot{p}^1 = 0 \\ \dot{p}^2 = \lambda p^1, \end{cases}$$

with the terminal condition

$$\text{(T)} \qquad \mathbf{p}(T) = \nabla g\big(x(T)\big).$$

In our case $g(x_1, x_2) = x_1 + q(T)x_2$, and hence

$$\text{(T)} \qquad \begin{cases} p^1(t) \equiv 1 \\ p^2(t) = \lambda(t - T) + q(T). \end{cases}$$

We then can solve for the costate:

$$\begin{cases} p^1 \equiv 1 \\ p^2 = \lambda(t - T) + q(T). \end{cases}$$

The maximization principle (M) tells us that

$$\text{(M)} \qquad \begin{aligned} H\big(\mathbf{x}(t), \mathbf{p}(t), \boldsymbol{\alpha}(t)\big) &= \max_{|a| \le M} \{p^1(t)(-\lambda x^2(t) - q(t)a) + p^2(t)a\} \\ &= -\lambda p^1(t)x^2(t) + \max_{|a| \le M} \big\{a\big(-q(t) + p^2(t)\big)\big\}. \end{aligned}$$

So

$$\alpha(t) = \begin{cases} M & \text{if } q(t) < p^2(t) \\ -M & \text{if } q(t) > p^2(t) \end{cases}$$

for $p^2(t) := \lambda(t - T) + q(T)$.

In some situations the amount of money on hand $x^1(\cdot)$ becomes negative for part of the time. The economic problem has a natural constraint $x_2 \ge 0$ (unless we can borrow with no interest charges) which we did not take into account in the mathematical model.

Now we return once again to our usual setting:

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big) \\ \mathbf{x}(0) = x^0, \end{cases}$$

$$\text{(P)} \qquad P[\boldsymbol{\alpha}(\cdot)] = \int_0^{\mathcal{T}} r\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big)\, dt$$

for $\mathcal{T} = \mathcal{T}[\boldsymbol{\alpha}(\cdot)]$, the first time that $x(\mathcal{T}) = x^1$. This is the fixed endpoint problem.

We introduce a new complication by asking that our dynamics $x(\cdot)$ must always remain within a given region $R \subset \mathbb{R}^n$. We will as above suppose that R has the explicit representation

$$R = \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$$

for a given function $g(\cdot) : \mathbb{R}^n \to \mathbb{R}$.

Definition (2.2.17):

It will be convenient to introduce the quantity

$$c(x, a) := \nabla g(x) \cdot \mathbf{f}(x, a).$$

Notice that

$$\text{if } x(t) \in \partial R \text{ for times } s_0 \leq t \leq s_1, \text{then } c\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big) \equiv 0 \quad (s_0 \leq t \leq s_1).$$

This is so since f is then tangent to ∂R, whereas ∇g is perpendicular.

Theorem (2.2.18): (Maximum Principle for State Constraints)

Let $\boldsymbol{\alpha}^*(\cdot), \mathbf{x}^*(\cdot)$ solve the control theory problem above. Suppose also that $\mathbf{x}^*(t) \in \partial R$ for $s_0 \leq t \leq s_1$.

Then there exists a costate function $\mathbf{p}^*(\cdot) : [s_0, s_1] \to \mathbb{R}^n$ such that (ODE) holds. There also exists $\lambda^*(\cdot) : [s_0, s_1] \to \mathbb{R}$ such that for times $s_0 \leq t \leq s_1$ we have

$$\text{(ADJ}'\text{)} \qquad \dot{\mathbf{p}}^*(t) = -\nabla_x H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big) + \lambda^*(t) \nabla_x c\big(\mathbf{x}^*(t), \boldsymbol{\alpha}^*(t)\big);$$

and

$$\text{(M}'\text{)} \qquad H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)\big) = \max_{a \in A} \{H(\mathbf{x}^*(t), \mathbf{p}^*(t), a) \mid c(\mathbf{x}^*(t), a) = 0\}.$$

To keep things simple, we have omitted some technical assumptions really needed for the Theorem to be valid.

Remarks and Interpretations (2.2.19):

(i) Let $A \subset \mathbb{R}^m$ be of this form:

$$A = \{a \in \mathbb{R}^m \mid g_1(a) \leq 0, \dots, g_s(a) \leq 0\}$$

for given functions $g_1, \ldots, g_s : \mathbb{R}^m \to \mathbb{R}$. In this case we can use Lagrange multipliers to deduce from (M') that

$$(\mathrm{M}'') \qquad \nabla_a H(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)) = \lambda^*(t)\nabla_a c(\mathbf{x}^*(t), \boldsymbol{\alpha}^*(t)) + \sum_{i=1}^{s} \mu_i^*(t)\nabla_a g_i(\mathbf{x}^*(t)).$$

The function $\lambda^*(\cdot)$ here is that appearing in (ADJ ').

If $\mathbf{x}^*(t)$ lies in the interior of R for say the time $0 \le t \le s_0$, then the ordinary Maximum Principle holds.

(ii) **Jump conditions**

In the situation above, we always have

$$\mathbf{p}^*(s_0 - 0) = \mathbf{p}^*(s_0 + 0),$$

where s_0 is a time that $\mathbf{x}^*$ hits ∂R. In other words, there is no jump in $\mathbf{p}^*$ when we hit the boundary of the constraint ∂R.

However,

$$\mathbf{p}^*(s_1 + 0) = \mathbf{p}^*(s_1 - 0) - \lambda^*(s_1)\nabla g(\mathbf{x}^*(s_1));$$

this says there is (possibly) a jump in $\mathbf{p}^*(\cdot)$ when we leave ∂R.

Example (2.2.20): Shortest Distance between two Points, Avoiding an Obstacle

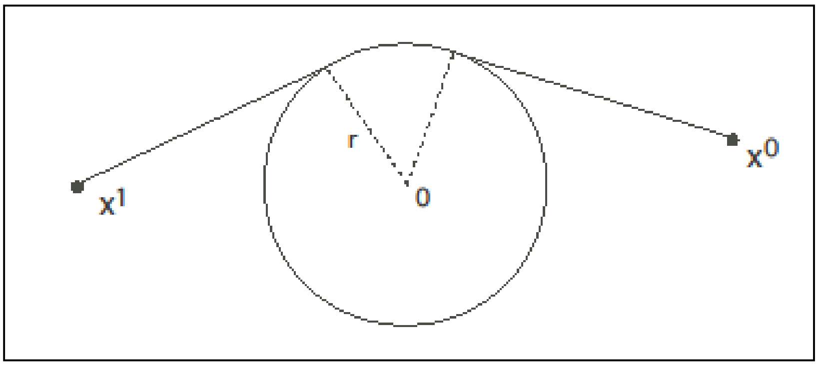

Figure (2.10)

What is the shortest path between two points that avoids the disk $B = B(0, r)$, as drawn?

Let us take

(ODE) $$\begin{cases}\dot{\mathbf{x}}(t) = \boldsymbol{\alpha}(t)\\ \mathbf{x}(0) = x^0\end{cases}$$

for $A = S^1$, with the payoff

(P) $$P[\boldsymbol{\alpha}(\cdot)] = -\int_0^{\mathcal{T}} |\dot{\mathbf{x}}|\, dt = -\text{ length of the curve } \mathbf{x}(\cdot).$$

We have

$$H(x, p, a) = \mathbf{f} \cdot p + r = p_1 a_1 + p_2 a_2 - 1.$$

Case (1): avoiding the obstacle

Assume $\mathbf{x}(t) \notin \partial B$ on some time interval. In this case, the usual Pontryagin Maximum Principle applies, and we deduce as before that

$$\dot{\mathbf{p}} = -\nabla_x H = 0.$$

Hence

(ADJ) $$\mathbf{p}(t) \equiv \text{constant} = p^0.$$

Condition (M) says

$$H\big(x(t), \mathbf{p}(t), \alpha(t)\big) = \max_{a \in S^1}(-1 + p_1^0 a_1 + p_2^0 a_2).$$

The maximum occurs for $\alpha = \frac{p^0}{|p^0|}$. Furthermore,

$$-1 + p_1^0 \alpha_1 + p_2^0 \alpha_2 \equiv 0;$$

and therefore $\alpha \cdot p^0 = 1$. This means that $|p^0| = 1$, and hence in fact $\alpha = p^0$. We have proved that the trajectory $\mathbf{x}(\cdot)$ is a straight line away from the obstacle.

Case (2): touching the obstacle

Suppose now $\mathbf{x}(t) \in \partial B$ for some time interval $s_0 \le t \le s_1$. Now we use the modified version of Maximum Principle, provided by Theorem (2.2.18).

First we must calculate $c(x, a) = \nabla g(x) \cdot \mathbf{f}(x, a)$. In our case,

$$R = \mathbb{R}^n - B = \{x \mid x_1^2 + x_2^2 \ge r^2\} = \{x \mid g \coloneqq r^2 - x_1^2 - x_2^2 \le 0\}.$$

Then $\nabla g = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix}$. Since $\mathbf{f} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, we have

$$c(x, a) = -2a_1x_1 - 2a_2x_2.$$

Now condition (ADJ ') implies

$$\dot{\mathbf{p}}(t) = -\nabla_x H + \lambda(t)\nabla_x c;$$

which is to say,

$$\begin{cases} \dot{p}^1 = 2\lambda\alpha^1 \\ \dot{p}^2 = -2\lambda\alpha^2. \end{cases} \tag{2.6}$$

Next, we employ the maximization principle (M'). We need to maximize

$$H(\mathbf{x}(t), \mathbf{p}(t), a)$$

subject to the requirements that $c(\mathbf{x}(t), a) = 0$ and $g_1(a) = a_1^2 + a_2^2 - 1 = 0$, since $A = \{a \in \mathbb{R}^2 \mid a_1^2 + a_2^2 = 1\}$. According to (M'') we must solve

$$\nabla_a H = \lambda(t)\nabla_a c + \mu(t)\nabla_a g_1;$$

that is,

$$\begin{cases} p^1 = \lambda(-2x^1) + \mu 2\alpha^1 \\ p^2 = \lambda(-2x^2) + \mu 2\alpha^2. \end{cases}$$

We can combine these identities to eliminate μ. Since we also know that $\mathbf{x}(t) \in \partial B$, we have $(x^1)^2 + (x^2)^2 = r^2$; and also $\boldsymbol{\alpha} = (\alpha^1, \alpha^2)^T$ is tangent to ∂B. Using these facts, we find after some calculations that

$$\lambda = \frac{p^2\alpha^1 - p^1\alpha^2}{2r}. \tag{2.7}$$

But we also know

$$p^2(\alpha^1)^2 + (\alpha^2)^2 = 1 \tag{2.8}$$

and

$$H \equiv 0 = -1 + p^1\alpha^1 + p^2\alpha^2;$$

hence

$$p^1\alpha^1 + p^2\alpha^2 \equiv 1. \tag{2.9}$$

We now have the five equations (2.6)-(2.9) for the five unknown functions $p^1, p^2, \alpha^1, \alpha^2, \lambda$ that depend on t. We introduce the angle θ, as illustrated, and note that $\frac{d}{d\theta} = r\frac{d}{dt}$. A calculation then confirms that the solutions are

$$\begin{cases} \alpha^1(\theta) = -\sin\theta \\ \alpha^2(\theta) = \cos\theta, \end{cases}$$

$$\lambda = -\frac{k+\theta}{2r},$$

and

$$\begin{cases} p^1(\theta) = k\cos\theta - \sin\theta + \theta\cos\theta \\ p^2(\theta) = k\sin\theta + \cos\theta + \theta\sin\theta \end{cases}$$

for some constant k.

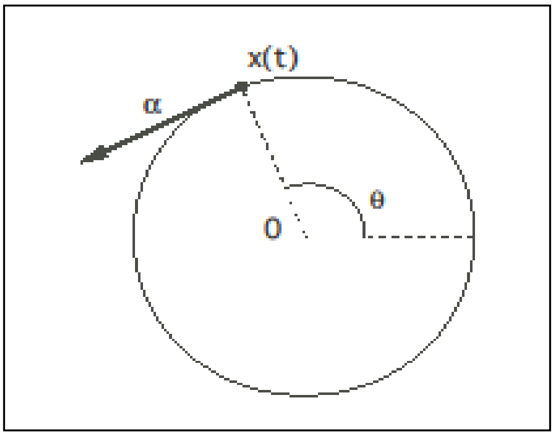

Figure (2.11)

Case (3): approaching and leaving the obstacle

In general, we must piece together the results from Case (1) and Case (2). So suppose now $\mathbf{x}(t) \in R = \mathbb{R}^2 - B$ for $0 \le t < s_0$ and $\mathbf{x}(t) \in \partial B$ for $s_0 \le t \le s_1$.

We have shown that for times $0 \le t < s_0$, the trajectory $\mathbf{x}(\cdot)$ is a straight line. For this case we have shown already that $p = \alpha$ and therefore

$$\begin{cases} p^1 \equiv -\cos\phi_0 \\ p^2 \equiv \sin\phi_0, \end{cases}$$

for the angle ϕ_0 as shown in the picture.

By the jump conditions, $\mathbf{p}(\cdot)$ is continuous when $\mathbf{x}(\cdot)$ hits ∂B at the time s_0, meaning in this case that

$$\begin{cases} k\cos\theta_0 - \sin\theta_0 + \theta_0\cos\theta_0 = -\cos\phi_0 \\ k\sin\theta_0 + \cos\theta_0 + \theta_0\sin\theta_0 = \sin\phi_0. \end{cases}$$

These identities hold if and only if

$$\begin{cases} k = -\theta_0 \\ \theta_0 + \phi_0 = \dfrac{\pi}{2}. \end{cases}$$

The second equality says that the optimal trajectory is tangent to the disk B when it hits ∂B.

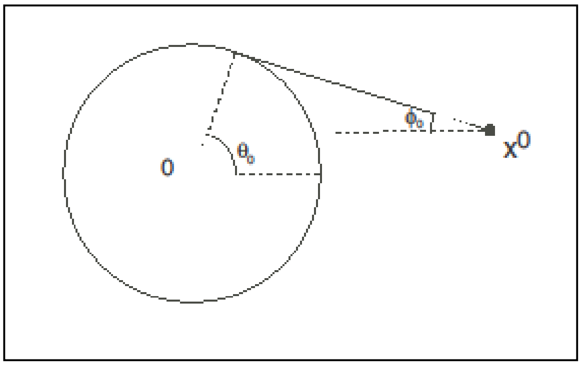

Figure (2.12)

We turn next to the trajectory as it leaves ∂B: see the next picture. We then have

$$\begin{cases} p^1(\theta_1^-) = -\theta_0\cos\theta_1 - \sin\theta_1 + \theta_1\cos\theta_1 \\ p^2(\theta_1^-) = -\theta_0\sin\theta_1 + \cos\theta_1 + \theta_1\sin\theta_1\,. \end{cases}$$

Now our formulas above for λ and k imply $\lambda(\theta_1) = \frac{\theta_0 - \theta_1}{2r}$. The jump conditions give

$$p(\theta_1^+) = p(\theta_1^-) - \lambda(\theta_1)\nabla g\big(\mathbf{x}(\theta_1)\big)$$

for $g(x) = r^2 - x_1^2 - x_2^2$. Then

$$\lambda(\theta_1)\nabla g\big(\mathbf{x}(\theta_1)\big) = (\theta_1 - \theta_0)\begin{pmatrix}\cos\theta_1 \\ \sin\theta_1\end{pmatrix}.$$

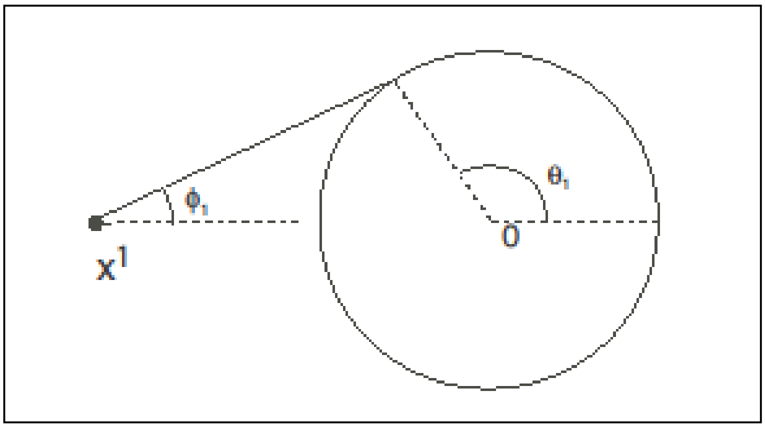

Figure (2.13)

Therefore

$$\begin{cases} p^1(\theta_1^+) = -\sin\theta_1 \\ p^2(\theta_1^+) = \cos\theta_1, \end{cases}$$

and so the trajectory is tangent to ∂B. If we apply usual Maximum Principle after $x(\cdot)$ leaves B, we find

$$\begin{cases} p^1 \equiv \text{constant} = -\cos\phi_1 \\ p^2 \equiv \text{constant} = -\sin\phi_1\,. \end{cases}$$

Thus

$$\begin{cases} -\cos\phi_1 = -\sin\theta_1 \\ -\sin\phi_1 = \cos\theta_1, \end{cases}$$

and so $\phi_1 + \theta_1 = \pi$.

We have carried out elaborate calculations to derive some pretty obvious conclusions in this example. It is best to think of this as a confirmation in a simple case of Theorem (2.2.18), which applies in far more complicated situations.

Now we turn to a simple model for ordering and storing items in a warehouse. Let the time period $T > 0$ be given, and introduce the variables

$$x(t) = \text{amount of inventory at time } t$$

$$\alpha(t) = \text{rate of ordering from manufacturers}, \alpha \geq 0,$$

$$d(t) = \text{customer demand (known)}$$

$$\gamma = \text{cost of ordering 1 unit}$$

$$\beta = \text{cost of storing 1 unit.}$$

Our goal is to fill all customer orders shipped from our warehouse, while keeping our storage and ordering costs at a minimum. Hence the payoff to be maximized is

$$\text{(P)} \qquad P[\alpha(\cdot)] = -\int_0^T \gamma\alpha(t) + \beta x(t)\, dt.$$

We have $A = [0, \infty)$ and the constraint that $x(t) \geq 0$. The dynamics are

$$\text{(ODE)} \qquad \begin{cases} \dot{x}(t) = \alpha(t) - d(t) \\ x(0) = x^0 > 0. \end{cases}$$

Let us just guess the optimal control strategy: we should at first not order anything $(\alpha = 0)$ and let the inventory in our warehouse fall off to zero as we fill demands; thereafter we should order just enough to meet our demands $(\alpha = d)$.

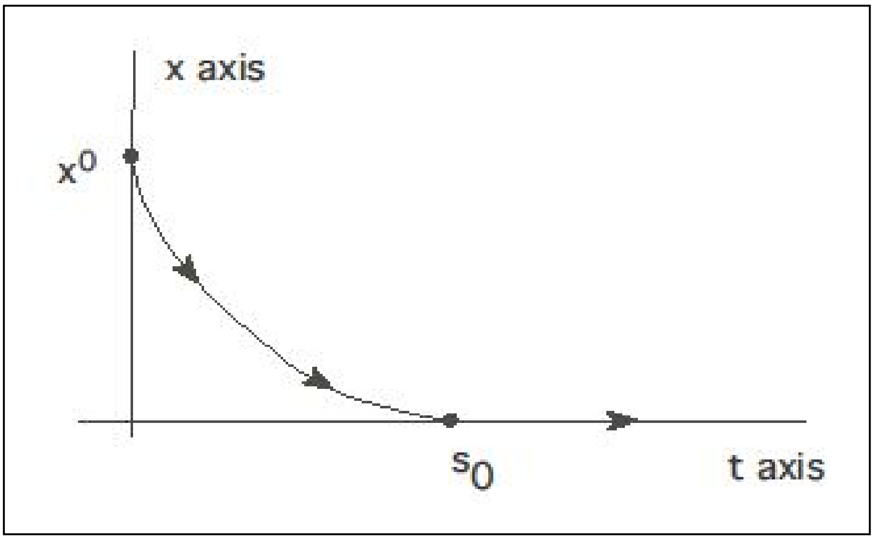

Figure (2.14)

We will prove this guess is right, using the Maximum Principle. Assume first that $x(t) > 0$ on some interval $[0, s_0]$. We then have

$$H(x, p, a, t) = \big(a - d(t)\big)p - \gamma a - \beta x;$$

and (ADJ) says $\dot{p} = \nabla_x H = \beta$. Condition (M) implies

$$H(x(t), p(t), t) = \max_{a\geq 0}\{-\gamma a - \beta x(t) + p(t)(a - d(t))\}$$

$$= -\beta x(t) - p(t)d(t) + \max_{a\geq 0}\{a(p(t) - \gamma)\}.$$

Thus

$$\alpha(t) = \begin{cases} 0 & \text{if } p(t) \leq \gamma \\ +\infty & \text{if } p(t) > \gamma. \end{cases}$$

If $\alpha(t) \equiv +\infty$ on some interval, then $P[\alpha(\cdot)] = -\infty$, which is impossible, because there exists a control with finite payoff. So it follows that $\alpha(\cdot) \equiv 0$ on $[0, s_0]$: we place no orders.

According to (ODE), we have

$$\begin{cases} \dot{x}(t) = -d(t) \quad (0 \leq t \leq s_0) \\ x(0) = x^0 \end{cases}$$

Thus s_0 is first time the inventory hits 0. Now since $x(t) = x^0 - \int_0^t d(s)\,ds$, we have $x(s_0) = 0$. That is, $\int_0^{s_0} d(s)\,ds = x^0$ and we have hit the constraint. Now use Pontryagin Maximum Principle with state constraint for times $t \geq s_0$

$$R = \{x \geq 0\} = \{g(x) := -x \leq 0\}$$

and

$$c(x, a, t) = \nabla g(x) \cdot f(x, a, t) = (-1)(a - d(t)) = d(t) - a.$$

We have

(M) $$H(x(t), p(t), \alpha(t), t) = \max_{a\geq 0}\{H(x(t), p(t), a, t) \mid c(x(t), a, t) = 0\}.$$

But $c(x(t), \alpha(t), t) = 0$ if and only if $\alpha(t) = d(t)$. Then (ODE) reads

$$\dot{x}(t) = \alpha(t) - d(t) = 0$$

and so $x(t) = 0$ for all times $t \geq s_0$.

We have confirmed that our guess for the optimal strategy was right.

Chapter (3)

Dynamic Programming and Game Theorem

Section (3.1): Derivation of Bellman's PDE and Dynamic Programming

We begin with some mathematical wisdom: " It is sometimes easier to solve a problem by embedding it within a larger class of problems and then solving the larger class all at once ".

Example (3.1.1): (A Calculus Example)

Suppose we wish to calculate the value of the integral

$$\int_0^\infty \frac{\sin x}{x}\, dx.$$

This is pretty hard to do directly, so let us as follows add a parameter α into the integral:

$$I(\alpha) := \int_0^\infty e^{-\alpha x} \frac{\sin x}{x}\, dx.$$

We compute

$$I'(\alpha) = \int_0^\infty (-x) e^{-\alpha x} \frac{\sin x}{x}\, dx = -\int_0^\infty \sin x e^{-\alpha x}\, dx = \frac{1}{\alpha^2 + 1},$$

where we integrated by parts twice to find the last equality. Consequently

$$I(\alpha) = -\arctan \alpha + C,$$

and we must compute the constant C. To do so, observe that

$$0 = I(\infty) = -\arctan(\infty) + C = -\frac{\pi}{2} + C,$$

and so $C = \frac{\pi}{2}$. Hence $I(\alpha) = -\arctan \alpha + \frac{\pi}{2}$, and consequently

$$\int_0^\infty e^{-\alpha x} \frac{\sin x}{x}\, dx = I(0) + \frac{\pi}{2}.$$

We want to adapt some version of this idea to the vastly more complicated setting of control theory. For this, fix a terminal time $T > 0$ and then look at the controlled dynamics

(ODE) $$\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \boldsymbol{\alpha}(s)) & (0 < s < T) \\ \mathbf{x}(0) = x^0, \end{cases}$$

with the associated payoff functional

(P) $$P[\boldsymbol{\alpha}(\cdot)] = \int_0^T r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T)).$$

We embed this into a larger family of similar problems, by varying the starting times and starting points:

$$\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \boldsymbol{\alpha}(s)) & (t < s < T) \\ \mathbf{x}(t) = x, \end{cases} \tag{3.1}$$

with

$$P_{x,t}[\boldsymbol{\alpha}(\cdot)] = \int_t^T r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T)). \tag{3.2}$$

Consider the above problems for all choices of starting times $0 \leq t \leq T$ and all initial points $x \in \mathbb{R}^n$.

Definition (3.1.2):

For $x \in \mathbb{R}^n, 0 \leq t \leq T$, define the value function $v(x,t)$ to be the greatest payoff possible if we start at $x \in \mathbb{R}^n$ at time t. In other words,

$$v(x,t) := \sup_{\boldsymbol{\alpha}(\cdot) \in \mathcal{A}} P_{x,t}[\boldsymbol{\alpha}(\cdot)] \quad (x \in \mathbb{R}^n, 0 \leq t \leq T). \tag{3.3}$$

Notice then that

$$v(x,T) = g(x) \quad (x \in \mathbb{R}^n). \tag{3.4}$$

Now we discuss the derivation of Hamilton-Jacobi-Bellman equation:

Our first task is to show that the value function v satisfies a certain nonlinear partial differential equation.

Our derivation will be based upon the reasonable principle that " it's better to be smart from the beginning, than to be stupid for a time and then become smart ". We want to convert this philosophy of life into mathematics.

To simplify, we hereafter suppose that the set A of control parameter values is compact.

Theorem (3.1.3): (Hamilton-Jacobi-Bellman Equation)

Assume that the value function v is a C^1 function of the variables (x, t). Then v solves the nonlinear partial differential equation

(HJB) $$v_t(x,t) + \max_{a \in A}\{\mathbf{f}(x,a) \cdot \nabla_x v(x,t) + r(x,a)\} = 0 \quad (x \in \mathbb{R}^n, 0 \leq t \leq T).$$

with the terminal condition

$$v(x,T) = g(x) \quad (x \in \mathbb{R}^n).$$

Remark (3.1.4):

We call (HJB) the Hamilton-Jacobi-Bellman equation, and can rewrite it as

(HJB) $$v_t(x,t) + H(x, \nabla_x v) = 0 \quad (x \in \mathbb{R}^n, 0 \leq t \leq T),$$

for the partial differential equations Hamiltonian

$$H(x,p) = \max_{a \in A} H(x,p,a) = \max_{a \in A}\{\mathbf{f}(x,a) \cdot p + r(x,a)\}$$

where $x, p \in \mathbb{R}^n$.

Proof:

1. Let $x \in \mathbb{R}^n, 0 \leq t < T$ and let $h > 0$ be given. As always

$$\mathcal{A} = \{\boldsymbol{\alpha}(\cdot) : [0,\infty) \to A \text{ measurable}\}.$$

Pick any parameter $a \in A$ and use the constant control

$$\boldsymbol{\alpha}(\cdot) \equiv a$$

for times $t \leq s \leq t + h$. The dynamics then arrive at the point $\mathbf{x}(t+h)$, where $t + h < T$. Suppose now a time $t + h$, we switch to an optimal control and use it for the remaining times $t + h \leq s \leq T$.

What is the payoff of this procedure? Now for times $t \leq s \leq t + h$, we have

$$\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), a) \\ \mathbf{x}(t) = x. \end{cases}$$

The payoff for this time period is $\int_t^{t+h} r(\mathbf{x}(s), a)\, ds$. Furthermore, the payoff incurred from time $t+h$ to T is $v(\mathbf{x}(t+h), t+h)$, according to the definition of the payoff function v. Hence the total payoff is

$$\int_t^{t+h} r(\mathbf{x}(s), a)\, ds + v(\mathbf{x}(x+h), t+h).$$

But the greatest possible payoff if we start from (x, t) is $v(x, t)$. Therefore

$$v(x,t) \geq \int_t^{t+h} r(\mathbf{x}(s), a)\, ds + v(\mathbf{x}(x+h), t+h). \tag{3.5}$$

2. We now want to convert this inequality into a differential form. So we rearrange (3.5) and divide by $h > 0$:

$$\frac{v(\mathbf{x}(x+h), t+h) - v(x,t)}{h} + \frac{1}{h}\int_t^{t+h} r(\mathbf{x}(s), a)\, ds \leq 0.$$

Let $h \to 0$:

$$v_t(x,t) + \nabla_x v(\mathbf{x}(t), t) \cdot \dot{\mathbf{x}}(t) + r(\mathbf{x}(t), a) \leq 0.$$

But $\mathbf{x}(\cdot)$ solves the ODE

$$\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), a) & (t \leq s \leq t+h) \\ \mathbf{x}(t) = x, \end{cases}$$

Employ this above, to discover:

$$v_t(x,t) + \mathbf{f}(x,a) \cdot \nabla_x v(x,t) + r(x,t) \leq 0.$$

This inequality holds for all control parameters $a \in A$, and consequently

$$\max_{a \in A}\{v_t(x,t) + \mathbf{f}(x,a) \cdot \nabla_x v(x,t) + r(x,a)\} \leq 0. \tag{3.6}$$

3. We next demonstrate that in fact the maximum above equals zero. To see this, suppose $\boldsymbol{\alpha}^*(\cdot), \mathbf{x}^*(\cdot)$ were optimal for the problem above. Let us utilize the optimal control $\boldsymbol{\alpha}^*(\cdot)$ for $t \leq s \leq t+h$. The payoff is

$$\int_t^{t+h} r\big(\mathbf{x}^*(s), \boldsymbol{\alpha}^*(s)\big)\, ds$$

and the remaining payoff is $v(\mathbf{x}^*(t+h), t+h)$. Consequently, the total payoff is

$$\int_t^{t+h} r(\mathbf{x}^*(s), \boldsymbol{\alpha}^*(s))\, ds + v(\mathbf{x}^*(t+h), t+h) = v(x,t).$$

Rearrange and divide by h:

$$\frac{v(\mathbf{x}^*(t+h), t+h) - v(x,t)}{h} + \frac{1}{h}\int_t^{t+h} r(\mathbf{x}^*(s), \boldsymbol{\alpha}^*(s))\, ds = 0.$$

Let $h \to 0$ and suppose $\boldsymbol{\alpha}^*(t) = a^* \in A$. Then

$$v_t(x,t) + \nabla_x v(x,t) \cdot \underbrace{\dot{\mathbf{x}}^*(t)}_{\mathbf{f}(x,a^*)} + r(x,a^*) = 0;$$

and therefore

$$v_t(x,t) + \mathbf{f}(x,a^*) \cdot \nabla_x v(x,t) + r(x,a^*) = 0$$

for some parameter value $a^* \in A$. This proves (HJB).

Now we study the Dynamic Programming Method:

Here is how to use the dynamic programming method to design optimal controls:

Step 1:

Solve the Hamilton-Jacobi-Bellman equation, and thereby compute the value function v.

Step 2:

Use the value function v and the Hamilton-Jacobi-Bellman PDE to design an optimal feedback control $\boldsymbol{\alpha}^*(\cdot)$, as follows. Define for each point $x \in \mathbb{R}^n$ and each time $0 \le t \le T$,

$$\boldsymbol{\alpha}(x,t) = a \in A$$

to be a parameter value where the maximum in (HJB) is attained. In other words, we select $\boldsymbol{\alpha}(x,t)$ so that

$$v_t(x,t) + \mathbf{f}(x,\boldsymbol{\alpha})(x,t) \cdot \nabla_x v(x,t) + r(x, \boldsymbol{\alpha}(x,t)) = 0.$$

Next we solve the following ODE, assuming $\boldsymbol{\alpha}(\cdot,t)$ is sufficiently regular to let us do so:

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}\big(\mathbf{x}^*(s), \boldsymbol{\alpha}(\mathbf{x}^*(\mathrm{s}), \mathrm{s})\big) & (t \le s \le T) \\ \mathbf{x}(t) = x. \end{cases}$$

Finally, define the feedback control

$$\boldsymbol{\alpha}^*(s) := \boldsymbol{\alpha}(\mathbf{x}^*(s), s). \tag{3.7}$$

In summary, we design the optimal control this way: If the state of system is x at time t, use the control which at time t takes on the parameter value $a \in A$ such that the minimum in (HJB) is obtained.

We demonstrate next that this construction does indeed provide us with an optimal control.

Theorem (3.1.5): (Verification of Optimality)

The control $\boldsymbol{\alpha}^*(\cdot)$ defined by the construction (3.7) is optimal.

Proof:

We have

$$P_{x,t}[\boldsymbol{\alpha}^*(\cdot)] = \int_t^T r\big(\mathbf{x}^*(s), \boldsymbol{\alpha}^*(s)\big)\, ds + g\big(\mathbf{x}^*(T)\big).$$

Furthermore according to the definition (3.7) of $\alpha(\cdot)$:

$$\begin{aligned} P_{x,t}[\boldsymbol{\alpha}^*(\cdot)] &= \int_t^T \Big(-v_t(\mathbf{x}^*(s), s) - \mathbf{f}\big(\mathbf{x}^*(s), \boldsymbol{\alpha}^*(s)\big) \cdot \nabla_x v(\mathbf{x}^*(s), s)\Big)\, ds + g\big(\mathbf{x}^*(T)\big) \\ &= -\int_t^T v_t(\mathbf{x}^*(s), s) + \nabla_x v(\mathbf{x}^*(s), s) \cdot \dot{\mathbf{x}}^*(s)\, ds + g\big(\mathbf{x}^*(T)\big) \\ &= -\int_t^T \frac{d}{ds} v(\mathbf{x}^*(s), s)\, ds + g\big(\mathbf{x}^*(T)\big) \\ &= -v(\mathbf{x}^*(T), T) + v(\mathbf{x}^*(t), t) + g\big(\mathbf{x}^*(T)\big) \\ &= -g\big(\mathbf{x}^*(T)\big) + v(\mathbf{x}^*(t), t) + g\big(\mathbf{x}^*(T)\big) \\ &= v(x,t) = \sup_{\boldsymbol{\alpha}(\cdot) \in \mathcal{A}} P_{x,t}\, [\boldsymbol{\alpha}(\cdot)]. \end{aligned}$$

That is,

$$P_{x,t}[\boldsymbol{\alpha}^*(\cdot)] = \sup_{\boldsymbol{\alpha}(\cdot)\in\mathcal{A}} P_{x,t}\,[\boldsymbol{\alpha}(\cdot)];$$

and so $\boldsymbol{\alpha}^*(\cdot)$ is optimal, as asserted.

Example (3.1.6): (Dynamics with Three Velocities)

Let us begin with a fairly easy problem:

$$\text{(ODE)} \qquad \begin{cases} \dot{x}(s) = \alpha(s) & (0 \le t \le s \le 1) \\ x(t) = x \end{cases}$$

where our set of control parameters is

$$A = \{-1,0,1\}.$$

We want to minimize

$$\int_t^1 |x(s)|\, ds,$$

and so take for our payoff functional

$$\text{(P)} \qquad P_{x,t}[\alpha(\cdot)] = \int_t^1 |x(s)|\, ds.$$

As our first illustration of dynamic programming, we will compute the value function $v(x,t)$ and confirm that it does indeed solve the appropriate Hamilton-Jacobi-Bellman equation. To do this, we first introduce the three regions:

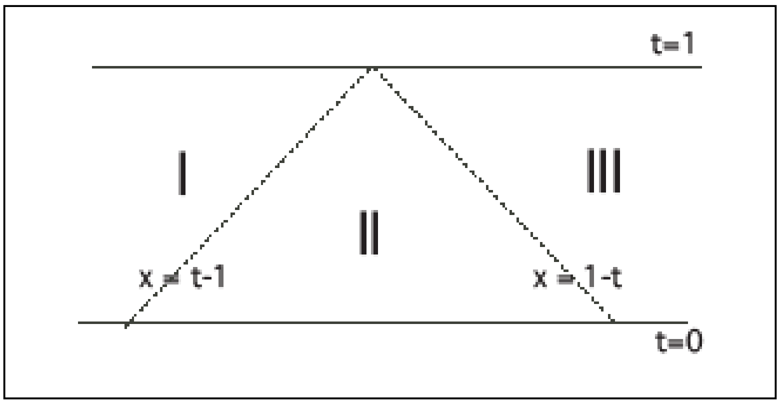

Figure (3.1)

- Region $I = \{(x,t) \mid x < t-1, 0 \le t \le 1\}$.

- Region $II = \{(x,t) \mid t-1 < x < 1-t, 0 \le t \le 1\}$.
- Region $III = \{(x,t) \mid x > 1-t, 0 \le t \le 1\}$.

We will consider the three cases as to which region the initial data (x,t) lie within.

Region I:

In this region, we should take $\alpha \equiv 1$, in which case we can similarly compute $v(x,t) = -\left(\frac{1-t}{2}\right)(-2x+t+1)$.

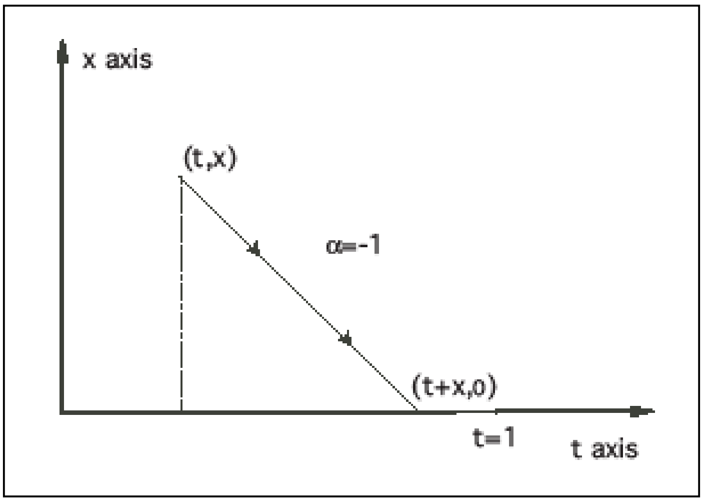

Figure (3.2) Optimal path in Region II

Region II:

In this region we take $\alpha \equiv \pm 1$, until we hit the origin, after which we take $\alpha \equiv 0$. We therefore calculate that $v(x,t) = -\frac{x^2}{2}$ in this region.

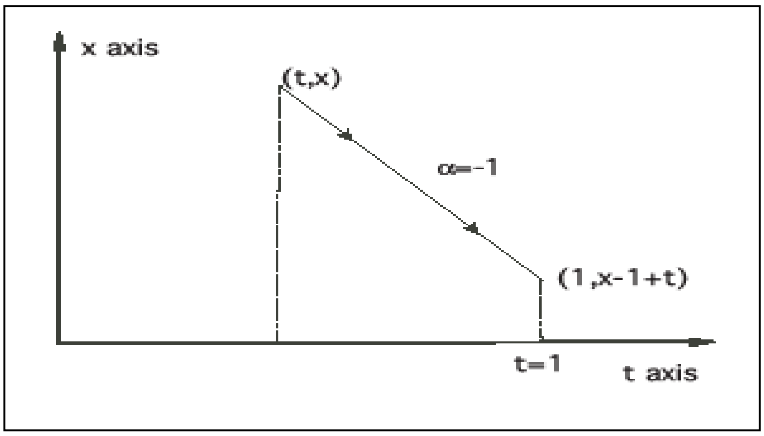

Figure (3.3) Optimal path in Region III

Region III:

In this case we should take $\alpha \equiv -1$, to steer as close to the origin 0 as quickly as possible. (See the next picture). Then

$$v(x,t) = -\text{ area under path taken} = (1-t)\frac{1}{2}(x+x+t+1)$$
$$= -\frac{(1-t)}{2}(2x+t+1).$$

Now the Hamilton-Jacobi-Bellman equation for our problem reads

$$v_t + \max_{a \in A}\{f \cdot \nabla_x v + r\} = 0 \tag{3.8}$$

for $f = a,\ r = -|x|$. We rewrite this as

$$v_t + \max_{a=\pm 1,0}\{a v_x\} = -|x| = 0;$$

and so

(HJB) $$v_t + |v_x| - |x| = 0.$$

We must check that the value function v, defined explicitly above in Regions I-III, does in fact solve this PDE, with the terminal condition that $v(x,1) = g(x) = 0$.

Now in Region II $v = -\frac{x^2}{2},\ v_t = 0,\ v_x = -x$. Hence

$$v_t + |v_x| + |x| = 0 + |-x| - |x| = 0 \text{ in Region II,}$$

in accordance with (HJB).

In Region III we have

$$v(x,t) = -\frac{(1-t)}{2}(2x+t-1);$$

and therefore

$$v_t = \frac{1}{2}(2x+t-1) - \frac{(1-t)}{2} = x-1+t, \quad v_x = t-1, \quad |t-1| = 1-t \geq 0.$$

Hence

$$v_t + |v_x| - |x| = x-1+t+|t-1| - |x| = 0 \text{ in Region III,}$$

because $x > 0$ there.

Likewise the Hamilton-Jacobi-Bellman PDE holds in Region I.

Remarks (3.1.7):

(i) In the example, v is not continuously differentiable on the borderline between Regions II and I or III.

(ii) In general, it may not be possible actually to find the optimal feedback control. For example, reconsider the above problem, but now with $A = \{-1,1\}$.We still have $\alpha = \operatorname{sgn}(v_x)$, but there is no optimal control in Region II.

Example (3.1.8): (Rocket Railroad Car)

Recall that the equations of motion in this model are

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \alpha, \quad |\alpha| \leq 1$$

and

$$P[\alpha(\cdot)] = -\text{time to reach}(0,0) = -\int_0^\tau 1 \, dt = -\tau.$$

To use the method of dynamic programming, we define $v(x_1, x_2)$ to be minus the least time it takes to get to the origin $(0,0)$, given we start at the point (x_1, x_2).

What is the Hamilton-Jacobi-Bellman equation? Note v does not depend on t, and so we have

$$\sup_{a \in A} \{\mathbf{f} \cdot \nabla_x v + r\} = 0$$

for

$$A = [-1,1], \quad \mathbf{f} = \begin{pmatrix} x_2 \\ a \end{pmatrix}, \quad r = -1$$

Hence our PDE reads

$$\max_{|a| \leq 1} \{x_2 v_{x_1} + a v_{x_2} - 1\} = 0;$$

and consequently

$$\text{(HJB)} \qquad \begin{cases} x_2 v_{x_1} + |v_{x_2}| - 1 = 0 & \text{in } \mathbb{R}^n \\ v(0,0) = 0. \end{cases}$$

We now confirm that v really satisfies (HJB). For this, define the regions

$$I := \left\{(x_1, x_2) \mid x_1 \geq -\frac{1}{2}x_2|x_2|\right\} \quad \text{and} \quad II := \left\{(x_1, x_2) \mid x_1 \leq -\frac{1}{2}x_2|x_2|\right\}.$$

A direct computation, the details of which we omit, reveals that

$$v(x) = \begin{cases} -x_2 - 2\left(x_1 + \frac{1}{2}x_2^2\right)^{\frac{1}{2}} & \text{in Region I} \\ x_2 - 2\left(-x_1 + \frac{1}{2}x_2^2\right)^{\frac{1}{2}} & \text{in Region II.} \end{cases}$$

In Region I we compute

$$v_{x_2} = -1 - \left(x_1 + \frac{x_2^2}{2}\right)^{-\frac{1}{2}} x_2,$$

$$v_{x_1} = -\left(x_1 + \frac{x_2^2}{2}\right)^{-\frac{1}{2}} x_2;$$

and therefore

$$x_2 v_{x_1} + |v_{x_2}| - 1 = -x_2\left(x_1 + \frac{x_2^2}{2}\right)^{-\frac{1}{2}} + \left[1 + x_2\left(x_1 + \frac{x_2^2}{2}\right)^{-\frac{1}{2}}\right] - 1 = 0.$$

This confirms that our (HJB) equation holds in Region I, and a similar calculation holds in Region II.

Now to calculate the Optimal control, since

$$\max_{|\alpha| \leq 1}\{x_2 v_{x_1} + a v_{x_2} + 1\} = 0,$$

the optimal control is

$$\alpha = \operatorname{sgn} v_{x_2}.$$

Example (3.1.9): (General Linear-Quadratic Regulator)

For this important problem, we are given matrices $M, B, D \in \mathbb{M}^{n\times n}$, $N \in \mathbb{M}^{n\times m}$, $C \in \mathbb{M}^{m\times m}$; and assume B, C, D are symmetric and nonnegative definite, and C is invertible.

We take the linear dynamics

(ODE) $$\begin{cases} \dot{\mathbf{x}}(s) = M\mathbf{x}(s) + N\boldsymbol{\alpha}(s) & (t \le s \le T) \\ \mathbf{x}(t) = x, \end{cases}$$

for which we want to minimize the quadratic cost functional

$$\int_t^T x(s)^T B\mathbf{x}(s) + \boldsymbol{\alpha}(s)^T C\boldsymbol{\alpha}(s)\, ds + \mathbf{x}(T)^T D\mathbf{x}(T).$$

So we must maximize the payoff

(P) $$P_{x,t}[\boldsymbol{\alpha}(\cdot)] = -\int_t^T x(s)^T B\mathbf{x}(s) + \boldsymbol{\alpha}(s)^T C\boldsymbol{\alpha}(s)\, ds - \mathbf{x}(T)^T D\mathbf{x}(T).$$

The control values are unconstrained, meaning that the control parameter values can range over all of $A = \mathbb{R}^m$.

We will solve by dynamic programming the problem of designing an optimal control. To carry out this plan, we first compute the Hamilton-Jacobi-Bellman equation

$$v_t + \max_{a\in\mathbb{R}^m}\{\mathbf{f}\cdot\nabla_x v + r\} = 0,$$

where

$$\begin{cases} \mathbf{f} = Mx + Na \\ r = -x^T Bx - a^T - a^T Ca \\ g = -x^T Dx. \end{cases}$$

Rewrite:

(HJB) $$v_t + \max_{a\in\mathbb{R}^m}\{(\nabla v)^T Na - a^T Ca\} + (\nabla v)^T Mx - x^T Bx = 0.$$

We also have the terminal condition

$$v(x, T) = -x^T Dx$$

Now for what value of the control parameter a is the minimum attained? To understand this, we define $Q(a) := (\nabla v)^T Na - a^T Ca$, and determine where Q has a minimum by computing the partial derivatives Q_{a_j} for $j = 1, \dots, m$ and setting them equal to 0. This gives the identitites

$$Q_{a_j} = \sum_{i=1}^{n} v_{x_i} n_{ij} - 2a_i c_{ij} = 0.$$

Therefore $(\nabla v)^T N = 2a^T C$, and then $2C^T a = N^T v$. But $C^T = C$. Therefore

$$a = \frac{1}{2} C^{-1} N^T \nabla_x v.$$

This is the formula for the optimal feedback control: It will be very useful once we compute the value function v.

Now to find the value function, we insert our formula $a = \frac{1}{2} C^{-1} N^T \nabla v$ into (HJB), and this PDE then reads

$$\text{(HJB)} \qquad \begin{cases} v_t + \dfrac{1}{4} (\nabla v)^T N C^{-1} N^T \nabla v + (\nabla v)^T M x - x^T B x = 0 \\ v(x, T) = -x^T D x. \end{cases}$$

Our next move is to guess the form of the solution, namely

$$v(x, t) = x^T K(t) x,$$

provided the symmetric $n \times n$-matrix valued function $K(\cdot)$ is properly selected. Will this guess work?

Now, since $-x^T K(T) x = -v(x, T) = x^T D x$, we must have the terminal condition that

$$K(T) = -D.$$

Next, compute that

$$v_t = x^T \dot{K}(t) x, \qquad \nabla_x v = 2K(t) x.$$

We insert our guess $v = x^T K(t) x$ into (HJB), and discover that

$$x^T \{ \dot{K}(t) + K(t) N C^{-1} N^T K(t) + 2K(t) M - B \} x = 0.$$

Look at the expression

$$\begin{aligned} 2x^T KMx &= x^T KMx + [x^T KMx]^T \\ &= x^T KMx + x^T M^T Kx. \end{aligned}$$

Then

$$x^T\{\dot{K} + KNC^{-1}N^T K + KM + M^T K - B\}x = 0.$$

This identity will hold if $K(\cdot)$ satisfies the matrix Riccati equation

$$\text{(R)} \qquad \begin{cases} \dot{K}(t) + K(t)NC^{-1}N^T K(t) + K(t)M + M^T K(t) - B = 0 \quad (0 \le t < T) \\ K(T) = -D \end{cases}$$

In summary, if we can solve the Riccati equation (R), we can construct an optimal feedback control

$$\boldsymbol{\alpha}^*(t) = C^{-1}N^T K(t)\mathbf{x}(t)$$

Furthermore, (R) in fact does have a solution, as explained for instance in the book of Fleming-Rishel [4].

Now we discuss Dynamic Programming and the Pontryagin Maximum Principle:

To illustrate the Method of Characteristics, assume $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and consider this initial-value problem for the Hamilton-Jacobi equation:

$$\text{(HJ)} \qquad \begin{cases} u_t(x,t) + H\big(x, \nabla_x u(x,t)\big) = 0 \quad (x \in \mathbb{R}^n, 0 < t < T) \\ u(x,0) = g(x). \end{cases}$$

A basic idea in PDE theory is to introduce some ordinary differential equations, the solution of which lets us compute the solution u. In particular, we want to find a curve $x(\cdot)$ along which we can, in principle at least, compute $u(x,t)$.

This section discusses this method of characteristics, to make clearer the connections between PDE theory and the Pontryagin Maximum Principle.

We have to note that:

$$\mathbf{x}(t) = \begin{pmatrix} x^1(t) \\ \vdots \\ x^n(t) \end{pmatrix}, \quad \mathbf{p}(t) = \nabla_x u(\mathbf{x}(t),t) = \begin{pmatrix} p^1(t) \\ \vdots \\ p^n(t) \end{pmatrix}.$$

Hence, to discuss the derivation of characteristic equations, we have

$$p^k(t) = u_{x_k}(\mathbf{x}(t), t),$$

and therefore

$$\dot{p}^k(t) = u_{x_k}(\mathbf{x}(t), t) + \sum_{i=1}^{n} u_{x_k x_k}(\mathbf{x}(t), t) \cdot \dot{x}^i.$$

Now suppose u solves (HJ). We differentiate this PDE with respect to the variable x_k:

$$u_{tx_k}(x, t) = -H_{x_k}\big(x, \nabla u(x, t)\big) - \sum_{i=1}^{n} H_{p_i}\big(x, \nabla u(x, t)\big) \cdot u_{x_k x_i}(x, t).$$

Let $x = \mathbf{x}(t)$ and substitute above:

$$\dot{p}^k(t) = -H_{x_k}(\mathbf{x}(t), \underbrace{\nabla_x u(\mathbf{x}(t), t))}_{\mathbf{p}(t)} + \sum_{i=1}^{n} (\dot{x}^i(t) - H_{pi}, \mathbf{x}(t), \underbrace{\nabla_x u(x, t)}_{\mathbf{p}(t)}) u_{x_k x_i}(\mathbf{x}(t), t).$$

We can simplify this expression if we select $\mathbf{x}(\cdot)$ so that

$$\dot{x}^i(t) = H_{p_i}\big(\mathbf{x}(t), \mathbf{p}(t)\big), \quad (1 \le i \le n);$$

then

$$\dot{p}^i(t) = H_{x_k}\big(\mathbf{x}(t), \mathbf{p}(t)\big), \quad (1 \le k \le n).$$

These are Hamilton's equations, already discussed in a different context in section (2.2):

$$\text{(H)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = \nabla_p H\big(\mathbf{x}(t), \mathbf{p}(t)\big) \\ \dot{\mathbf{p}}(t) = -\nabla_x H\big(\mathbf{x}(t), \mathbf{p}(t)\big). \end{cases}$$

We next demonstrate that if we can solve (H), then this gives a solution to PDE (HJ), satisfying the initial conditions $u = g$ on $t = 0$. Set $p^0 = \nabla g(x^0)$. We solve (H), with $x(0) = x^0$ and $\mathbf{p}(0) = p^0$. Next, let us calculate

$$\begin{aligned} \frac{d}{dt} u(\mathbf{x}(t), t) &= u_t(\mathbf{x}(t), t) + \nabla_x u(\mathbf{x}(t), t) \cdot \dot{\mathbf{x}}(t) \\ &= -H\,(\underbrace{\nabla_x u(\mathbf{x}(t), t)}_{\mathbf{p}(t)}, \mathbf{x}(t)) + \underbrace{\nabla_x u(\mathbf{x}(t), t)}_{\mathbf{p}(t)} \cdot \nabla_p H\big(\mathbf{x}(t), \mathbf{p}(t)\big) \end{aligned}$$

$$= -H(\mathbf{x}(t), \mathbf{p}(t)) + \mathbf{p}(t) \cdot \nabla_p H(\mathbf{x}(t), \mathbf{p}(t))$$

Note also $u(\mathbf{x}(0), 0) = u(x^0, 0) = g(x^0)$. Integrate, to compute u along the curve $\mathbf{x}(\cdot)$

$$u(\mathbf{x}(t), t) = \int_0^t -H + \nabla_p H \cdot \mathbf{p}\, ds + g(x^0).$$

This gives us the solution, once we have calculated $\mathbf{x}(\cdot)$ and $\mathbf{p}(\cdot)$.

To study the connections between Dynamic Programming and the Pontryagin Maximum Principle, let us return now to our usual control theory problem, with dynamics

(ODE) $$\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \boldsymbol{\alpha}(s)) & (t \leq s \leq T) \\ \mathbf{x}(t) = x, \end{cases}$$

and payoff

(P) $$P_{x,t}[\boldsymbol{\alpha}(\cdot)] = \int_t^T r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T)).$$

As above, the value function is

$$v(x, t) = \sup_{\boldsymbol{\alpha}(\cdot)} P_{x,t}[\boldsymbol{\alpha}(\cdot)].$$

The next theorem demonstrates that the costate in the Pontryagin Maximum Principle is in fact the gradient in x of the value function v, taken along an optimal trajectory:

Theorem (3.1.10): (Costates and Gradients)

Assume $\boldsymbol{\alpha}^*(\cdot), \mathbf{x}^*(\cdot)$ solve the control problem (ODE), (P).

If the value function v is C^2, then the costate $\dot{\mathbf{p}}^*(\cdot)$ occuring in the Maximum Principle is given by

$$\mathbf{p}^*(s) = \nabla_x v(\mathbf{x}^*(s), s) \quad (t \leq s \leq T).$$

Proof:

1. As usual, suppress the superscript *. Define $\mathbf{p}(t) := \nabla_x v(\mathbf{x}(t), t)$.

We claim that $p(\cdot)$ satisfies conditions (ADJ) and (M) of the Pontryagin Maximum Principle. To confirm this assertion, look at

$$\dot{p}^i(t) = \frac{d}{dt} v_{x_i}(\mathbf{x}(t), t) = v_{x_i t}(\mathbf{x}(t), t) + \sum_{j=1}^{n} v_{x_i x_j}(\mathbf{x}(t), t)\dot{x}^j(t).$$

We know v solves

$$v_t(x,t) + \max_{a \in A}\{\mathbf{f}(x,a) \cdot \nabla_x v(x,t) + r(x,a)\} = 0;$$

and, applying the optimal control $\boldsymbol{\alpha}(\cdot)$, we find:

$$u_t(\mathbf{x}(t), t) + \mathbf{f}\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big) \cdot \nabla_x v(\mathbf{x}(t), t) + r\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big) = 0.$$

2. Now freeze the time t and define the function

$$h(x) := v_t(x,t) + \mathbf{f}\big(x, \boldsymbol{\alpha}(t)\big) \cdot \nabla_x v(x,t) + r\big(x, \boldsymbol{\alpha}(t)\big) \leq 0.$$

Observe that $h\big(\mathbf{x}(t)\big) = 0$. Consequently $h(\cdot)$ has a maximum at the point $x = \mathbf{x}(t)$; and therefore for $i = 1, \ldots, n$,

$$\begin{aligned} 0 &= h_{x_i}\big(\mathbf{x}(t)\big) \\ &= v_{tx_i}(\mathbf{x}(t), t) + \mathbf{f}_{x_i}\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big) \cdot \nabla_x v(\mathbf{x}(t), t) + \mathbf{f}\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big) \\ &\cdot \nabla_x v_{x_i}(\mathbf{x}(t), t) + r_{x_i}\big(\mathbf{x}(t), \boldsymbol{\alpha}(t)\big). \end{aligned}$$

Substitute above:

$$\dot{p}^i(t) = v_{x_i t} + \sum_{i=1}^{n} v_{x_i x_j} f_j = v_{x_i t} + \mathbf{f} \cdot \nabla_x v_{x_i} = -\mathbf{f}_{x_i} \cdot \nabla_x v - r_{x_i}.$$

Recalling that $\mathbf{p}(t) = \nabla_x v(\mathbf{x}(t), t)$, we deduce that

$$\dot{\mathbf{p}}(t) = -(\nabla_x \mathbf{f})\mathbf{p} - \nabla_x r.$$

Recall also

$$H = \mathbf{f} \cdot p + r, \quad \nabla_x H = (\nabla_x \mathbf{f})p + \nabla_x r.$$

Hence

$$\dot{\mathbf{p}}(t) = -\nabla_x H\big(\mathbf{p}(t), \mathbf{x}(t)\big), \mathbf{x}(t),$$

which is (ADJ).

3. Now we must check condition (M). According to (HJB),

$$v_t(\mathbf{x}(t),t) + \max_{a\in A}\left\{\mathbf{f}(\mathbf{x}(t),a)\cdot \underbrace{\nabla v(\mathbf{x}(t),t)}_{\mathbf{p}(t)} + r(\mathbf{x}(t),t)\right\} = 0,$$

and maximum occurs for $a = \boldsymbol{\alpha}(t)$. Hence

$$\max_{a\in A}\{H(\mathbf{x}(t),\mathbf{p}(t),a)\} = H\big(\mathbf{x}(t),\mathbf{p}(t),\boldsymbol{\alpha}(t)\big);$$

and this is assertion (M) of the Maximum Principle.

The foregoing provides us with another way to look at transversality conditions:

(i) Free endpoint problem:

Recall that we stated earlier in Theorem (2.2.8) that for the free endpoint problem we have the condition

(T) $$\mathbf{p}^*(T) = \nabla g\big(\mathbf{x}^*(T)\big)$$

for the payoff functional

$$\int_t^T r\big(\mathbf{x}(s),\boldsymbol{\alpha}(s)\big)\,ds + g\big(\mathbf{x}(T)\big).$$

To understand this better, note $\mathbf{p}^*(s) = -\nabla v(\boldsymbol{x}^*(s),s)$. But $v(x,t) = g(x)$, and hence the foregoing implies

$$\mathbf{p}^*(T) = \nabla_x v(\mathbf{x}^*(T),T) = \nabla g\big(\mathbf{x}^*(T)\big).$$

(ii) Constrained initial and target sets:

Recall that for this problem we stated in Theorem (2.2.9) the transversality conditions that

(T) $$\begin{cases}\mathbf{p}^*(0)\text{is perpendicular to } T_0\\ \mathbf{p}^*(\tau^*)\text{is perpendicular to } T_1\end{cases}$$

when τ^*. denotes the first time the optimal trajectory hits the target set X_1.

Now let v be the value function for this problem:

$$v(x) = \sup_{\boldsymbol{\alpha}(\cdot)} P_x[\boldsymbol{\alpha}(\cdot)],$$

with the constraint that we start at $x^0 \in X_0$ and end at $x^1 \in X_1$ But then v will be constant on the set X_0 and also constant on X_1. Since ∇v is perpendicular to any level surface, ∇v is therefore perpendicular to both ∂X_0 and ∂X_1. And since

$$\mathbf{p}^*(t) = \nabla v(\mathbf{x}^*(t)),$$

this means that

$$\begin{cases} \mathbf{p}^* \text{ is perpendicular to } \partial X_0 \text{ at } t = 0, \\ \mathbf{p}^* \text{ is perpendicular to } \partial \mathrm{X}_1 \text{ at } t = \tau^*. \end{cases}$$

Section (3.2): Differential Games

Definitions (3.2.1):

We introduce in this section a model for a two-person, zero-sum differential game. The basic idea is that two players control the dynamics of some evolving system, and one tries to maximize, the other to minimize, a payoff functional that depends upon the trajectory.

What are optimal strategies for each player? This is a very tricky question, primarily since at each moment of time, each player's control decisions will depend upon what the other has done previously.

We begin with a Model Problem:

Let a time $0 \leq t < T$ be given, along with sets $A \subseteq \mathbb{R}^m, B \subseteq \mathbb{R}^l$ and a function $\mathbf{f} : \mathbb{R}^n \times A \times B \to \mathbb{R}^n$.

Definition (3.2.2):

A measurable mapping $\boldsymbol{\alpha}(\cdot): [t, T] \to A$ is a control for player I (starting at time t). A measurable mapping $\boldsymbol{\beta}(\cdot): [t, T] \to B$ is a control for player II.

Corresponding to each pair of controls, we have corresponding dynamics:

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \boldsymbol{\alpha}(s), \boldsymbol{\beta}(s)) & (t \leq s \leq T) \\ \mathbf{x}(t) = x, \end{cases}$$

the initial point $x \in \mathbb{R}^n$ being given.

Definition (3.2.3):

The payoff of the game is

(P) $$P_{x,t}[\boldsymbol{\alpha}(\cdot),\boldsymbol{\beta}(\cdot)] = \int_t^T r\big(\mathbf{x}(s),\boldsymbol{\alpha}(s),\boldsymbol{\beta}(s)\big)\,ds + g\big(\mathbf{x}(T)\big).$$

Player I, whose control is $\boldsymbol{\alpha}(\cdot)$, wants to maximize the payoff functional $\boldsymbol{P}[\cdot]$. Player II has the control $\boldsymbol{\beta}(\cdot)$ and wants to minimize $\boldsymbol{P}[\cdot]$. This is a two-person, zero-sum differential game.

We intend now to define value functions and to study the game using dynamic programming.

Definition (3.2.4):

The sets of controls for the game of the game are

$$A(t) := \{\boldsymbol{\alpha}(\cdot) : [t,T] \to A, \boldsymbol{\alpha}(\cdot) \text{ measurable}\}$$

$$B(t) := \{\boldsymbol{\beta}(\cdot) : [t,T] \to B, \boldsymbol{\beta}(\cdot) \text{ measurable}\}.$$

We need to model the fact that at each time instant, neither player knows the other's future moves. We will use concept of strategies, as employed by Varaiya and Elliott-Kalton. The idea is that one player will select in advance, not his control, but rather his responses to all possible controls that could be selected by his opponent.

Definitions (3.2.5):

(i) A mapping $\Phi : B(t) \to A(t)$ is called a strategy for player I if for all times $t \leq s \leq T$,

$$\boldsymbol{\beta}(\tau) \equiv \widehat{\boldsymbol{\beta}}(\tau) \quad \text{for } t \leq \tau \leq s$$

implies

$$\Phi[\boldsymbol{\beta}](\tau) \equiv \Phi[\widehat{\boldsymbol{\beta}}](\tau) \quad \text{for } t \leq \tau \leq s \tag{3.9}$$

We can think of $\Phi[\boldsymbol{\beta}]$ as the response of player I to player II's selection of control $\boldsymbol{\beta}(\cdot)$. Condition (3.9) expresses that player I cannot foresee the future.

(ii) A strategy for player II is a mapping $\Psi : A(t) \to B(t)$ such that for all times $t \leq s \leq T$,

$$\boldsymbol{\alpha}(\tau) \equiv \hat{\boldsymbol{\alpha}}(\tau) \quad \text{for } t \leq \tau \leq s$$

implies

$$\boldsymbol{\Psi}[\boldsymbol{\alpha}](\tau) \equiv \Psi[\hat{\boldsymbol{\alpha}}](\tau) \quad \text{for } t \leq \tau \leq s.$$

Definition (3.2.6):

The sets of strategies are

$$\mathcal{A}(t) := \text{ strategies for player } I \text{ (starting at } t)$$

$$\mathcal{B}(t) := \text{ strategies for player } II \text{ (starting at } t).$$

Finally, we introduce value functions:

Definition (3.2.7):

The lower value function is

$$v(x,t) := \inf_{\Psi \in \mathcal{B}(t)} \sup_{\boldsymbol{\alpha}(\cdot) \in A(t)} P_x[\boldsymbol{\alpha}(\cdot), \boldsymbol{\Psi}[\boldsymbol{\alpha}](\cdot)], \tag{3.10}$$

and the upper value function is

$$u(x,t) := \sup_{\Phi \in \mathcal{A}(t)} \inf_{\boldsymbol{\beta}(\cdot) \in B(t)} P_x[\Phi[\boldsymbol{\beta}](\cdot), \boldsymbol{\beta}(\cdot)], \tag{3.11}$$

One of the two players announces his strategy in response to the other's choice of control, the other player chooses the control. The player who " plays second ", i.e., who chooses the strategy, has an advantage. In fact, it turns out that always

$$v(x,t) \leq u(x,t).$$

Now we discuss Dynamic Programming, Isaacs' equations:

Theorem (3.2.8): (PDE for the Upper and Lower Value Functions)

Assume u, v are continuously differentiable. Then u solves the upper Isaacs' equation

$$\begin{cases} u_t + \min_{b \in B} \max_{a \in A}\{\mathbf{f}(x,a,b) \cdot \nabla_x u(x,t) + r(x,a,b)\} = 0 \\ u(x,T) = g(x), \end{cases} \tag{3.12}$$

and v solves the lower Isaacs' equation

$$\begin{cases} v_t + \mathrm{mac}_{a\in A} \min_{b\in B}\{\mathbf{f}(x,a,b)\cdot \nabla_x v(x,t) + r(x,a,b)\} = 0 \\ v(x,T) = g(x). \end{cases} \tag{3.13}$$

Isaacs' equations are analogs of Hamilton-Jacobi-Bellman equation in two. person, zero-sum control theory. We can rewrite these in the forms

$$u_t + H^+(x, \nabla_x u) = 0$$

for the upper PDE Hamiltonian

$$H^+(x,p) := \min_{b\in B} \max_{a\in A}\{\mathbf{f}(x,a,b)\cdot p + r(x,a,b)\};$$

and

$$v_t + H^-(x, \nabla_x v) = 0$$

for the lower PDE Hamiltonian

$$H^-(x,p) := \max_{a\in A} \min_{b\in B}\{\mathbf{f}(x,a,b)\cdot p + r(x,a,b)\}.$$

Interpretations and Remarks (3.2.9):

(i) In general, we have

$$\max_{a\in A} \min_{b\in B}\{\mathbf{f}(x,a,b)\cdot p + r(x,a,b)\} < \min_{b\in B} \max_{a\in A}\{\mathbf{f}(x,a,b)\cdot p + r(x,a,b)\},$$

and consequently $H^-(x,p) < H^+(x,p)$. The upper and lower Isaacs' equations are then different PDE and so in general the upper and lower value functions are not the same: $u \neq v$.

The precise interpretation of this is tricky, but the idea is to think of a slightly different situation in which the two players take turns exerting their controls over short time intervals. In this situation, it is a disadvantage to go first, since the other player then knows what control is selected. The value function u represents a sort of " infinitesimal" version of this situation, for which player **I** has the advantage. The value function v represents the reverse situation, for which player **II** has the advantage.

If however

$$\max_{a\in A} \min_{b\in B}\{\mathbf{f}(\cdots)\cdot p + r(\cdots)\} = \min_{b\in B} \max_{a\in A}\{\mathbf{f}(\cdots)\cdot p + r(\cdots)\}, \tag{3.14}$$

for all p, x, we say the game satisfies the minimax condition, also called Isaacs' condition. In this case it turns out that $u \equiv v$ and we say the game has value.

(ii) As in dynamic programming from control theory, if (3.14) holds, we can solve Isaacs' equation for $u \equiv v$ and then, at least in principle, design optimal controls for players I and II.

(iii) To say that $\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}^*(\cdot)$ are optimal means that the pair $(\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}^*(\cdot))$ is a saddle point for $P_{x,t}$. This means

$$P_{x,t}[\boldsymbol{\alpha}(\cdot), \boldsymbol{\beta}^*(\cdot)] \leq P_{x,t}[\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}^*(\cdot)] \leq P_{x,t}[\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}(\cdot)] \tag{3.15}$$

for all controls $\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}^*(\cdot)$.Player *I* will select $\boldsymbol{\alpha}^*(\cdot)$ because he is afraid *II* will play $\boldsymbol{\beta}^*(\cdot)$.Player *II* will play $\boldsymbol{\beta}^*(\cdot)$ because she is afraid *I* will play $\boldsymbol{\alpha}^*(\cdot)$.

To study the Games and the Pontryagin Maximum Principle, assume the minimax condition (3.14) holds and we design optimal $\alpha^*(\cdot), \beta^*(\cdot)$ as above. Let $\mathrm{x}^*(\cdot)$ denote the solution of the ODE (3.9), corresponding to our controls $\alpha^*(\cdot), \beta^*(\cdot)$. Then define

$$\mathbf{p}^*(t) := \nabla_x v(\mathbf{x}^*(t), t) = \nabla_x u(\mathbf{x}^*(t), t).$$

It turns out that

$$\text{(ADJ)} \qquad \dot{\mathbf{p}}^*(t) = -\nabla_x H\big(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}^*(\cdot)\big)$$

for the game-theory Hamiltonian

$$H(x, p, a, b) := \mathbf{f}(x, a, b) \cdot p + r(x, a, b).$$

Now to discuss Statement of Problem, we assume that two opponents I and II are at war with each other. Let us define

$$x^1(t) = \text{supply of resources for } I$$

$$x^2(t) = \text{supply of resources for } II.$$

Each player at each time can devote some fraction of his/her efforts to direct attack, and the remaining fraction to attrition (= guerrilla warfare). Set $A = B = [0,1]$, and define

$$\alpha(t) = \text{ fraction of } I'\text{s effort devoted to attrition}$$

$$1-\alpha(t) = \text{fraction of } I'\text{s effort devoted to attack}$$

$$\beta(t) = \text{fraction of } II'\text{s effort devoted to attrition}$$

$$1-\beta(t) = \text{fraction of } II's \text{ effort devoted to attack.}$$

We introduce as well the parameters

$$m_1 = \text{ rate of production of war material for } I$$

$$m_2 = \text{rate of production of war material for } II$$

$$c_1 = \text{effectiveness of } II'\text{s weapons against } I'\text{s production}$$

$$c_2 = \text{effectiveness of } I'\text{s weapons against } II'\text{s production}$$

We will assume

$$c_2 > c_1,$$

a hypothesis that introduces an asymmetry into the problem.

The dynamics are governed by the system of ODE

$$\begin{cases} \dot{x}^1(t) = m_1 - c_1\beta(t)x^2(t) \\ \dot{x}^2(t) = m_2 - c_2\alpha(t)x^1(t). \end{cases} \tag{3.16}$$

Let us finally introduce the payoff functional

$$P[\alpha(\cdot),\beta(\cdot)] = \int_0^T \big(1-\alpha(t)\big)x^1(t) - \big(1-\beta(t)\big)x^2(t)\,dt$$

the integrand recording the advantage of *I* over *II* from direct attacks at time t.Player *I* wants to maximize P, and player *II* wants to minimize P.

Now to applying Dynamic Programming, first, we check the minimax condition, for $n = 2, p = (p_1, p_2)$:

$$\begin{aligned} &\mathbf{f}(x,a,b)\cdot p + r(x,a,b) \\ &\qquad = (m_1 - c_1 b x_2)p_1 + (m_2 - c_2 a x_2)p_2 + (1-a)x_1 - (1-b)x_2 \\ &\qquad = m_1p_1 + m_2p_2 + x_1 - x_2 + a(-x_1 - c_2x_1p_2) + b(x_2 - c_1x_2p_1). \end{aligned}$$

Since a and b occur in separate terms, the minimax condition holds. Therefore $v \equiv u$ and the two forms of the Isaacs' equations agree:

$$v_t + H(x, \nabla_x v) = 0,$$

for

$$H(x, p) := H^+(x, p) = H^-(x, p).$$

We recall $A = B = [0,1]$ and $p = \nabla_x v$, and then choose $a \in [0,1]$ to maximize

$$a x_1 \left(-1 - c_2 v_{x_1}\right).$$

Likewise, we select $b \in [0,1]$ to minimize

$$b x_2 \left(1 - c_1 v_{x_1}\right).$$

Thus

$$\alpha = \begin{cases} 1 & \text{if} \quad -1 - c_2 v_{x_2} \geq 0 \\ 0 & \text{if} \quad -1 - c_2 v_{x_2} < 0, \end{cases} \tag{3.17}$$

and

$$\beta = \begin{cases} 1 & \text{if} \quad -1 - c_1 v_{x_1} \geq 0 \\ 0 & \text{if} \quad -1 - c_1 v_{x_1} < 0. \end{cases} \tag{3.18}$$

So if we knew the value function v, we could then design optimal feedback controls for *I*, *II*.

It is however hard to solve Isaacs' equation for v, and so we switch approaches.

Now to applying the Maximum Principle, assume $\alpha(\cdot), \beta(\cdot)$ are selected as above, and $\mathbf{x}(\cdot)$ corresponding solution of the ODE (3.16). Define

$$\mathbf{p}(t) := \nabla_x v(\mathbf{x}(t), t).$$

By results stated above, $\mathbf{p}(\cdot)$ solves the adjoint equation

$$\dot{\mathbf{p}}(t) = -\nabla_x H\big(\mathbf{x}(t), \mathbf{p}(t), \alpha(\cdot), \beta(\cdot)\big) \tag{3.19}$$

for

$$H(x, p, a, b) = p \cdot \mathbf{f}(x, a, b) + r(x, a, b)$$

$$= p_1(m_1 - c_1 b x_2) + p_2(m_2 - c_2 a x_1) + (1-a)x_1 - (1-b)x_2.$$

Therefore (3.19) reads

$$\begin{cases} \dot{p}^1 = \alpha - 1 + p^2 c_2 \alpha \\ \dot{p}^2 = 1 - \beta + p^1 c_1 \beta, \end{cases} \tag{3.20}$$

with the terminal conditions $p^1(T) = p^2(T) = 0$.

We introduce the further notation

$$s^1 := -1 - c_2 v_{x_2} = -1 - c_2 p^2, \quad s^2 := 1 - c_1 v_{x_1} = 1 - c_1 p^1;$$

so that, according to (3.17) and (3.18), the functions s^1 and s^2 control when player *I* and player *II* switch their controls. In the following we will study s^1 and s^2

(1) Dynamics for s^1 and s^2

Our goal now is to find ODE for s^1, s^2. We compute

$$\dot{s}^1 = -c_2 \dot{p}^2 = c_2(\beta - 1 - p^1 c_1 \beta) = c_2\big(-1 + \beta(1 - p^1 c_1)\big) = c_2(-1 + \beta s^2)$$

and

$$\dot{s}^2 = -c_1 \dot{p}^1 = c_1(1 - \alpha - p^2 c_2 \alpha) = c_1\big(1 + \alpha(-1 - p^2 c_2)\big) = c_1(1 + \alpha s^1).$$

Therefore

$$\begin{cases} \dot{s}^1 = c_2(-1 + \beta s^2), \quad s^1(T) = -1 \\ \dot{s}^2 = c_1(1 + \alpha s^1), \quad s^2(T) = 1. \end{cases} \tag{3.21}$$

Recall from (3.17) and (3.18) that

$$\alpha = \begin{cases} 1 & \text{if} \quad s_1 \geq 0 \\ 0 & \text{if} \quad s_1 < 0, \end{cases}$$

$$\beta = \begin{cases} 1 & \text{if} \quad s_2 \geq 0 \\ 0 & \text{if} \quad s_2 < 0. \end{cases}$$

Consequently, if we can find s^1, s^2, then we can construct the optimal controls α and β.

(2) Calculating s^1 and s^2

We work backwards from the terminal time T. Since at time T, we have $s^1 < 0$ and $s^2 > 0$, the same inequalities hold near T. Hence we have $\alpha = \beta \equiv 0$ near T, meaning a full attack from both sides.

Next, let $t^* < T$ be the first time going backward from T at which either *I* or *II* switches stategy. Our intention is to compute t^*. On the time interval $[t^*, T]$, we have $\alpha(\cdot) \equiv \beta(\cdot) \equiv 0$. Thus (6.21) gives

$$\dot{s}^1 = -c_2, \ \ s^1(T) = -1, \quad \dot{s}^2 = c_1, \ \ s^2(T) = 1;$$

and therefore

$$s^1(t) = -1 + c_2(T - t), \quad s^2(t) = 1 + c_1(t - T)$$

for times $t^* \le t \le T$. Hence s^1 hits 0 at time $T - \frac{1}{c_2}$; s^2 hits 0 at time $T - \frac{1}{c_1}$. Remember that we are assuming $c_2 > c_1$. Then $T - \frac{1}{c_1} < T - \frac{1}{c_2}$, and hence

$$t^* = T - \frac{1}{c_2}.$$

Now define$t_* < t^*$. to be the next time going backward when player *I* or player *II* switches. On the time interval $[t_*, t^*]$, we have $\alpha \equiv 1, \ \beta \equiv 0$. Therefore the dynamics read:

$$\begin{cases} \dot{s}^1 = -c_2, \quad s^1(t^*) = 0 \\ \dot{s}^2 = c_1(1 + s^1), \quad s^2(t^*) = 1 - \dfrac{c_1}{c_2}. \end{cases}$$

We solve these equations and discover that

$$\begin{cases} s^1(t) = -1 + c_2(T - t) \\ s^2(t) = 1 - \dfrac{c_1}{2c_2} - \dfrac{c_1 c_2}{2}(t - T)^2. \quad (t_* \le t \le t^*). \end{cases}$$

Now $s^1 > 0$ on $[t_*, t^*]$ for all choices of t_*. But $s^2 = 0$ at

$$t_* := T - \frac{1}{c_2}\left(\frac{2c_2}{c_1} - 1\right)^{1/2}.$$

If we now solve (3.21) on $[0, t_*]$ with $\alpha = \beta \equiv 1$, we learn that s_1, s_2 do not change sign.

We have assumed that $x_1 > 0$ and $x_2 > 0$ for all times t. If either x_1 or x_2 hits the constraint, then there will be a corresponding Lagrange multiplier and everything becomes much more complicated.

Chapter (4)

Stochastic Control Theory

Section (4.1): Stochastic probability theory, Brownian motion

Stochastic Differential Equations. We begin with a brief overview of random differential equations. Consider a vector field $\mathbf{f}: \mathbb{R}^n \to \mathbb{R}^n$ and the associated ODE

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathrm{x}(t)) & (t > 0) \\ \mathbf{x}(0) = x^0. \end{cases} \tag{4.1}$$

In many cases a better model for some physical phenomenon we want to study is the stochastic differential equation

$$\begin{cases} \dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t)) + \sigma_{\xi}(t) & (t > 0) \\ \mathbf{X}(0) = x^0, \end{cases} \tag{4.2}$$

where $\xi(\cdot)$ denotes a "white noise" term causing random fluctuations. We have switched notation to a capital letter $\boldsymbol{X}(\cdot)$ to indicate that the solution is random. A solution of (4.2) is a collection of sample paths of a stochastic process, plus probabilistic information as to the likelihoods of the various paths.

Now assume $\mathbf{f}: \mathbb{R}^n \times A \to \mathbb{R}^n$ and turn attention to the controlled stochastic differential equation:

$$\text{(SDE)} \qquad \begin{cases} \dot{\mathbf{X}}(s) = \mathbf{f}(\mathbf{X}(s), \mathbf{A}(s)) + \xi(s) \\ \mathbf{X}(t) = x^0, \end{cases} \qquad (t \le s \le T)$$

Definition (4.1.1):

(i) A control $\mathbf{A}(\cdot)$ is a mapping of $[t, T]$ into A, such that for each time $t \le s \le T$, $\mathbf{A}(s)$ depends only on s and observations of $\mathbf{X}(\tau)$ for $t \le \tau \le s$.

(ii) The corresponding payoff functional is

$$(P) \qquad P_{x,t}[\mathbf{A}(\cdot)] = E\left\{\int_t^T r(\mathbf{X}(s), \mathbf{A}(s))ds + g(\mathbf{X}(T))\right\},$$

the expected value over all sample paths for the solution of (SDE). As usual, we are given the running payoff r and terminal payoff g.

Our goal is to find an optimal control $\mathbf{A}^*(\cdot)$, such that

$$P_{x,t}[\mathbf{A}^*(\cdot)] = \max_{\mathbf{A}(\cdot)} P_{x,t}[\mathbf{A}(\cdot)].$$

To discuss the dynamic programming, we will adapt a dynamic programming methods, so, we firstly define the value function

$$\upsilon(x,t) \coloneqq \sup_{\mathbf{A}(\cdot)} P_{x,t}[\mathbf{A}(\cdot)].$$

The overall plan to find an optimal control $\mathbf{A}^*(\cdot)$ will be (i) to find a Hamilton-Jacobi-Bellman type of PDE that υ satisfies, and then (ii) to utilize a solution of this PDE in designing $\mathbf{A}^*$.

It will be particularly interesting to see [2] how the stochastic effects modify the structure of the Hamilton-Jacobi-Bellman (HJB) equation, as compared with the deterministic case [4].

Definition (4.1.2):

A probability space is a triple $(\Omega, \mathcal{F}, P)$, where

(i) Ω is a set,

(ii) $\mathcal{F}$ is a σ-algebra of subsets of Ω,

(iii) P is a mapping from $\mathcal{F}$ into $[0,1]$ such that $P(\emptyset) = 0, P(\Omega) = 1$, and $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$, provided $A_i \cap A_j = \emptyset$ for all $i \neq j$.

A typical point in Ω is denoted "ω" and is called a sample point. A set $A \in \mathcal{F}$ is called an event. We call P a probability measure on Ω, and $P(A) \in [0, 1]$ is probability of the event A.

Definition (4.1.3):

A random variable **X** is a mapping $X: \Omega \to \mathbb{R}$ such that for all $t \in \mathbb{R}$

$$\{\omega | X(\omega) \leq t\} \in \mathcal{F}.$$

We mostly employ capital letters to denote random variables. Often the dependence of X on ω is not explicitly displayed in the notation.

Definition (4.1.4):

Let X be a random variable, defined on some probability space $(\Omega, \mathcal{F}, P)$. The expected value of X is

$$E[X] := \int_\Omega X \, dP.$$

Example (4.1.5):

Assume $\Omega \subset \mathbb{R}^m$, and $P(A) = \int_A f \, d\omega$ for some function $f: \mathbb{R}^m \to [0, \infty)$, with $\int_\Omega f \, d\omega = 1$. We then call f the density of the probability P, and write "$dP = f \, d\omega$". In this case,

$$E[X] = \int_\Omega X f \, d\omega.$$

Definition (4.1.6):

We define also the variance

$$\mathrm{Var}(X) = E[(X - E(X))^2] = E[X^2] - (E[X])^2.$$

Example (4.1.7):

A random variable X is called normal (or Gaussian) with mean μ, variance σ^2 if for all $-\infty \leq a < b \leq \infty$

$$P(a \leq X \leq b) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx,$$

We write "X is $N(\mu, \sigma^2)$".

Definitions (4.1.8):

(i) Two events $A, B \in \mathcal{F}F$ are called independent if

$$P(A \cap B) = P(A)P(B).$$

(ii) Two random variables X and Y are independent if

$$P(X \leq t \text{ and } Y \leq s) = P(X \leq t)P(Y \leq s)$$

for all $t, s \in \mathbb{R}$. In other words, X and Y are independent if for all t, s the events $A = \{X \leq t\}$ and $B = \{Y \leq s\}$ are independent.

Definition (4.1.9):

A stochastic process is a collection of random variables $X(t)(0 \leq t < \infty)$, each defined on the same probability space $(\Omega, \mathcal{F}, P)$.

The mapping $t \mapsto X(t, \omega)$ is the ω-th sample path of the process.

Definition (4.1.10):

A real-valued stochastic process $W(t)$ is called a Wiener process or Brownian motion if:

(i) $W(0) = 0$,
(ii) each sample path is continuous,
(iii) $W(t)$ is Gaussian with $\mu = 0, \sigma^2 = t$ (that is, $W(t)$ is $N(0, t)$),
(iv) for all choices of times $0 < t_1 < t_2 < \cdots < t_m$ the random variables

$$W(t_1), W(t_2) - W(t_1), \ldots, W(t_m) - W(t_{m-1})$$

are independent random variables.

Assertion (iv) says that W has "independent increments".

Interpretation (4.1.11):

We heuristically interpret the one-dimensional "white noise" $\xi(\cdot)$ as equalling $\frac{dW(t)}{dt}$. However, this is only formal, since for almost all ω, the sample path $t \mapsto W$ is in fact nowhere differentiable.

Definition (4.1.12):

An n-dimensional Brownian motion is

$$\mathbf{W}(t) = (W^1(t), W^2(t), \ldots, W^n(t))^T$$

when the $W^i(t)$ are independent one-dimensional Brownian motions.

We use boldface below to denote vector-valued functions and stochastic processes.

Section (4.2): Stochastic calculus, Itô chain rule

We discuss next how to understand stochastic differential equations, driven by "white noise". Consider first of all

$$\begin{cases} \dot{\mathbf{X}}(t) = \mathbf{f}\big(\mathbf{X}(t)\big) + \sigma_{\xi}(t) \quad (t > 0) \\ \mathbf{X}(0) = x^0, \end{cases} \tag{4.3}$$

where we informally think of $\boldsymbol{\xi} = \dot{\boldsymbol{W}}$.

Definition (4.2.1):

A stochastic process $\mathbf{X}(\cdot)$ solves (4.3) if for all times $t \geq 0$, we have

$$\mathbf{X}(t) = x^0 + \int_0^t f\big(\mathbf{X}(s)\big)ds + \sigma\mathbf{W}(t). \tag{4.4}$$

Remarks (4.2.2):

(i) It is possible to solve (4.4) by the method of successive approximation. For this, we set $\mathbf{X}^0(\cdot) \equiv x$, and inductively define

$$\mathbf{X}^{k+1}(t) \coloneqq x^0 + \int_0^t f\big(\mathbf{X}^k(s)\big)ds + \sigma\mathbf{W}(t).$$

It turns out that $\mathbf{X}^k(s)$ converges to a limit $\mathbf{X}(s)$ for all $t \geq 0$ and $\mathbf{X}(\cdot)$ solves the integral identity (4.4).

(ii) Consider a more general SDE

$$\dot{\mathbf{X}}(t) = \mathbf{f}\big(\mathbf{X}(t)\big) + \mathbf{H}\big(\mathbf{X}(t)\big)\xi(t) \qquad (t > 0), \tag{4.5}$$

which we formally rewrite to read:

$$\frac{d\mathbf{X}(t)}{dt} = \mathbf{f}\big(\mathbf{X}(t)\big) + \mathbf{H}\big(\mathbf{X}(t)\big)\frac{d\mathbf{W}(t)}{dt}$$

and then

$$d\mathbf{X}(t) = \mathbf{f}\big(\mathbf{X}(t)\big)dt + \mathbf{H}\big(\mathbf{X}(t)\big)d\mathbf{W}(t).$$

This is an Itô stochastic differential equation. By analogy with the foregoing, we say $\mathbf{X}(\cdot)$ is a solution, with the initial condition $\mathbf{X}(0) = x^0$, if

$$\mathbf{X}(t) = x^0 + \int_0^t \mathbf{f}(\mathbf{X}(s))ds + \int_0^t \mathbf{H}(\mathbf{X}(s)) \cdot d\mathbf{W}(s)$$

for all times $t \geq 0$. In this expression $\int_0^t \mathbf{H}(\mathbf{X}) \cdot d\mathbf{W}$ is called an Itô stochastic integral.

Remark (4.2.3):

Given a Brownian motion $\mathbf{W}(\cdot)$ it is possible to define the Itô stochastic integral

$$\int_0^t \mathbf{Y} \cdot d\mathbf{W}$$

for processes $\mathbf{Y}(\cdot)$ having the property that for each time $0 \leq s \leq t$ "$\mathbf{Y}(s)$ depends on $W(\tau)$ for times $0 \leq \tau \leq s$, but not on $W(\tau)$ for times $s \leq \tau$ Such processes are called "nonanticipating".

We will not here explain the construction of the Itô integral, but will just record one of its useful properties:

$$E\left[\int_0^t \mathbf{Y} \cdot d\mathbf{W}\right] = 0. \tag{4.6}$$

Once the Itô stochastic integral is defined, we have in effect constructed a new calculus, the properties of which we should investigate. This section explains that the chain rule in the Itô calculus contains additional terms as compared with the usual chain rule in one and higher dimensions.

(1) One Dimension: We suppose that $n = 1$ and

$$\begin{cases} d\mathbf{X}(t) = A(t)dt + B(t)d\boldsymbol{W}(t) \qquad (t \geq 0) \\ X(0) = x^0. \end{cases} \tag{4.7}$$

The expression (4.7) means that

$$X(t) = x^0 + \int_0^t A(s)ds + \int_0^t B(s)\, dW(s)$$

for all times $t \geq 0$.

Let $u: \mathbb{R} \to \mathbb{R}$ and define

$$Y(t) \coloneqq u(X(t)).$$

We ask: what is the law of motion governing the evolution of Y in time? Or, in other words, what is $dY(t)$?

It turns out, quite surprisingly, that it is incorrect to calculate

$$dY(t) = d\left(u\big(X(t)\big)\right) = u'\big(X(t)\big)dX(t) = u'\big(X(t)\big)\big(A(t)dt + B(t)dW(t)\big).$$

We try again and make use of the heuristic principle that "$dW = (dt)^{1/2}$". So let us expand u into a Taylor's series, keeping only terms of order dt or larger. Then

$$\begin{aligned} dY(t) &= d\left(u\big(X(t)\big)\right) \\ &= u'\big(X(t)\big)\, dX(t) + \frac{1}{2}u''\big(X(t)\big)dX(t)^2 + \frac{1}{6}u'''\big(X(t)\big)dX(t)^3 + \cdots \\ &= u'\big(X(t)\big)[A(t)dt + B(t)dW(t)] + \frac{1}{2}u''\big(X(t)\big)[A(t)dt + B(t)dW(t)]^2 + \cdots, \end{aligned}$$

the last line following from (4.7). Now, formally at least, the heuristic that $dW = (dt)^{1/2}$ implies

$$\begin{aligned} [A(t)dt + B(t)dW(t)]^2 &= A(t)^2dt^2 + 2A(t)B(t)dtdW(t) + B^2(t)dW(t)^2 \\ &= B^2(t)dt + o(dt). \end{aligned}$$

Thus, ignoring the $o(dt)$ term, we derive the one-dimensional Itô chain rule

$$\begin{aligned} dY(t) &= d\left(u\big(X(t)\big)\right) \\ &= \left[u'\big(X(t)\big)A(t) + \frac{1}{2}B^2(t)u''\big(X(t)\big)\right]dt + u'\big(X(t)\big)B(t)dW(t). \end{aligned} \tag{4.8}$$

This means that for each time $t > 0$

$$\begin{aligned} u\big(X(t)\big) &= Y(t) \\ &= Y(0) + \int_0^t \left[u'\big(X(s)\big)A(s) + \frac{1}{2}B^2(s)u''\big(X(s)\big)\right]ds \\ &\quad + \int_0^t u'\big(X(s)\big)B(s)dW(s). \end{aligned}$$

(2) Higher Dimensions: We turn now to stochastic differential equations in higher dimensions. For simplicity, we consider only the special form

$$\begin{cases} d\mathbf{X}(t) = \mathbf{A}(t)dt + \sigma d\boldsymbol{W}(t) & (t \geq 0) \\ \mathbf{X}(0) = x^0. \end{cases} \tag{4.9}$$

We write

$$X(t) = \left(X^1(t), X^2(t), \ldots, X^n(t)\right)^T.$$

The stochastic differential equation means that for each index i, we have $dX^i(t) = A^i(t)dt + \sigma dW^i(t)$.

Hence to explain the ITô Chain Rule again, let $u: \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$ and put

$$Y(t) \coloneqq u(\mathbf{X}(t), t).$$

What is dY? Similarly to the computation above, we calculate

$$\begin{aligned} dY(t) &= d[u(X(t), t)] \\ &= u_t(X(t), t)dt + \sum_{i=1}^{n} u_{x_i}(X(t), t)dX^i(t) \\ &+ \frac{1}{2}\sum_{i,j=1}^{n} u_{x_i x_j}(X(t), t)dX^i(t)dX^j(t). \end{aligned}$$

Now use (4.9) and the heuristic rules that

$$dW^i = (dt)^{1/2} \quad \text{and} \quad dW^i\, dW^j = \begin{cases} dt & \text{if i=j} \\ 0 & \text{if i≠j.} \end{cases}$$

The second rule holds since the components of $d\mathbf{W}$ are independent. Plug these identities into the calculation above and keep only terms of order dt or larger:

$$\begin{aligned} d\mathbf{Y}(t) &= u_t(X(t), t)dt + \sum_{i=1}^{n} u_{x_i}(\mathbf{X}(t), t)[A^i(t)dt + \sigma dW^i(t)] \\ &+ \frac{\sigma^2}{2}\sum_{i=1}^{n} u_{x_i x_j}(\mathbf{X}(t), t)dt \qquad (4.10) \\ &= u_t(\mathbf{X}(t), t) + \nabla_x u(X(t), t) \cdot [\mathbf{A}(t)dt + \sigma d\mathbf{W}(t)] + \frac{\sigma^2}{2}\Delta u(\mathbf{X}(t), t)dt. \end{aligned}$$

This is Itô's chain rule in n-dimensions. Here

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$$

denotes the Laplacian.

Now we discuss two applications to PDE;

(1) A stochastic representation formula for harmonic functions: Consider a region $U \subseteq \mathbb{R}^n$ and the boundary-value problem

$$\begin{cases} \Delta u = 0 & (x \in U) \\ u = g & (x \in \partial U) \end{cases} \tag{4.11}$$

where, as above, $\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ is the Laplacian. We call u a harmonic function.

We develop a stochastic representation formula for the solution of (4.11). Consider the random process $\mathbf{X}(t) = \mathbf{W}(t) + x$; that is,

$$\begin{cases} d\mathbf{X}(t) = d\mathbf{W}(t) & (t > 0) \\ \mathbf{X}(0) = 0 \end{cases}$$

and $W(\cdot)$ denotes an n-dimensional Brownian motion. To find the link with the PDE (4.11), we define $Y(t) := u(\mathbf{X}(t))$. Then Itô's rule (4.10) gives

$$dY(t) = \nabla u\big(\mathbf{X}(t)\big) \cdot d\mathbf{W}(t) + \frac{1}{2} \Delta u\big(\mathbf{X}(t)\big) dt$$

Since $\Delta u \equiv 0$, we have

$$dY(t) = \nabla u(\mathbf{X}(t)) \cdot d\mathbf{W}(t);$$

which means

$$u\big(\mathbf{X}(t)\big) = Y(t) = Y(0) + \int_0^t \nabla u\big(\mathbf{X}(s)\big) \cdot d\mathbf{W}(s).$$

Let τ denote the (random) first time the sample path hits ∂U. Then, putting $t = \tau$ above, we have

$$u(x) = u\big(\mathbf{X}(\tau)\big) - \int_0^\tau \nabla u \cdot d\mathbf{W}(s).$$

But $u\big(\mathbf{X}(\tau)\big) = g\big(\mathbf{X}(\tau)\big)$, by definition of τ . Next, average over all sample paths:

$$u(x) = E\big[g\big(\mathbf{X}(\tau)\big)\big] - E\left[\int_0^\tau \nabla u \cdot d\mathbf{W}\right].$$

The last term equals zero, according to (4.6). Consequently,

$$u(x) = E\big[g\big(\mathbf{X}(\tau)\big)\big].$$

Interpretation (4.2.4):

Consider all the sample paths of the Brownian motion starting at x and take the average of $g\big(\mathbf{X}(\tau)\big)$. This gives the value of u at x.

(2) A time-dependent problem: We next modify the previous calculation to cover the terminal-value problem for the inhomogeneous backwards heat equation:

$$\begin{cases} u_t(x,t) + \dfrac{\sigma^2}{2}\Delta u(x,t) = f(x,t) \qquad (x \in \mathbb{R}^n, 0 \leq t < T) \\ u(x,T) = g(x). \end{cases} \tag{4.12}$$

Fix $0 \leq t < T$. We introduce the stochastic process

$$\begin{cases} d\mathbf{X}(s) = \sigma d\mathbf{W}(s) \quad (st \geq) \\ \mathbf{X}(t) = x. \end{cases}$$

Use Itô's chain rule (4.10) to compute $du(\mathbf{X}(s), s)$:

$$du(\mathbf{X}(s), s) = u_s(\mathbf{X}(s), s)ds + \nabla_x u(\mathbf{X}(s), s) \cdot d\mathbf{X}(s) + \frac{\sigma^2}{2}\Delta u(\mathbf{X}(s), s).$$

Now integrate for times $t \leq s \leq T$, to discover

$$u(\mathbf{X}(T), T) = u(\mathbf{X}(t), t) + \int_t^T \frac{\sigma^2}{2}\Delta u(\mathbf{X}(s), s) + u_s(\mathbf{X}(s), s)ds$$
$$+ \int_t^T \sigma \nabla_x u(\mathbf{X}(s), s) \cdot d\mathbf{W}(s).$$

Then, since u solves (4.12):

$$u(x,t) = E\left(g\big(\mathbf{X}(T)\big) - \int_t^T f(\mathbf{X}(s), s)ds\right).$$

Section (4.3): Dynamic Programming and Application

We now turn our attention to controlled stochastic differential equations, of the form

$$\text{(SDE)}\qquad \begin{cases} d\mathbf{X}(s) = \mathbf{f}\big(\mathbf{X}(s), \mathbf{A}(s)\big)ds + \sigma d\mathbf{W}(s) \quad (t \le s \le T) \\ \mathbf{X}(t) = x. \end{cases}$$

Therefore

$$\mathbf{X}(\tau) = x + \int_t^\tau \mathbf{f}\big(\mathbf{X}(s), \mathbf{A}(s)\big)ds + \sigma[\mathbf{W}(\tau) - \mathbf{W}(t)]$$

for all $t \le \tau \le T$. We introduce as well the expected payoff functional

$$(P)\qquad P_{x,t}[\mathbf{A}(\cdot)] \coloneqq E\left\{\int_t^T r\big(\mathbf{X}(\tau), \mathbf{A}(s)\big)ds + g\big(\mathbf{X}(T)\big)\right\}.$$

The value function is

$$v(x,t) \coloneqq \sup_{\mathbf{A}(\cdot)} P_{x,t}[\mathbf{A}(\cdot)].$$

We will employ the method of dynamic programming. To do so, we must:

(i) Find a PDE satisfied by v, and then

(ii) (ii) Use this PDE to design an optimal control $\mathbf{A}^*(\cdot)$.

Now to find a PDE for the value function, let $\mathbf{A}(\cdot)$ be any control, and suppose we use it for times $t \le s \le t + h, h > 0$, and thereafter employ an optimal control. Then

$$v(x,t) \ge E\left\{\int_t^{t+h} r\big(\mathbf{X}(s), \mathbf{A}(s)\big)\, ds + v(\mathbf{X}(t+h), t+h)\right\}, \tag{4.13}$$

and the inequality in (4.13) becomes an equality if we take $\mathbf{A}(\cdot) = \mathbf{A}^*(\cdot)$, an optimal control.

Now from (4.13) we see for an arbitrary control that

$$\begin{aligned} 0 &\ge E\left\{\int_t^{t+h} r\big(\mathbf{X}(s), \mathbf{A}(s)\big)\, ds + v(\mathbf{X}(t+h), t+h) - v(x,t)\right\} \\ &= E\left\{\int_t^{t+h} r\, ds\right\} + E\{v(\mathbf{X}(t+h), t+h) - v(x,t)\} \end{aligned}$$

Recall next Itô's formula:

$$
\begin{aligned}
dv(\mathbf{X}(s),s) &= v_t(\mathbf{X}(s),s)ds + \sum_{i=1}^{n} v_{x_i}(\mathbf{X}(s),s)\, dX^i(s) \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n} v_{x_i x_j}(\mathbf{X}(s),s)\, dX^i(s)dX^j(s) \\
&= v_t ds + \nabla_x v \cdot \big(\mathbf{f}ds + \sigma d\mathbf{W}(s)\big) + \frac{\sigma^2}{2}\Delta v ds. \qquad (4.14)
\end{aligned}
$$

This means that

$$
\begin{aligned}
&v(\mathbf{X}(t+h),t+h) - v(\mathbf{X}(t),t) \\
&\qquad = \int_t^{t+h}\left(v_t + \nabla_x v \cdot \mathbf{f} + \frac{\sigma^2}{2}\Delta v\right)ds + \int_t^{t+h} \sigma \nabla_x v \cdot d\mathbf{W}(s)\,;
\end{aligned}
$$

and so we can take expected values, to deduce

$$
E[v(\mathbf{X}(t+h),t+h) - v(x,t)] = E\left[\int_t^{t+h}\left(v_t + \nabla_x v \cdot \mathbf{f} + \frac{\sigma^2}{2}\Delta v\right)ds\right]. \qquad (4.15)
$$

We derive therefore the formula

$$
0 \geq E\left[\int_t^{t+h}\left(r + v_t + \nabla_x v \cdot \mathbf{f} + \frac{\sigma^2}{2}\Delta v\right)ds\right].
$$

Divide by h:

$$
\begin{aligned}
0 \geq E\Bigg[\frac{1}{h}\int_t^{t+h} & r\big(\mathbf{X}(s),\mathbf{A}(s)\big) + v_t(\mathbf{X}(s),s) + f\big(\mathbf{X}(s),\mathbf{A}(s)\big)\cdot \nabla_x v(\mathbf{X}(s),s) \\
&+ \frac{\sigma^2}{2}\Delta v(\mathbf{X}(s),s)ds\Bigg].
\end{aligned}
$$

If we send $h \to 0$, recall that $\mathbf{X}(t) = x$ and set $\mathbf{A}(t) := a \in A$, we see that

$$
0 \geq r(x,a) + v_t(x,t) + \mathbf{f}(x,a)\cdot \nabla_x v(x,t) + \frac{\sigma^2}{2}\Delta v(x,t).
$$

The above identity holds for all x, t, a and is actually an equality for the optimal control. Hence

$$\max_{a\in A}\left\{v_t+\mathbf{f}\cdot\nabla_x v+\frac{\sigma^2}{2}\Delta v+r\right\}=0.$$

In summary, we have shown that the value function v for our stochastic control problem solves this PDE:

$$\text{(HJB)}\quad\begin{cases}v_t(x,t)+\dfrac{\sigma^2}{2}\Delta v(x,t)+\max_{a\in A}\{\mathbf{f}(x,a)\cdot\nabla_x v(x,t)+r(x,a)\}=0\\ v(x,T)=g(x).\end{cases}$$

This semilinear parabolic PDE is the stochastic Hamilton-Jacobi-Bellman equation.

Assume now that we can somehow solve the (HJB) equation, and therefore know the function v. We can then compute for each point (x,t) a value $a\in A$ for which $\nabla_x v(x,t)\cdot\mathbf{f}(x,a)+r(x,a)$ attains its maximum. In other words, for each (x,t) we choose $a=\alpha(x,t)$ such that

$$\max_{a\in A}[\mathbf{f}(x,a)\cdot\nabla_x v(x,t)+r(x,a)]$$

occurs for $a=\alpha(x,t)$. Next solve

$$\begin{cases}d\mathbf{X}^*(s)=\mathbf{f}\big(\mathbf{X}^*(s),\alpha(\mathbf{X}^*(s),s)\big)ds+\sigma d\mathbf{W}(s)\\ \mathbf{X}^*(t)=x.\end{cases}$$

assuming this is possible. Then $\mathbf{A}^*(s)=\alpha(\mathbf{X}^*(s),s)$ is an optimal feedback control.

Following is an interesting example worked out by Merton. In this model we have the option of investing some of our wealth in either a risk-free bond (growing at a fixed rate) or a risky stock (changing according to a random differential equation). We also intend to consume some of our wealth as time evolves. As time goes on, how can we best (i) allot our money among the investment opportunities and (ii) select how much to consume?

We assume time runs from 0 to a terminal time T. Introduce the variables

$X(t)$ = wealth at time t (random)

$b(t)$ = price of a risk-free investment, say a bond

$S(t)$ = price of a risky investment, say a stock (random)

$\alpha^1(t)$ = fraction of wealth invested in the stock

$\alpha^2(t)$ = rate at which wealth is consumed.

Then

$$0 \leq \alpha^1(t) \leq 1, \quad 0 \leq \alpha^2(t) \quad (0 \leq t \leq T) \tag{4.16}$$

We assume that the value of the bond grows at the known rate $r > 0$:

$$db = rbdt; \tag{4.17}$$

whereas the price of the risky stock changes according to

$$dS = RSdt + \sigma SdW. \tag{4.18}$$

Here r, R, σ are constants, with

$$R > r > 0, \qquad \sigma \neq 0.$$

This means that the average return on the stock is greater than that for the risk-free bond.

According to (4.17) and (4.18), the total wealth evolves as

$$dX = (1 - \alpha^1(t))Xr\,dt + \alpha^1(t)X(Rdt + \sigma dW) - \alpha^2(t)dt. \tag{4.19}$$

Let

$$Q := \{(x,t) | 0 \leq t \leq T, \qquad x \geq 0\}$$

and denote by τ the (random) first time $\mathbf{X}(\cdot)$ leaves Q. Write $\mathbf{A}(t) = (\alpha^1(t), \alpha^2(t))^T$ for the control.

The payoff functional to be maximized is

$$P_{x,t}[\mathbf{A}(\cdot)] = E\left(\int_t^\tau e^{-\rho s} F(\alpha^2(s))\,ds\right),$$

where F is a given utility function and $\rho > 0$ is the discount rate.

Guided by theory similar to that developed [2], we discover that the corresponding (HJB) equation is

$$u_t \max_{0 \leq a_1 \leq 1, a_2 \geq 0} \left\{\frac{(a_1 x \sigma)^2}{2} u_{xx} + \big((1 - a_1)xr + a_1 xR - a_2\big)u_x + e^{-\rho t} F(a_2)\right\} = 0 \tag{4.20}$$

with the boundary conditions that

$$u(0,t) = 0, u(x,T) = 0. \tag{4.21}$$

We compute the maxima to find

$$\alpha^{1*} = \frac{-(R-r)u_x}{\sigma^2 x u_{xx}}, \quad F'(\alpha^{2*}) = e^{\rho t} u_x, \tag{4.22}$$

provided that the constraints $0 \leq \alpha^{1*} \leq 1$ and $0 \leq \alpha^{2*}$ are valid: we will need to worry about this later. If we can find a formula for the value function u, we will then be able to use (4.22) to compute optimal controls.

To go further (finding an explicit solution), we assume the utility function F has the explicit form

$$F(a) = a^\gamma \quad (0 < \gamma < 1)$$

Next we guess that our value function has the form

$$u(x,t) = g(t)x^\gamma,$$

for some function g to be determined. Then (4.22) implies that

$$\alpha^{1*} = \frac{-(R-r)u_x}{\sigma^2(1-\gamma)}, \quad \alpha^{2*} = [e^{\rho t} g(t)]^{\frac{1}{\gamma-1}} x.$$

Plugging our guess for the form of u into (4.20) and setting $a_1 = \alpha^{1*}, a_2 = \alpha^{2*}$, we find

$$\left(g'(t) + \nu\gamma g(t) + (1-\gamma)g(t)\left(e^{\rho t} g(t)\right)^{\frac{1}{\gamma-1}}\right)x^\gamma = 0$$

for the constant

$$\nu := \frac{(R-r)^2}{2\sigma^2(1-\gamma)} + r.$$

Now put

$$h(t) := \left(e^{\rho t} g(t)\right)^{\frac{1}{\gamma-1}}$$

to obtain a linear ODE for h. Then we find

$$g(t) = e^{-\rho t}\left[\frac{1-\gamma}{\rho - \nu\gamma}\left(1 - e^{\frac{-(\rho-\nu\gamma)(T-t)}{1-\gamma}}\right)\right]^{1-\gamma}.$$

If $R - r \leq \sigma^2(1-\gamma)$, then $0 \leq \alpha^{1*} \leq 1$ and $\alpha^{2*} \geq 0$ as required.

References

[1] B. D. Craven, Control and Optimization, Chapman & Hall, 1995.

[2] B. K. Oksendal, Stochastic Differential Equations: An Introduction with Applications, 4th ed., Springer, 1995.

[3] E. B. Lee and L. Markus, Foundations of Optimal Control Theory, Wiley, 1967.

[4] F. Clarke, Optimization and Nonsmooth Analysis, Wiley-Interscience, 1983.

[5] F. Clarke, Methods of Dynamic and Nonsmooth Optimization, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, 1989.

[6] G. Knowles, An Introduction to Applied Optimal Control, Academic Press, 1981.

[7] G. Oster and E. O. Wilson, Caste and Ecology in Social Insects, Princeton University Press.

[8] J. Lewin, Differential Games: Theory and methods for solving game problems with singular surfaces, Springer, 1994.

[9] J. Macki and A. Strauss, Introduction to Optimal Control Theory, Springer, 1982.

[10] L. C. Evans, An Introduction to Stochastic Differential Equations, lecture notes available at http://math.berkeley.edu/˜ evans/SDE.course.pdf.

[11] L. Hocking, Optimal Control: An Introduction to the Theory with Applications, Oxford University Press, 1991.

[12] L. S. Pontryagin, V. G. Boltyanski, R. S. Gamkrelidze and E. F. Mishchenko, The Mathematical Theory of Optimal Processes, Interscience, 1962.

[13] M. Bardi and I. Capuzzo-Dolcetta, Optimal Control and Viscosity Solutions of Hamilton- Jacobi-Bellman Equations, Birkhauser, 1997.

[14] N. Barron and R. Jensen, The Pontryagin maximum principle from dynamic programming and viscosity solutions to first-order partial differential equations,

[15] N. V. Krylov, Controlled Diffusion Processes, Springer, 1980.

[16] R. Isaacs, Differential Games: A mathematical theory with applications to warfare and pursuit, control and optimization, Wiley, 1965 (reprinted by Dover in 1999).

[17] W. Fleming and R. Rishel, Deterministic and Stochastic Optimal Control, Springer, 1975.

[18] W. Fleming and M. Soner, Controlled Markov Processes and Viscosity Solutions, Springer, 1993.