Sudan University of Science and Technology
College of Graduate Studies

# Turnpike Theory Properties in Optimal Mathematical Control with Applications

## خصائص نظرية مدخل الحاجز في التحكم الرياضي الأمثل مع التطبيقات

A Thesis submitted in Fulfillment Requirements for the Degree of Ph.D in Mathematics

**By:**

**Safa Mohammed Ahmed Mohammed**

**Supervisor:**

**Dr.Abdelrhim Bashir Hamid**

**December- 2020**

# Dedication

To My parents.

Who always support

To My Brothers and My Sister

Who always encourage me

To my best friends

Who always stand beside and help me

# Acknowledgments

Thank and praise is to Allah ,the almighty for giving health, and patient to complete this work. I would like to thank my Family especially my Parents for encouraging and supporting me all the time. Also I would like to take this opportunity to thank all my Teachers and my research supervisor, Dr. Abdelrhim Bashir Hamid for giving me the opportunity to work with her and guiding and helping me throughout this research and other courses. Very special thanks one due to my Co- Supervisor Dr. Belgis Abdelaziz for her encourage and support. For guiding and supporting me by her experience and valuable advices to accomplish this research.

# Abstract

Optimal control is a mathematical method to fined values for a systems variables, so that these values lead the system to follow an optimal path or curve that achieves the maximum or minimum values for a characteristic or cost function. The Turnpike phenomenon appears in several variational and optimal control problems, arising in engineering and economic growth. We say that a problem has a turnpike property when the optimal solutions converge to certain path during most of time, this path is known as the turnpike of problem. In this research we discussed number of recent results concerning turnpike properties in the calculus of variations and optimal control problems. The optimal trajectory is shown to remains exponentially close to the steady-state solution of an associated static optimal control problem ,but also corresponding adjoint vector of the Pontryagin maximum principle. We provide a general version of a turnpike theorem, valuable for nonlinear dynamics without any specific assumption ,and for very general terminal conditions. We characterized turnpike properties of the dynamics in terms of the system matrices related to the linear quadratic problem. These characterizations lead to new necessary conditions for the turnpike properties under consideration, and thus eventually to necessary and sufficient conditions in terms of spectral criteria and matrix inequalities.

## الخلاصة

يعتبر التحكم الأمثل طريقة رياضية لإيجاد قيم  لمتغيرات نظام ما بحيث تقود هذه القيم النظام لتتبع مسار او منحنى أمثل يحقق القيم القصوى او الدنيا لخاصية او دالة الكلفة. تظهر ظاهرة مدخل الحاجز في العديد من مسائل التحكم الأمثل والمتغيرات الناشئة في الهندسة والنمو الاقتصادي ،نقول أن المشكلة لها خاصية مدخل الحاجز عندما تتقارب الحلول المثلى الى مسار معين خلال معظم الوقت يعرف هذا المسار بأسم مسالة مدخل الحاجز.ناقشنا في هذا البحث عددا من النتائج الحديثة المتعلقة بخصائص مدخل الحاجز في حساب التغييرات ومسائل التحكم الأمثل. وتبين أن المسار الأمثل يظل قريبا بشكل كبير من حل الحالة المستقرة لمشكلة التحكم المثلى الثابتة المرتبطة ، أيضا يقابل المتجه المساعد لمبدأ بونترياجين الأقصى .لقد قدمنا نسخة عامة من نموذج نظرية مدخل الحاجز وهو ذو قيمة للديناميكيات غير الخطية دون أفتراض محدد للشروط النهائية العامة ،نحن نميز الخصائص الديناميكية من حيث مصفوفات النظام المتعلقة بالمسألة التربيعية الخطية ،تؤدي هذه الخصائص الي شروط ضرورية جديدة لخصائص مدخل الحاجز في ظل الظروف ، وبالتالي إلى الشروط الضرورية والكافية من حيث المعايير الطيفية ومتباينات المصفوفات.

# Introduction

In chapter one we study introduction content basic concept , in chapter two we study some basic in calculus of variations. Classical solutions to minimization problems in the calculus of variations are prescribed by boundary by value problem involving certain types differential equations , known as the associated Euler-Lagrange equations. The mathematical techniques that have been developed to handle such optimization problems are fundamental in many areas of mathematics ,physics , engineering ,and other applications. The history of the calculus of variation is tightly interwoven with the history of mathematics. The field has drawn the attention of a remarkable range of mathematical luminaries , beginning with Newton and Leibniz ,then initiated as a subject in its own right by Bernoulli brother Jakob and Johann. The first major developments appeared in the work of Euler, Lagrange, and Laplace. In the nineteenth century, Hamilton, Jacobi, Dirichlet, and Hilbert are but a few of the outstanding contributors. In modern times, the calculus of variations has continued to occupy center stage, witnessing major theoretical advances, along with wide-ranging applications in physics, engineering and all branches of mathematics. Minimization problems that can be analyzed by the calculus of variations serve to characterize the equilibrium configurations of almost all continuous physical systems, ranging through elasticity, solid and fluid mechanics, electro-magnetism, gravitation, quantum mechanics ,string theory, and many, many others. Many geometrical configurations, such as minimal surfaces, can be conveniently formulated as optimization problems.

In chapter three we study mathematical optimal control, control system design is generally a trial and error process in Classical which various methods of analysis are used iteratively to determine the design parameters of an "acceptable" system. Acceptable performance is generally defined in terms of time and frequency domain criteria such as rise time, settling time, peak overshoot, gain and phase margin, and bandwidth. Radically different performance criteria must be satisfied, however, by the complex, multiple-input, multiple-output systems required to meet the demands of modern technology. For example, the design of a spacecraft attitude control system that minimizes fuel expenditure is not amenable to solution by classical methods. A new and direct approach to the synthesis of these complex systems, called optimal control theory, has been made feasible by the development of the digital computer. The objective of optimal control theory is to determine the control signals that will cause a process to satisfy the physical constraints and at the same time minimize (or maximize) some performance criterion. Later, we shall give a more explicit mathematical statement of " the optimal control problem " ,but first let us consider the matter of problem formulation.

In chapter four we study turnpike theory and it is properties, Turnpike Theory refers to a set of economic theories about the optimal path of accumulation (often capital accumulation ) in a system , depending on the initial and final levels. In the context of a macroeconomic exogenous growth model.

In chapter five we study turnpike properties in calculus of variation and optimal control, In this chapter we survey our results of the turnpike property for some classes of variational and optimal control problems. To have property means the approximate solution of the problems are determined mainly by objective functions and essentially independent of the choice of interval and endpoint conditions except in regions close to the endpoint. We discuss necessary and sufficient conditions for turnpike properties of approximate solutions for variational problems and discrete-time optimal control problems. Turnpike properties have been established long time ago in finite-dimensional optimal control problem arising in econometry.

In chapter six we study main results an conclusion.

## The Contents

# List of Figure

# Chapter One

# Introduction

## 1. Background

Turnpike properties are well known in the field of mathematical economics . The term was first coined by Samuelson who showed that an efficient expanding economy would for most of the time be in the vicinity of a balanced equilibrium path (von Neumann path) .Turnpike properties are studied for optimal control problems on finite time intervals $[T_1, T_2]$ such that $T_1 < T_2$. Here $T_1$, $T_2$ are real numbers in the case of continuous –time problems and are integers in the case of discrete-time problems .Solutions of such problems (trajectories or path ) always depend on an optimality criterion usually determined by a cost function and some initial conditions .In the turnpike theory we study the structure of solutions when a cost function (an optimality criterion) is fixed while $T_1$, $T_2$ and the data vary . To have turnpike properties means ,that the solutions of a problem are determined mainly by the a cost function (optimality criterion ) , and are essentially independent of the choice of time interval and data , except in regions close to the endpoints of the time interval . If a real number $t$ does not belong to these regions ,then the value of a solution at the point $t$ is closed to a “turnpike” – a trajectory ( path) which is defined on the infinite time interval and depends only on the a cost function (optimality criterion ). This phenomenon has the following interpretation .If one wishes to reach a point A from a point B by a car in an optimal way , then one should enter onto a turnpike, spend most of one’s time on it and then leave the turnpike to reach the required point .P.A.Samuelson discovered the turnpike phenomenon in a specific situation in 1948 . In further studies turnpike results were obtained under certain rather strong assumptions on an optimality criterion .Usually it was assumed that an objective function is convex , as a function of all its variables and does not depend on the time variable $t$.In this case it was shown that the “turnpike” is a stationary trajectory (a singleton ). Since convexity assumptions usually hold for models of economic growth ,turnpike theory has many applications in mathematical economics .

## 1.1 Mathematical Control Theory[31]

Mathematical control theory is the area of application-oriented mathematics the deals with the basic principles underlying the analysis and design of control systems. To control an object means to influence its behavior so as to achieve a desired goal . In order to implement this influence, engineers build devices that incorporate various mathematical techniques .These devices range from Watt’s steam engine governor designed during the English Industrial Revolution ,to the sophisticated microprocessor controllers found in consumer items- such as players and automobiles or in industrial robots and airplane autopilots .While on the one hand one wants to understand the fundamental limitations that mathematics imposes on what is achievable , irrespective of that precise technology being used , it is also true that technology may well influence the type of question to be asked and the choice of mathematical model . An example of this is the use of difference rather than differential equations when one is interested in digital control .Roughly speaking ,there have been two main lines of work in control theory , which sometime have seemed to proceed in very different directions but which are in fact complementary. One of these is based on the ideas that a good model of the object to be controlled is available and that one wants to some now optimize its behavior . For instance physical principles and engineering specifications can be and are used in order to calculate that trajectory of a spacecraft which minimizes total travel time or fuel consumption . The techniques here are closely related to the classical calculus of variations and to other areas of optimization theory , the end result is typically a preprogrammed fight plan .the other main line of work is that based on the constraints imposed by uncertainty about the model or about the environment in which the object operates . The central tool here is the use of feedback in order to correct for deviations from the desired behavior . For instance , various feedback control systems are used during actual space flight in order to compensate for errors from the precomputed trajectory mathematically , stability theory , dynamical systems , and especially the theory of functions of complex variable, have had a strong influence on this approach . It is widely recognized today that these two broad lines of work deal just with different aspects of the some problems.

## 1.2 Control Systems

**Definition of Systems 1.1[21]**

A system is a combination of components that act together and perform a certain objective . A system need not be physical .The concept of the system can be applied to abstract , dynamic phenomena such as those encountered economic .The word system should , therefore, be interpreted to imply physical , biological , economic , and the like , systems.

**Definition of Control System 1.2[2]**

A control system is an arrangement of physical components connected or related in such manner as to command ,direct ,or regular itself or another system.

Control systems abound in our environment , but before exemplifying, we define to terms: input and output , which help in identifying delineating , or defining a control system.

**Definition of input 1.3**[2]

The input is the stimulus , excitation or command applied to a control system ,typically form an external energy source.

**Definition of output 1.4[2]**

The output is the actual response obtained from a control system . It may or may not be equal to the specified response implied by the. Input and out can have many different forms inputs for examples may be physical variables or more abstract quantities such as reference , set point or desired values for the output of the control system.

**Definition 1.5 :**

**Controlled Variable and Control Signal or Manipulated Variable [21]**

The controlled variable is the quantity or condition that is measured and controlled . the control signal or manipulated variable is the quantity or condition that is varied by the controller so as to affect the value of the controlled variable .Normally , the controlled variable is the output of the system .Control means measuring the value of the controlled variable of the system and applying the control signal to the system to correct or limit deviation of the measured value form a desired value.

**Definition plants 1.6 [21]**

A plant may be a piece of equipment , perhaps just a set of machine parts functioning together ,the purpose of which is perform a particular operation.

**Definition Feedback Control 1.7**[21]

Feedback control refers to an operation that , in the presence of disturbances, tends to reduce the difference between the output of a system and some reference input and does so on the basis of this difference .Here only unpredictable disturbances can always be compensated for within the system.

## 1.3 Closed Loop Control Versus Open Loop Control[2]

### 1.3.1 Feedback Control Systems :

A system that maintains a prescribed relationship between the output and the reference input by comparing then and using difference as a means of control is called a feedback control system. An example would be a room temperature control system .Feedback control systems are not limited engineering but can be found in various non engineering fields as well . The human body ,for instance, is highly advanced feedback control system. Both body temperature and blood pressure are kept constant by means of physiological feedback . In fact , feedback performs a vital function :It makes the human body relatively insensitive to external disturbances , thus enabling it to function properly in a changing environment.

**Definition of open loop control 1.8**

An open-loop control system is one in which the control action is independent of the output.

**Definition of closed loop control 1.9**

A closed-loop control system is one in which the control action is somehow dependent on the output.

## 1.4 Analog and Digital Control Systems[2]

The signal in control system ,for example , the input and output waveforms , are typically functions of some independent variable , usually time , denoted $t$ .

**Definition 1.10:**

A signal dependent on a continuum of values of the independent variable $t$ is called a continuous –time signal or , more generally , a continuous –date or (less frequently ) an analog signal.

**Definition1.11**

A signal defined at ,or interest at ,only discrete (distinct ) instants of the independent variable $t$ (upon which it depends). Is called discrete-time , a discrete date , a sampled –date ,or a digital signal .

**Definition 1.12**

Continuous –time control systems also called continuous-date control systems also called continuous-data control systems or , analog control systems , contain or process only continuous –time (analog ) signals and components.

**Definition 1.13**

Discrete-time control systems , also called discrete –data control systems , or sampled data control systems , have discrete –time signals or components at one or more points in the systems .

## 1.5 Control System Models or Representations[2]

To solve a control systems problem we must put the specifications or description of the system configuration and its components into a form amenable to analysis or design . Three basic representations (models) components and systems are used extensively in the study of control systems

1- Mathematical models , in the form of differential equations ,difference equations ,and / or other mathematical relations ,for example , Laplace
2- Block diagrams

**1.5.1 Mathematical Models** [21]:

Mathematical models may assume many different forms depending on the particular system and the particular circumstances , one mathematical model may better suited than other models. For example, in optimal control problems it is advantageous to use state- space representations . On other hand , for the transient-response or frequency-response analysis of single input , single output ,linear , time – invariant systems , the transfer function representation may be more convenient than any other .Once a mathematical model of a system is obtained , various analytical and computer tools can be used for analysis and synthesis purposes .

**1.5.2 Linear Systems** [21]

A system is called linear if the principle of superposition applies. The principle of superposition states that the response produced by the simultaneous application of two different forcing functions the sum of the two individual responses .Hence ,for the linear system ,the response to several inputs can be calculated by treating one input at a time and adding the results . It is this principle that allows one to build up complicated solutions to the linear differential equation from simple solutions .In an experimental investigation of a dynamic system , if cause and effect are proportional , thus implying that the principle of superposition holds , then the system can be considered linear .

### 1.5.3 Linear Time-Invariant Systems and Linear Time- Varying Systems[21]

A differential equation is linear if the coefficients are constants or functions only of the independent variable . Dynamic systems that are composed of linear time – Invariant lumped-parameter components may be described by linear time-invariant differential equations –that is ,constant- coefficient differential equations. Such systems are called linear time – Invariant systems .Systems That are represented by differential equations whose coefficients are functions of time are called linear time varying systems .

### 1.5.4 Nonlinear Systems[21]

A system is nonlinear if the principle of superposition does not apply .Thus , for a nonlinear system the response to two inputs cannot be calculated by treating on input at a time and adding the results .

Example of nonlinear differential equations are

$$\frac{d^2x}{dt^2} + \left(\frac{dx}{dt}\right)^2 + x = A\sin wt$$

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

Although many physical relationships are often represented by linear equations, in most cases actual relationships are not quite linear , in most cases actual relationships are not quite linear ,in fact ,a careful study of physical systems reveals that that even so –called “linear system” are really linear only in limited operating ranges , in practice , many electromechanical systems , hydraulic systems , pneumatic systems, and so on , involve nonlinear relationships among , the variable .For example , the output of a component may saturate for large input signals . There may be a dead space that affects small signals. Square-law nonlinearity may occur in some components . For instance , dampers used in physical systems may be linear for law –velocity operations but may become nonlinear at high velocities , and the damping force may become proportional to the square of the operating velocity .Note that some important control systems , are nonlinear for signals of any size. For example in on-off control systems , the control action is either on or off ,and there is on linear relationship between the input and output of the controller . Procedure for finding the solutions of problems involving such nonlinear systems , in general , are extremely complicated . Because of this mathematical difficulty attached to nonlinear systems , one often finds it necessary to introduce "equivalent" linear systems in place of nonlinear once . such equivalent linear systems are valid for only a limited range of operation. Once a nonlinear system is approximated by linear mathematical model , a number of linear tools may be applied for analysis and design purposes.

## 1.6 Transfer Function and Impulse Response Function[21]

### 1.6.1 Transfer Function

The transfer function of a liner ,time- invariant differential equation system is defined as the ratio of the Laplace transform of the output (response function) to the Laplace transform of the input (deriving function) under the assumption that all initial conditions are zero. Consider the linear time-invariant system defined by the following differential equation£

$$a_0 y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1}\dot{y} + a_n y$$

$$= b_0 x^{(m)} + b_1 x^{(m-1)} + \dots + b_{m-1}\dot{x} + b_m x \qquad (n \geq m)$$

Where $y$ is the output of the system and $x$ is the input .The transfer function of this system is the ratio of the Laplace transform output to the Laplace transform input when all initial conditions are zero ,or

$$\text{Transfer Function}=\left.\frac{\mathcal{L}[output]}{\mathcal{L}[input]}\right|_{zero\ initial\ conditions}$$

$$=\frac{Y(s)}{X(s)}=\frac{b_0s^m+b_1s^{m-1}+\cdots+b_{m-1}s+b_m}{a_0s^n+a_1s^{n-1}+\cdots+a_{n-1}s+a_n}$$

By using the concept of transfer function , it possible to represent system dynamics by algebraic equations in $s$ .If the highest power of $s$ in the denominator of the transfer function is equal to $n$ , the system is called an $nth$- order system.,

### 1.6.2 Convolution Integral

For a linear ,time-invariant system the transfer function $G(x)$ is

$$G(s)=\frac{Y(s)}{X(s)} \qquad (1.1)$$

Where $X(s)$ is Laplace transform of the input to the system and $Y(s)$ is the is Laplace transform of the output of the system , where we assume that all initial conditions involved are zero . It follows that the output $Y(s)$ can be written as the product of $G(s)$ and $X(s)$ , or

$$Y(s)=G(s)X(s)$$

Note that the multiplication in the complex domain is equivalent to convolution in the time domain , so the inverse Laplace transform of equation ( 1.1 ) is given by the following convolution integral:

$$y(t)=\int_0^t x(\tau)g(t-\tau)d\tau$$

$$=\int_0^t g(\tau)x(t-\tau)d\tau$$

Where both $g(t)$ and $x(t)$ are 0 for $t<0$ .

### 1.6.3 Impulse Response Function

Consider the output (response) of a linear time invariant system to unit – impulse input when the initial conditions are zero . Since The Laplace transform of the unit – impulse function is unity ,the Laplace transform of output of the system is

$$Y(s) = G(s) \qquad (1.2)$$

The Laplace transform of output given by equation (1.2 ) gives the impulse response of the system. The inverse Laplace transform of $G(s)$ . or

$$\mathcal{L}^{-1}[G(s)] = g(t)$$

Is called the impulse-response function. The function $g(t)$called the weighting function of the system.

## 1.7 The Laplace Transform

The Laplace transform relates time functions to frequency-dependent functions of complex variable.

**Definition1.14**

Let $f(t)$ be a real function of a real variable $t$ defined for $t > 0$.Then

$$\mathcal{L}[f(t)] = f(s) = \lim_{\substack{T\to\infty \\ \epsilon\to\infty}} \int_{\epsilon}^{T} f(t)e^{-st}dt = \int_{0^+}^{\infty} f(t)e^{-st}dt \quad 0 < \epsilon < T$$

Is called the Laplace transform of $f(t)$ . S is a complex variable defined by $s = \sigma + j\omega$ , where $\sigma$ and $\omega$ are a real variables and $j = \sqrt{-1}$ .

**Definition1.15**

If $f(t)$ is defined and single-valued for $t > 0$ and $F(\sigma)$ is absolutely convergent for some real number $\sigma_0$ , that is ,

$$\int_{0^+}^{\infty} |f(t)| e^{-\sigma_0 t} dt = \lim_{\substack{T\to\infty \\ \epsilon\to 0}} |f(t)| e^{-\sigma_0 t} dt < +\infty \qquad 0 < \epsilon < T$$

Then $f(t)$ is Laplace transformable for $Re(s) > \sigma_0$ .

**Example 1.1**

The Laplace transform of $e^{-t}$ is

$$\mathcal{L}[e^{-t}] = \int_{0^+}^{\infty} e^{-t}\, e^{-st} dt = \frac{-1}{(s+1)} e^{-(s+1)t}\Big|_{0^+}^{\infty} = \frac{1}{s+1} \text{ for } Re(s) > -1.$$

### 1.7.1 The Inverse Laplace Transform

The Laplace transforms a problem from the real variable time domain into the complex variable S-domain . After a solution of the transformed problem has been obtained in terms of $s$ ,it is necessary to “invert” this transform to obtain the time domain solution . The transformation from the s-domain into the t-domain is called the inverse Laplace transform.

**Definition1.16**

Let $F(s)$ be the Laplace transform of a function $f(t)$ , $t > 0$ . The contour integral

$$\mathcal{L}^{-1}[F(s)] \equiv f(t) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} F(s) e^{st}\, ds$$

Where $j = \sqrt{-1}$ and $c > \sigma_0$ is called the inverse Laplace transform.

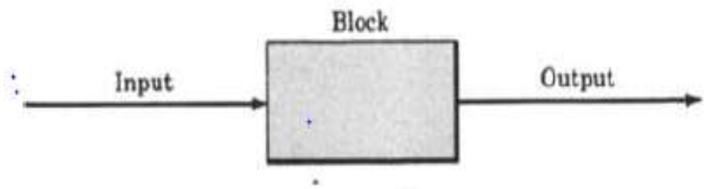
**Example 1.2**

The inverse Laplace transform of the functions $1/s+1$ and $1/s+3$ are

$$\mathcal{L}^{-1}\left[\frac{1}{s+1}\right] = e^{-t} \qquad , \qquad \mathcal{L}^{-1}\left[\frac{1}{s+3}\right] = e^{-3t}$$

## 1.8 Block Diagrams: Fundamentals[21]

A block diagram is shorthand , pictorial representation of the cause and- effect relationship between the input and output of a physical system. The simplest form of the block diagram is single block. With one input and one output ,as shown fig (1.1).




**Fig(1.1)**:block diagram is single input and single output

### 1.8.1 Summing Point

A circle with across in the symbol that indicates a summing operation . The plus or minus sign at each arrowhead indicated whether that signal is to be added or subtracted . It is important that the quantities being added or subtracted have the same dimensions and the same units.

### 1.8.2 Branch Point

A branch point is a point from which the signal from a block goes concurrently to other blocks or summing points .

### 1.8.3 Block Diagram of a Closed Loop System

Figures 1.2 shows an example of a block diagram of a closed –loop system . The output $C(s)$ is feedback to summing point , where it is compared with the reference input $R(s)$ . The output of the block , $C(s)$ in this case , is obtained by multiplying the transfer function $G(s)$ by the input to the block , $E(s)$. Any linear control system may be represent by a block diagram consisting of a blocks , summing points ,and branch points . When the output is fed bock to the summing point for comparison with the input . it is necessary to convert the form of the output signal to that of the input signal. For example , in a temperature control system , the output signal is usually the controlled temperature . The output signal , which has the dimension of temperature , must be converted to a force or position or voltage before it can be compared with the input signal. This conversion is accomplished by the feedback element whose transfer function is $H(s)$, as shown in figure 1.3 . The role of the feedback element is to modify the output before it is compared with the input.

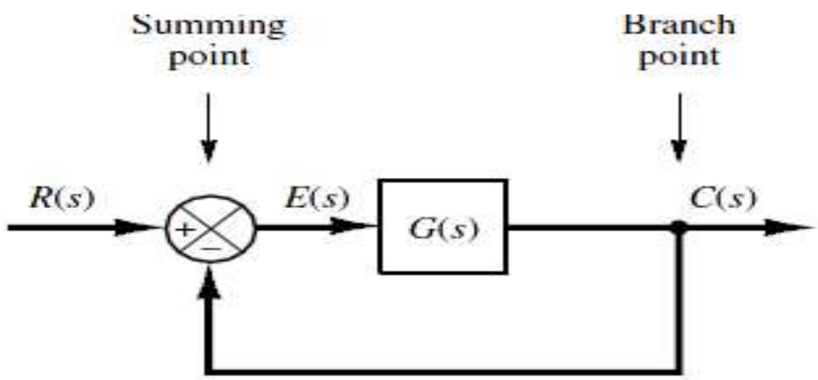



**Figure 1.2 :**Block diagram of a closed-loop system

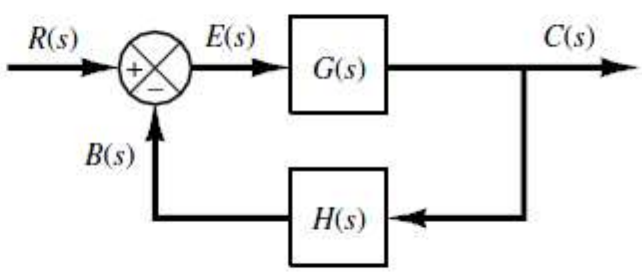



**Figure 1.3**: Closed-loop system

In the present example the feedback signal that the summing point for comparison with the input is $B(s) = H(s)C(s)$.

### 1.8.4 Open-Loop Transfer Function and Feedforward Transfer Function

Referring to figure 1.3 the ratio of the feedback signal $B(s)$ to the actuating error signal $E(s)$ is called the open-loop transfer function. That is

$$\text{Open –loop transfer function} = \frac{B(s)}{E(s)} = G(s)H(s)$$

The ratio of the output $C(s)$ to the actuating error signal $E(s)$ is called the feedforward function, so that

$$\text{Feedforward transfer function} = \frac{C(s)}{E(s)} = G(s)$$

If the feedback transfer function $H(s)$is unity, then the open –loop transfer function and the feedforward transfer function are same .

### 1.8.5 Closed Loop Transfer Function

For the system shown in figure 1.3 the output $C(s)$ and input $R(s)$ are related as follows: since

$$C(s) = G(s)E(s)$$

$$E(s) = R(s) - B(s)$$

$$= R(s) - H(s)C(s)$$

Eliminating $E(s)$ from these equations gives

$$C(s) = G(s)[R(s) - H(s)C(s)]$$

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1+G(s)H(s)} \qquad (1.\ 3)$$

The transfer function relation $C(s)$ to $R(s)$ is called the closed-loop system dynamic to the feedforward elements and feedback element . From equation (1.3 ) , $C(s)$ is given by

$$C(s) = \frac{G(s)}{1 + G(s)H(s)} R(s)$$

Thus the output of the closed loop systems clearly depends on the both the closed-loop transfer function ad nature of the input .

## 1.9 Modeling in State Space[21]

### 1.9.1 Modern Control Theory

The modern trend in engineering systems is toward greater complexity ,due mainly to the requirements of complex tasks and good accuracy . Complex systems may have multiple inputs and multiple outputs and may be time varying . Because of the necessity of meeting increasingly stringent requirements on the performance of control systems ,the increase in system complexity ,and easy access to large scale computers .This new approach is based on the concept of state.

**Definition State 1.19**

The state of a dynamic system is the smallest set of variables (called state variables) such that knowledge of these variable at $t = t_0$,together with knowledge of the input for $t \geq t_0$,completely determines the behavior of the system for any time $t \geq t_0$. Note that the concept to state is by no means limited to physical systems. It is applicable to biological systems , economic systems , social systems , and others .

**Definition State Variables 1.20**

The state variables a dynamic systems are the variables making least $n$ variables $x_1, x_2, \dots, x_n$ are needed to completely describe the behavior of a dynamic system ( so that once the input is given for $t \geq t_0$ and the initial state at $t = t_0$ is specified , the future state of the system is completely determined ) ,then such $n$ variables are a set of state variables . Note that state variables need do not be physically measurable or observable quantities . Variables that do not represent physical quantities and those that are neither measurable nor observable can be chosen as state variables . Such freedom in choosing state variables is an advantage of the state-space methods . Practically , however ,it is convenient to choose easily measurable quantities for state variables ,if this is possible at all , because optimal control laws will require the feedback of all state variables with suitable weighting.

**Definition State Vector 1.21 :**

If $n$ state variables are need to completely describe the behavior of a given system , then $n$ state variables can be considered the $n$ components of a vector$x$. Such a vector is called state vector . A state vector is thus a vector that determines uniquely the system state $x(t)$ for any time $t \geq t_0$ ,once the stats at $t = t_0$ ,is given and the input $u(t)$ for $t \geq t_0$ is specified .

**Definition State Space 1.22 :**

The $n-$dimensional space whose coordinate axes consist of the $x_1 axis, x_2 axis, \dots, x_n axis$ are state variables , is called a state space . Any state can be represented by a point in the state space .

**Definition State Space Equations 1.23**

In state- space analysis are concerned with three types of variable that are involved in the modeling of dynamic systems: input variable ,output variable , and state variable .The state- space representation for a given system is not unique , except that the number of state variables is the same for any of the different state- space representations of same system. Assume that a multiple- input ,multiple-output system involves $n$ integrator . Assume also that there are $r$ inputs $u_1(t), u_2(t), \dots, u_r(t)$ and $m$ output $y_1(t), y_2(t), \dots, y_m(t)$. Define $n$ output of the integrators as stat variable : $x_1(t), x_2(t), \dots, x_n(t)$ then the system may be described

$$\begin{aligned}
\dot{x}_1(t) &= f_1(x_1, x_2, \dots, x_n;\ u_1, u_2, \dots u_r; t) \\
\dot{x}_2(t) &= f_2(x_1, x_2, \dots, x_n;\ u_1, u_2, \dots u_r; t) \qquad (1.4) \\
&\vdots \\
\dot{x}_n(t) &= f_n(x_1, x_2, \dots, x_n;\ u_1, u_2, \dots u_r; t)
\end{aligned}$$

The output $y_1(t), y_2(t), \dots, y_m(t)$ of the system may be given by

$$\begin{aligned}
y_1(t) &= g_1(x_1, x_2, \dots, x_n;\ u_1, u_2, \dots u_r; t) \\
y_2(t) &= g_2(x_1, x_2, \dots, x_n;\ u_1, u_2, \dots u_r; t) \qquad (1.5) \\
&\vdots \\
y_m(t) &= g_m(x_1, x_2, \dots, x_n;\ u_1, u_2, \dots u_r; t)
\end{aligned}$$

If we define

$$x(t)=\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} \quad , \quad f(x,u,t)=\begin{bmatrix} f_1(x_1,x_2,...,x_n;u_1,u_2,...,u_r;t) \\ f_2(x_1,x_2,...,x_n;u_1,u_2,...,u_r;t) \\ \vdots \\ f_n(x_1,x_2,...,x_n;u_1,u_2,...u_r;t) \end{bmatrix},$$

$$y(t)=\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix} \quad , \quad g(x,u,t)=\begin{bmatrix} g_1(x_1,x_2,...,x_n;u_1,u_2,...,u_r;t) \\ g_2(x_1,x_2,...,x_n;u_1,u_2,...,u_r;t) \\ \vdots \\ g_n(x_1,x_2,...,x_n;u_1,u_2,...,u_r;t) \end{bmatrix} , \; u(t)=\begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_r(t) \end{bmatrix} ,$$

The equations (1.4) and (1.5) become

$$\dot{x}(t) = f(x,u,t) \tag{1.6}$$

$$y(t) = g(x,u,t) \tag{1.7}$$

Where equation (1.6) is the state equation and equation (1.7) is the output equation . If vector function $f$ and / or involve time $t$ explicitly , then the system is called a time varying system . If equation (1.6) and(1.7) are linearized about the operating state , then we have the following linearized state equation and output equation:

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{1.8}$$

$$y(t) = C(t)x(t) + D(t)u(t) \tag{1.9}$$

Where $A(t)$ is called the state matrix , $B(t)$ the input matrix , $C(t)$ the output matrix and the $D(t)$ the direct transmission matrix. If vector function $f$ and $g$ do not involve time explicitly then the system is called a time invariant system . In this case ,equations (1.8) and (1.9) can be simplified to

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{1.10}$$

$$\dot{y}(t) = Cx(t) + Du(t) \tag{1.11}$$

Equation (1.10) is state equation of the linear ,time – invariant system and equation (1.11) is the output equation for the same system.

## 1.10 Fundamental Concepts of the Calculus of Variations[19]

### 1.10.1 Function and Functional

**(a) Function :** A variable $x$ is a function of variable quantity $t$ (written $x(t) = f(t)$) , if to every value $x$ ; i.e. , we have a correspondence : to number $t$ there corresponds a number $x$ . Note that here $t$ need not be always time but any independent variable.

**(b) Functional**: A variable quantity J is a functional dependent a function $f(x)$ , written as $J = J(f(x))$. If to each function $f(x)$ , there corresponds a number $J$ . Functional depend on several functions .

**Example 1.3**

Let $x(t) = 2t^2 + 1$ . Then

$$J(x(t)) = \int_0^1 (x(t))\,dt = \int_0^1 (2t^2 + 1)dt = \frac{5}{3} \qquad (1.12)$$

Is the area under the curve $x(t)$.If $v(t)$ velocity of a vehicle , then

$$J\big(v(t)\big) = \int_{t_0}^{t_f} v(t)dt \qquad (1.13)$$

Is the path traversed by the vehicle . Thus , here $x(t)$ and $v(x)$ are functions of $t$ ,and $J$ is a functional of $x(t)$ and $v(t)$ .

### 1.10.2 Increment

We consider here increment of a function and a functional .

**(a) Increment of a function :** In order to consider optimal values of a function , we need the definition of an increment

**Definition1.24**

The increment of the function $f$, denoted by $\Delta f$ is defined as

$$\Delta f \triangleq f(t + \Delta t) - f(t)$$

Form the definition that $\Delta f$ depends on both the independent variable $t$ and the increment of the independent variable $\Delta t$, and hence strictly speaking , we need to write the increment of a function as $\Delta f(t, \Delta t)$ .

**Example 1.4**

$$\text{If } f(t) = (t_1 + t_2)^2$$

Find the increment of the function $f(t)$

**Solution**

$$\Delta f \triangleq f(t + \Delta t) - f(t)$$

$$= (t_1 + \Delta t_1 + t_2 + \Delta t_2)^2 - (t_1 + t_2)^2$$

$$= (t_1 + \Delta t_1)^2 + (t_2 + \Delta t_2)^2 + 2(t_1 + \Delta t_1)(t_2 + \Delta t_2) - (t_1^2 + t_2^2 + 2t_1 t_2) .$$

$$= (t_1 + t_2)\Delta t_1 + 2(t_1 + t_2)\Delta t_2 + (\Delta t_1)^2 + (\Delta t_2)^2 + 2\Delta t_1 \Delta t_2.$$

**(b) Increment of a functional**

**Definition1.25**

The increment of the functional $J$ ,denoted by $\Delta J$ , is defined as

$$\Delta J \triangleq J(x(t) + \delta x(t)) - J(x(t)) \qquad (1.14\ )$$

Here $\delta x(t)$ is called the variation of the function $x(t)$ . Since the increment of a functional is dependent upon the function $x(t)$ and its variation $\delta\big(x(t)\big)$ , strictly speaking , we need to write the increment as $\Delta J(x(t), \delta(x(t))$ .

**Example1.5**

Fined the increment of the functional

$$J = \int_{t_0}^{t_f} [2x^2(t) + 1]\, dt \qquad (1.15\ )$$

The increment of $J$ is given by

$$\Delta J \triangleq (x(t) + \delta x(t) - J(x(t)) \ ,$$

$$= \int_{t_0}^{t_f} [2(x(t) + \delta x(t))^2 + 1]\, dt - \int_{t_0}^{t_f} [2x^2(t) + 1]\, dt,$$

$$= \int_{t_0}^{t_f} [4x(t)\delta x(t) + 2(\delta x(t))^2]\, dt\ . \qquad (1.16)$$

**1.10.3 Differential and Variation**

**(a)Differential of a function**

Let as define at a point $t^*$ the increment of the function $f$

$$\Delta f \triangleq f(t^* + \Delta t) - f(t^*) \qquad (1.17)$$

By expanding $f(t^* + \Delta t)$ in a Taylor series a bout $t^*$ we get

$$\Delta f = f(t^*) + \left(\frac{df}{dt}\right)_* \Delta t + \frac{1}{2!}\left(\frac{d^2 f}{dt^2}\right)_* (\Delta t)^2 + \cdots - f(t^*) \qquad (1.18)$$

Neglecting the higher order terms in $\Delta t$ .

$$\Delta f = \left(\frac{df}{dt}\right)_* \Delta t = \dot{f}\,(t^*)\Delta t = df \qquad (1.\ 19)$$

Here , $df$ is called the differential of $f$ at the point $(t^*)$ . $\dot{f}(t^*)$ is the derivative or slope of $f$ at $t^*$ . the differential $df$ is the first order approximation to increment $\Delta t$.

**(b)Variation of functional**

Consider the increment a functional

$$\Delta J \triangleq J(x(t) + \delta x(t)) - J(x(t)) \qquad (1.20)$$

Expanding $J(x(t) + \delta x(t))$ in a Taylor series , we get

$$\Delta J = J(x(t) + \delta x(t) + \frac{1}{2!}\frac{\partial^2 J}{\partial x^2}\big(\delta x(t)\big)^2 + \cdots - J(x(t))$$

$$= \frac{\partial J}{\partial x}\delta x(t) + \frac{1}{2!}\frac{\partial^2 J}{\partial x^2}\big(\delta x(t)\big)^2 + \ldots$$

$$= \delta J + \delta^2 J + \cdots \qquad (1.21)$$

Where

$$\delta J = \frac{\partial J}{\partial x}\delta x(t) \qquad \text{and} \qquad \partial^2 J = \frac{1}{2!}\frac{\partial^2 J}{\partial x^2}\big(\delta x(t)\big)^2 \qquad (1.22)$$

Are called the first variation and the second variation of the functional $J$ , respectively . the variation $J$ is linear part of the increment $\Delta J$

### 1.10.4 Linearity of Functionals

**Definition 1.26:**

$f$ is a linear function of $q$ if and only if it satisfies the principle of homogeneity

$$f(\alpha q)) = \alpha f(q)$$

For all q $\in \wp$ . and for all real numbers $\alpha$ such that $\alpha q \in \wp$ ,and the principle of additivity

$$f\left(q^{(1)} + q^{(2)}\right) = f\left(q^{(1)}\right) + f(q^{(2)}) \qquad (1.23)$$

For all $q^{(1)}$ , $q^{(2)}$ and $q^{(1)}+ q^{(2)}$ in $\wp$

**Definition1.27**

$J$ is a linear functional of $x$ if and only if it satisfies the principle of

$$J(\alpha x) = \alpha J(x) \qquad (1.24)$$

For all $x \in \Omega$ and for all real numbers $\propto$ such that $\alpha x \in \Omega$ and the principle of additivity

$$J\left(x^{(1)} + x^{(2)}\right) = J\left(x^{(1)}\right) + J\left(x^{(2)}\right) \qquad (1.25)$$

For all $x^{(1)}$ , $x^{(2)}$ and $x^{(1)}+ x^{(2)}$ in $\Omega$ .

### 1.10.5 Closeness of Functions :

**Definition 1.28**

The norm of a function is a rule of correspondence assigns to each function $x \in \Omega$ , defined for $t \in [t_0 , t_f]$ , a real number . The norm of $x$ , denoted by $\|x\|$ , satisfies the following properties :

1- $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x(t) = 0$ for all $t \in \left[t_0, t_f\right]$ .

2- $\|\propto x\| = |\propto| . \|x\|$ for all real number $\propto$

3- $\left\|x^{(1)} + x^{(2)}\right\| \leq \left\|x^{(1)}\right\| + \left\|x^{(2)}\right\|$ . (1.26)

**Definition 1.29**

The functional $J[y]$ is said to be continuous at point $\hat{y} \in \mathcal{R}$ if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|J[y] - J[\hat{y}]| < \varepsilon$$

Provided that $\|y - \hat{y}\| < \delta$.

**1.11 Optimal Control** [10]

**Definition 1.30**

Let I ⊆ ℝ be an interval (finite or infinite) .We say a finite-valued function $u: I \rightarrow \mathbb{R}$ is piecewise continuous if it is continuous at each $t \in$ I , with the possible exception of at most a finite number of $t$, and if u is equal to either its left or right limit at every $t \in$ I .

**Definition 1.31**

Let $x: I \rightarrow \mathbb{R}$ be continuous on I and differentiable at all but finitely points of I .Further, suppose that $x'$ is continuous wherever it is defined . Then , we say $x$ is piecewise differentiable .

**Definition1.32**

Let $k$: I → ℝ .We say is continuously differentiable if $k'$ exists and is continuous on I .

**Definition1.33**

A function $k$ is called Lipschitz if there exists a constant $c$ (particular to $k$) such that $|k(t_1) - k(t_2)| \leq c|t_1 - t_2|$ for all points $t_1, t_2$ in the domain of $k$. The constant $c$ is called the Lipschitz constant of $k$ .

**Definition1.34**

A control history which satisfies the control constraints during the entire time interval $[t_0, t_1]$ is called an admissible control .

Suppose $u: I \to \mathbb{R}$ is piecewise continuous . Let $g: \mathbb{R}^3 \to \mathbb{R}$ be continuous in three variables . Then ,by the solution $x$ of differential equation

$$x'(t) = g(t, x(t), u(t)) \qquad (1.27)$$

In our basic optima control problem for ordinary differential equations , we use $u(t)$ for the control and $x(t)$ for the state . The state variable satisfies a differential equation which depends on the control variable :

$$x'(t) = g\big(t, x(t), u(t)\big)$$

As the control function is changed, the solution to the differential equation will change. Thus, we can view the control-to-state relationship as a map $u(t) \mapsto x = x(u)$. Our basic optimal control problem consists of finding a piecewise continuous control $u(t)$ and the associated state variable $x(t)$ to maximize the given objective functional , i.e.,

$$\max_u \int_{t_0}^{t_1} f\big(t, x(t), u(t)\big)dt$$

Subject to $x'(t) = g\big(t, x(t), u(t)\big)$

$$x(t_0) = x_0 \text{ and } x(t_1) \text{ free .} \qquad (1.28)$$

Such a maximizing control is called an optimal control . By $x(t_1)$ free , it is meant that the value of $x(t_1)$ is unrestricted .For our purposes , $f$ and $g$ will always be continuously differentiable functions , the associated states will always be piecewise differentiable . A piecewise continuous solution $u$ is called optimal control and the solution $x$ of the associated system of boundary value problems is said to be optimal trajectory.

**Definition 1.35**

If the optimal control is determined as a function of time for a specified initial state value, that is,

$$u^*(t) = e(x(t_0), t)$$

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

Thus the optimal open-loop control is optimal only for a particular initial state value, whereas, if the optimal control law is known, the optimal control history starting from any state value can be generated.

## 1.12 Eigenvalue Problems[17]

### 1.12.1 An Introduction to Coupled System

The simplest example is a system of linear differential equations of the from

$$\frac{dx}{dt} = ax + by \tag{1.29}$$

$$\frac{dy}{dt} = cx + dy$$

We note that this system is coupled. We cannot solve

We note that this system is coupled. We cannot solve either equation without knowing either x(t) or y(t). A much easier problem would be to solve an uncoupled system like

$$\frac{dx}{dt} = \lambda_1 x$$

$$\frac{dy}{dt} = \lambda_2 y$$

The solutions are quickly found to be

$$x(t) = c_1 e^{\lambda_1 t}$$

$$y(t) = c_2 e^{\lambda_2 t}$$

Here, $c_1$ and $c_2$ are two arbitrary constants.

We can determine particular solutions of the system by specifying $x(\, t_0) = x_0$ and $y(t_0) = y_0$ at some time $t_0$ .Thus,

$$x(t) = x_0 e^{\lambda_1 t}$$

$$y(t) = y_0 e^{\lambda_1 t}$$

We write the coupled system as

$$\frac{d}{dt} x = Ax$$

And the uncoupled system as

$$\frac{d}{dt} y = \Lambda y$$

Where

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Is a diagonal matrix.

Now, we seek a transformation between **x** and **y** that will transform the coupled system into the uncoupled system. Define the sought transformation as

$$x = sy$$

Inserting this transformation into the coupled system, we have

$$\frac{d}{dt}Sy = ASy$$

Because $S$ is a constant matrix ,

$$\frac{d}{dt}y = ASy \qquad (1.30)$$

If $S$ is invertible ,then $y = S^{-1}x$ .So , we can multiply both sides of equation (1.30) by $S^{-1}$ to obtain

$$\frac{d}{dt}y = S^{-1}ASy$$

Because we are seeking an uncoupled system of the form

$$\frac{d}{dt}y = \Lambda \mathrm{y}$$

We will require that

$$S^{-1}AS = \Lambda \qquad (1.31)$$

The expression $S^{-1}AS$ is called a similarity transformation of matrix A.

Multiplying equation (1.31) by $S^{-1}$ , we have

$$AS = S\Lambda$$

In particular, the columns of $S$ (denoted v) satisfy equations of the form

$$A\mathrm{v} = \lambda \mathrm{v}$$

for each $\lambda$ on the diagonal of $\Lambda$. This is an equation for vectors **v** and numbers $\lambda$ given matrix $A$. This is called an eigenvalue problem. The vectors are called eigenvectors and the numbers $\lambda$ are called eigenvalues. In principle,we can solve the eigenvalue problem and this will lead us to solutions of the uncoupled system of differential equation.

### 1.12.2 Eigenvalue problem

We seek nontrivial solutions to the eigenvalue problem

$$A\mathrm{v} = \lambda \mathrm{v} \qquad (1.32)$$

We are given the matrix $A$ is the $2 \times 2$ matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ .the eigenvalue problem Equation (1.32) takes the form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Multiplying the matrices, we obtain the homogeneous algebraic system

$$(a - \lambda)v_1 + bv_2 = 0$$

$$cv_1 + (d - \lambda)v_2 = 0.$$

The solution of such a system would be unique if the determinant of the system is not zero. However, this would give the trivial solution $v_1 = 0, v_2 = 0$, To get a nontrivial solution, we need to force the determinant to be zero:

$$0 = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc$$

This is a quadratic equation for the eigenvalues that would lead to nontrivial solutions. This is called the eigenvalue equation. Expanding the right-hand side of the equation, we find

$$\lambda^2 - (a + d)\lambda + ad - bc = 0$$

Thus, the eigenvalues correspond to the solutions of the characteristic equation for the system. Once we find the eigenvalues, then we solve the homogeneous system for $v_1$ in terms of $v_2$, or vice versa. There are possibly an infinite number solutions to the algebraic system with representing parallel vectors in the

plane. So, we need only pick one representative eigenvector as the solution for each eigenvalue.

The method for solving problem consists of just a few simple steps .

### 1.12.3 Solving Eigenvalue Problems

a)Write the coefficient matrix;

b)Find the eigenvalues from the equations $\det(A - \lambda I) = 0$ ,and

c)Solve the linear system $(A - \lambda I)v = 0$ for each $\lambda$

**Example1.6:**

Determine the eigenvalue and eigenvector for

$$A = \begin{pmatrix} 1 & -2 \\ -3 & 2 \end{pmatrix}$$

Let v $= \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ .Then the eigenvalue problem can be writte as the system

$$(1 - \lambda)v_1 - 2v_2 = 0$$

$$-3v_1 + (2-\lambda)v_2 = 0 \qquad (1.33)$$

The eigenvalue equation must hold:

$$\det(A - \lambda I) = 0 \qquad (1.34\ )$$

$$\begin{vmatrix} 1-\lambda & -2 \\ -3 & 2-\lambda \end{vmatrix} = 0$$

Computing the determinant, we have

$$\lambda^2 - 3\lambda - 4 = 0\ .$$

$$(\lambda - 4)(\lambda + 1) = 0$$

So ,the eigenvaules are $\lambda = 4, -1$

The sconed step is to find the eigenvector .

We first insert $\lambda = 4$ into the system .Solution would be

$v_1 = 3$ , $v_2 = -2$ .For $\lambda = -1$, the system becomes

$$2v_1 - 2v_2 = 0$$

$$-3v_1 + 3v_2 = 0$$

We get the same equation $v_1 = 1$ , $v_2 = 1$ .

**1.13 Matrix Formultion of Planar Systems[17]:**

This is a first –order vector differential equation

$$\dot{x} = Ax$$

Formally ,we can write the solutions of

$$x = x_0 e^{At}.$$

### 1.13.1 Classification of the Solutions for Two Linear First Order Differential Equations

**Case 1:Two real , distinct roots.**

Solve the eigenvalue problem $A\mathrm{v} = \lambda\mathrm{v}$ for each eigenvalue obtaining two eigenvectors$v_1, v_2$ . Then write the general solution as a linear combination $x(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2$.

**Case II: One repeated root.**

Solve the eigenvalue problem for one $A\mathrm{v} = \lambda\mathrm{v}$ eigenvalue $\lambda$, obtaining the first eigenvector $\mathrm{v}_1$. One then needs a second linearly independent solution. This is obtained by solving the nonhomogeneous problem $A\mathrm{v}_2 - \lambda\mathrm{v}_2 = \mathrm{v}_1$ for $\mathrm{v}_2$ .

The general solution is then given by $x(t) = c_1 e^{\lambda t} v_1 + c_2 e^{\lambda t}(v_2 + t v_1)$ .

**Case III: Two complex conjugate roots.**

Solve the eigenvalue problem $A\mathrm{v} = \lambda\mathrm{v}$ for one eigenvalue , $\lambda = \alpha + i\beta$ , obtaining one eigenvector **v**. Note that this eigenvector may have complex entries. Thus, one can write the vector

$$y(t) = e^{\lambda t}\mathrm{v} = \mathrm{e}^{\alpha t}(cos\beta t + isine\beta t)\mathrm{v}$$

Then the general solution can be written as $c_1 y_1(t) + c_1 y_2(t)$.

### 1.13.2 Planar Systems-Summary

Here we summarize some of these cases.

| Type | Eigenvalues | Stability |
|---|---|---|
| Node | Real λ ,same signs | $\lambda > 0$ ,stable |
| Saddle | Real λ ,opposite signs | Mostly unstable |
| Center | λ Pure imaginary | $-$ |
| Focus/spiral | Complex $\lambda, Re(\lambda) \neq 0$ | $Re(\lambda > 0)$ ,stable |
| Degenerate node | Repeated roots | $\lambda > 0$ ,stable |
| Line of Equilibria | One zero eigenvalue | $\lambda > 0$ ,stable |

**Table 1.1**: List of typical behaviors in planar system

## 1.14 Non Linear Dynamics Systems[18]

We consider the nonlinear system

$$\dot{x} = F(x, y)$$

$$\dot{y} = G(x, y) \qquad (1.35)$$

Where $F$ and G are smooth functions. We assume the solution exists for all $t \geq 0$ and is unique when initial data is provided. Typically, explicit solutions to (1.35) cannot found. However, as we will see, we can construct a phase portrait for the nonlinear system without finding the solutions

**Definition 1. 36** : A critical point $(x^*, y^*)$ ,(also called an equilibrium, Fixed, or stationary point) satisfies

$F(x^*, y^*) = 0 = G(x^*, y^*)$.

**Definition 1.37**

A trajectory starting at $x_0, y_0$ is the set

$$\{(x(t), y(t)) \big| t \geq 0, \big(x(0), y(0)\big) = (x_0, y_0)\}\ .$$

Here are some facts about the solutions to the nonlinear system, (1.35):

• If the initial data is a critical point, then the solution remains at the critical point for all time.

• Trajectories cannot intersect.

• In particular, trajectories not starting at a critical point can never intersect (in finite time) a critical point.

• The interesting dynamics occur near the critical points.

It turns out that we can usually figure out the behavior of the nonlinear system near a critical point. The rest of the phase portrait can usually be deduced from this information. To determine the behavior near a critical point, we will linearize the nonlinear system around the critical point and use our knowledge of linear systems. We hope the full nonlinear system inherits the behavior of the linearized system. As we shall see, this is frequently the case.

### 1.14.1 The Linearization

To find the linear approximation to the function $f(x)$ near a point $x^*$, we of course use Taylor's theorem:

$$f(x) \approx f(x^*) + f'(x^*)(x - x^*)$$

There is a similar version of this theorem for functions depending on two variables. Indeed

$$F(x, y) \approx F(x^*, y^*) + \frac{\partial F}{\partial x}(x^*, y^*)(x - x^*) + \frac{dF}{dy}(x^*, y^*)(y - y^*).$$

At a critical point , $F(x^*, y^*) = 0 = G(x^*, y^*)$ .So the linear approximation to (1.35) is

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} \frac{\partial F}{\partial x}(x^*, y^*) & \frac{\partial F}{\partial x}(x^*, y^*) \\ \frac{\partial G}{\partial x}(x^*, y^*) & \frac{\partial G}{\partial x}(x^*, y^*) \end{pmatrix} \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix}.$$

We can clean this up a little if we set $u = x - x^*$ and $v = y - y^*$. Then the linearization near $(x^*, y^*)$becomes

$$\begin{pmatrix} u \\ v \end{pmatrix}' = \begin{pmatrix} \frac{\partial F}{\partial x}(x^*, y^*) & \frac{\partial F}{\partial x}(x^*, y^*) \\ \frac{\partial G}{\partial x}(x^*, y^*) & \frac{\partial G}{\partial x}(x^*, y^*) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

$$:= J \begin{pmatrix} u \\ v \end{pmatrix}.$$

The matrix J is called the Jacobian.

Before stating the theorems relating the behavior of the linearized system to the full nonlinear system, we need first summarize the behavior of linear systems. We observe three types of behaviors in linear systems. They are (note the origin is always a critical point in a linear system)

• all solutions converge to the origin as t $\to \infty$. This happens when the eigenvalues are negative or the have negative real part. In this case we call the origin asymptotically stable.

• solutions near the origin stay near the origin for all time. This happens when the eigenvalues are purely complex or there is an eigenvalue which is zero while the other is negative. In this case we call the origin stable..

• if neither of the above two occur, we call the origin unstable. That is, at least one trajectory leaves the vicinity of the origin.

Here are the relevant theorems.

**Theorem1.1**: (Poincar´e-Lyapunov) Suppose is a critical point$x^*, y^*$ of the nonlinear

system (1.35), and suppose the Re(λ), the real part of the eigenvalues of J (the linearization) are negative. Then the critical point is locally asymptotically stable.

**Theorem1.2 :** Suppose $x^*, y^*$is a critical point, and the real part of at least one eigenvalue of J is positive. Then the critical point is unstable.

These theorems only describe the local behavior of solutions near a fixed point. They don't say what happens to the phase portrait. For that we have

**Theorem1. 3** : (Grobman-Hartman) Suppose $x^*, y^*$ is a hyperbolic critical point (i.e. the real part of the eigenvalues of J are not zero). Then the phase portrait of the linearization and the nonlinear equations are locally homeomorphic.

This just says the phase portraits of the linearization and nonlinear equations are similar provided none of the eigenvalues of the linearization have zero real part.

**1.14.2 Robust Cases:**

• Sources or Repellers: both eigenvalues have positive real part.

• Sinks or Attractors: both eigenvalues have negative real part.

• Saddles: one eigenvalues is positive and the other negative.

**1.14.3 Marginal Cases:**

• Focus or Center: eigenvalues are pure imaginary. Linearized system does NOT describe the nonlinear system.

• Zero Eigenvalue: usually results from non isolated critical points. If the other eigenvalue is positive, the critical point is unstable. If the other eigenvalue is negative, the linearization may NOT describe the nonlinear system.

# Chapter Two

# Calculus of Variations

## 2.1 Introduction

Classical solutions to minimization problems in the calculus of variations are prescribed by boundary by value problem involving certain types differential equations , known as the associated Euler-Lagrange equations . The mathematical techniques that have been developed to handle such optimization problems are fundamental in many areas of mathematics ,physics , engineering ,and other applications .The history of the calculus of variation is tightly interwoven with the history of mathematics . The field has drawn the attention of a remarkable range of mathematical luminaries , beginning with Newton and Leibniz ,then initiated as a subject in its own right by Bernoulli brother Jakob and Johann . The first major developments appeared in the work of Euler, Lagrange, and Laplace. In the nineteenth century, Hamilton, Jacobi, Dirichlet, and Hilbert are but a few of the outstanding contributors. In modern times, the calculus of variations has continued to occupy center stage, witnessing major theoretical advances, along with wide-ranging applications in physics, engineering and all branches of mathematics. Minimization problems that can be analyzed by the calculus of variations serve to characterize the equilibrium configurations of almost all continuous physical systems, ranging through elasticity, solid and fluid mechanics, electro-magnetism, gravitation, quantum mechanics ,string theory, and many, many others. Many geometrical configurations, such as minimal surfaces, can be conveniently formulated as optimization problems.

## 2.2 Optimum of Function and Functional [80]

We give some definitions for optimum or extremum (maximum or minimum) of a function and a functional . The variation plays the same role in determining optimal value of a functional as the differential does in finding extremal or optimal of a function .

**Definition2.1 Optimum of a function**

A function $f(t)$ is said to have a relative optimum at the point $t^*$ if there is a positive parameter $\epsilon$ such that for all points in a domain $D$ that satisfy $|t - t^*| < \epsilon$ ,the increment of $f(t)$ has the same sign (positive or negative). In other words , if

$$\Delta f = f(t) - f(t^*) \geq 0 \qquad (2 .1)$$

Then , $f(t^*)$is relative local minimum. If

$$\Delta f = f(t) - f(t^*) \leq 0 \qquad (2.2)$$

Then , $f(t^*)$is relative local maximum. If the previous relations are valid for arbitrarily large $\epsilon$ , then $f(t^*)$ is said to have a global absolute optimum .

It is well known that the necessary condition for optimum of a function is that the (first) differential vanishes , i.e. $df = 0$ . The sufficient condition

1.for minimum is that is that the second differential is positive i.e. , $d^2f > 0$ ,and

2. for maximum is that the second differential is negative i.e. $d^2f < 0$ . If $d^2f = 0$ , it corresponds to stationary (or inflection) point .

**Definition2.2 Optimum of a functional**

A functional $J$ is said to have a relative optimum at $x^*$ if there is a positive $\epsilon$ such that for all function $x$ in a domain $\Omega$ which satisfy $|x - x^*| < \epsilon$ , then increment of $J$ has the same sign. In other word , if

$$\Delta J = J(x) - J(x^*) \geq 0 \qquad (2.3)$$

Then $J(x^*)$ is a relative minimum , if

$$\Delta J = J(x) - J(x^*) \leq 0 \qquad (2.4)$$

Then $J(x^*)$ is relative maximum . If the above relations are satisfied for arbitrarily large $\epsilon$ , then $J(x^*)$ is a global absolute optimum.

## 2.3 The First Variation[28]

### 2.3.1 Function of Several Variable

The definitions for local and global extrema in n dimensions are formally the same as for one- variable case. Let $\Omega \subseteq \mathbb{R}^n$ be region and suppose that $f: \Omega \to \mathbb{R}$ . For $\epsilon > 0$ and $X = (x_1, x_2, \dots, x_n)$ ,

let $B(X;\epsilon) = \left\{ \hat{X} \in \mathbb{R}^n : \left|\hat{x}_1 - x_1\right|^2 + \left|\hat{x}_2 - x_2\right|^2 + \cdots \left|\hat{x}_n - x_n\right|^2 < \epsilon^2 \right\}$.

The function $f: \Omega \to \mathbb{R}$ has a global maximum ( global minimum) on $\Omega$ at $X \in \Omega$ If $f(\hat{x}) \leq f(x)$ ( $f(\hat{x}) \geq f(x)$ ) for $\hat{X} \in \Omega$ . The function $f$ has a local maximum ( local minimum) at $\Omega$ if there exists a number $\epsilon > 0$ such that for any $\hat{X} \in B(X;\epsilon) \subset \Omega$, $f(\hat{x}) \leq f(x)$ ( $f(\hat{x}) \geq f(x)$ ).

Necessary conditions for a smooth function of two independent variables to have local extrema can be derived from considerations similar to those used in the single-variable case . Suppose that $f: \Omega \to \mathbb{R}$ is smooth function on the region $\Omega \subseteq \mathbb{R}^2$ and that $f$ has local extremum at $x = (x_1 , x_2) \in \Omega$ .Then There an $\epsilon > 0$ Such that , $f(\hat{x}) - f(x)$ does not change sign for all $\hat{X} \in B(X;\epsilon)$ .

Let , $\hat{x} = x + \epsilon\eta$ where $\eta = (\eta_1, \eta_2) \in \mathbb{R}^2$ . For $\epsilon$ small , Taylor's Theorem implies

$$f(\hat{x}) = f(x) + \epsilon\left\{\eta_1 \frac{\partial f(x)}{\partial x_1} + \eta_2 \frac{\partial f(x)}{\partial x_2}\right\}$$

$$+ \frac{\epsilon^2}{2!}\left\{\eta_1^2 \frac{\partial^2 f(x)}{\partial x_1^2} + 2\eta_1\eta_2 \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} + \eta_2^2 \frac{\partial^2 f(x)}{\partial x_2^2}\right\} + O(\epsilon^3),$$

And the sign of , $f(\hat{x}) - f(x)$ is given by the linear term in the Taylor expansion unless this terms is zero. If x is a local extremum we must therefore have that

$$(\eta_1, \eta_2) \cdot \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_1}\right) = 0 \qquad (2.5)$$

For all $\eta \in \mathbb{R}^2$ . If $f$ has a local extremum at $x$ that

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

Geometrically equation (2.5) implies that the tangent plane to graph of $f$ is horizontal at a local extremum point $x$ at which $\nabla f(x) = 0$ yare called stationary points.

**2.4 The Euler Lagrange Equation[28]**

Let $J: X \to \mathbb{R}$ be a functional defined on the function space $(X, \|\cdot\|)$ and let $S \subseteq X$ . The functional $J$ is said have local maximum in $S$ at $y \in S$ it there exists an $\epsilon > 0$ such that $J(\hat{y}) - J(y) \leq 0$ for all $\hat{y} \in S$ such that $\|\hat{y} - y\| < \epsilon$ . The functional $J$ is said to have a local minimum in $S$ is at $y \in S$ y is local maximum in for $-J$ .The set $S$ is a set of functions satisfying certain boundary conditions .

Function $\hat{y} \in S$ in an $\epsilon$-neighbourhood of a function $y \in S$ can be represented in convenient way as a perturbation of $y$ . If $\hat{y} \in S$ and $\|\hat{y} - y\| < \epsilon$ , then there is some $\eta \in X$ such that

$$\hat{y} = y + \epsilon\eta$$

All functions in an $\epsilon$-neighbourhood of functions of $y$ can generated from a suitable set $H_\epsilon$ of functions $\eta$ . The set $H_\epsilon$ is thus defined

$$H_\epsilon = \{\eta \in X : y + \epsilon\eta \in S \text{ and } \|\eta\| < 1\} .$$

The auxiliary set $H_\epsilon$ can thus be replaced by the set

$$H = \{\eta \in X : y + \epsilon\eta \in S\} .$$

For purposes of analysis.

At this stage we specialize to a particular class of problem called the fixed endpoint variational problem , and work with the vector space $C^2[x_0 , x_1]$ . That consists of functions on $[x_0 , x_1]$ that have continuous second derivatives .

Let $J = C^2[x_0 , x_1] \to \mathbb{R}$ be functional of the form

$$J(y) = \int_{x_0}^{x_1} f(x , y, y')dx$$

Where $f$ if function assumed to have at least second-order continuous partial derivative with respect to $x , y \, and \, y'$ .Given two values $y_0 , y_1 \in \mathbb{R}$ , the fixed endpoint variational problem consists $y \in C^2[x_0 , x_1]$ such that

$y(x_0) = y_0$ , $y(x_1) = y_1$ and $J$ has a local extremum in $S$ at $y \in S$ . Here

$$S = \{y \in C^2[x_0 , x_1]: y(x_0) = y_0 \text{ } and \text{ } y(x_1) = y_1\} \text{ and}$$

$$H = \{\eta \in C^2[x_0 , x_1]: \eta(x_0) = \eta(x_1) = 0\}$$

For any $\hat{y} \in S$ there $\eta \subset H$ such that $\hat{y} = y + \epsilon\eta$

$$f(x, \hat{y}, \hat{y}') = f(x, y + \in \eta, y' + \in \eta')$$

$$= f(x, y, y') + \epsilon \left\{\eta \frac{\partial f}{\partial y} + \eta' \frac{\partial f}{\partial y'}\right\} + O(\epsilon^2)$$

$$J(\hat{y}) - J(y) = \int_{x_0}^{x_1} f(x, \hat{y}, \hat{y}')dx - \int_{x_0}^{x_1} f(x , y , y')dx$$

$$= \left\{\left(f(x , y , y') + \epsilon\left\{\eta \frac{\partial f}{\partial y} + \eta' \frac{\partial f}{\partial y'}\right\} + O(\epsilon^2)\right) - f(x, y , y')dx\right\}$$

$$= \epsilon \int_{x_0}^{x_1} \left(\eta \frac{\partial f}{\partial y} + \eta' \frac{\partial f}{\partial y'}\right) dx + O(\epsilon^2)$$

$$= \epsilon\delta J(\eta, y) + O(\epsilon^2)$$

The quantity

$$\delta J(\eta, y) = \int_{x_0}^{x_1} \left( \eta \frac{\partial f}{\partial y} + \eta' \frac{\partial f}{\partial y'} \right) dx$$

Is called the first variation of $J$ .

Unless $\delta J(\eta, y) = 0$ for all $\eta \in H$ . If $J(y)$ is a local maximum

Then

$$\delta J(\eta, y) = \int_{x_0}^{x_1} \left( \eta \frac{\partial f}{\partial y} + \eta' \frac{\partial f}{\partial y'} \right) dx = 0 \tag{2.6}$$

For all $\eta \in H$

$$\int_{x_0}^{x_1} \left( \eta' \frac{\partial f}{\partial y'} dx \right) = \eta \frac{\partial f}{\partial y'} \bigg|_{x_0}^{x_1} - \int_{x_0}^{x_1} \eta \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) dx$$

$$= - \int_{x_0}^{x_1} \eta \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) dx$$

Where we have used conditions $\eta(x_0) = 0$ and $\eta(x_1) = 0$ . Equation (2.6) can thus written

$$\int_{x_0}^{x_1} \eta \left\{ \frac{\partial f}{\partial y} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) \right\} dx = 0 \tag{2.7}$$

Now

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) = \frac{\partial f}{\partial y} - \frac{\partial^2 f}{\partial x \partial y'} - \frac{\partial^2 f}{\partial y \partial y'} y' - \frac{\partial^2 f}{\partial y' \partial y'} y'',$$

And given that $f$ has at least two continuous derivatives for any fixed $y \in C^2[x_0 , x_1]$ the function $E: [x_0 , x_1] \to \mathbb{R}$ defined by

$$E(x) = \frac{\partial f}{\partial y} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right)$$

Is continuous the interval $[x_0 , x_1]$ .

**Theorem 2.1**

Let $J = C^2[x_0 , x_1] \to \mathbb{R}$ be functional of the form

$$J(y) = \int_{x_0}^{x_1} f(x , y, y')dx$$

Where $f$ has continuous partial derivative of second-order with respect to $x , y \; and \; y'$ and $x_0 < x_1$ . Let

$$S = \{y \in C^2[x_0 , x_1]: y(x_0) = y_0 \;\; and \; y(x_1) = y_1 \} ,$$

Where $y_0$ and $y_1$ are given real number . If $y \in S$ is an extremal for $J$ , then

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0 \qquad (2.8)$$

**Note**

$$\frac{d}{dx}\left(f_{y'}\right) = f_{xy'} + f_{yy'} \cdot y' + f_{y'y'} \cdot y''$$

$$\frac{\partial f}{\partial y} = f_y \qquad , \quad \frac{\partial f}{\partial y'} = f_{y'}$$

$$\left(f_{y'y'}\right) \cdot y'' + \left(f_{yy'}\right) \cdot y' + \left(f_{xy'}\right) - f_y = 0 \qquad (2.9)$$

For all $x \in [x_0 , x_1]$. Equation (2.8) is a second-order (generally nonlinear) ordinary differential equation that any (smooth) extremal $y$ must satisfy .

This differential equation (2.8) is called the Euler –Lagrange equation. The boundary values associated with this equation for the fixed endpoint problem are

$$y(x_0) = y_0 \qquad , y(x_1) = y_1 \qquad (2.10)$$

The Euler- Lagrange equation is the infinite –dimensional analogue of the equation $\nabla f(x) = 0$ . In the transition from finite to infinite dimensional , an algebraic condition for determined of points $X \in \mathbb{R}^n$ which might lead to local extrema is replaced by a boundary –value problem involving a second- order differential equation.

## 2.5 Some Special Cases[28]

**Case I : No Explicit $y$ dependence**

Suppose That the functional is of the form

$$J(y) = \int_{x_0}^{x_1} f(x, y')dx$$

Where the variable $y$ does not appear explicitly in the integrand . Evidently the Euler- Lagrange equation reduces to

$$\frac{\partial f}{\partial y'} = c_1 \qquad (2.11)$$

In principle , equation (2.11) is solvable for $y'$ ,provided $\partial^2 f / \partial y'^2 \neq 0$ so that equation (2.11) could be recast in the form

$$y' = g(x, c_1) ,$$

For some function $g$ and the integrand , in practice , however , solving equation (2.11) for $y'$ can prove formidable if not impossible , and there may be several solutions available.

**Case II :No Explicit $x$ dependence**

Another simplification is available when the integrand does not contain the independent variable $x$ explicitly .

**Theorem 2.2**

Let $J$ be a functional of the form

$$J(y) = \int_{x_0}^{x_1} f(y, y')dx$$

And define the function $H$ by

$$H(y, y') = y' \frac{\partial f}{\partial y'} - f$$

Then $H$ is constant along any extremal $y$

**Proof**

Suppose that $y$ is an extremal for $J$ . Now ‘

$$\frac{d}{dx}H(y,y') = \frac{d}{dx}\left(y'\frac{\partial f}{\partial y'} - f\right)$$

$$= y''\frac{\partial f}{\partial y'} - y'\frac{d}{dx}\frac{\partial f}{\partial y'} - \left(y'\frac{\partial f}{\partial y} + y''\frac{\partial f}{\partial y'}\right)$$

$$= y'\left(\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y}\right),$$

And since $y$ is an extremal ,the Euler – Lagrange equation (2.8) is satisfied ; Hence ,

$$\frac{d}{dx}H(y,y') = 0$$

Consequently , $H$ must be constant alonge an extrmal. Note that the function $H$ depends only on $y$ and $y'$ , and Thus the equation

$$\frac{d}{dx}H(y,y') = const$$

Is a first –order differential equation for the extremely .

**2.6 The Second Variation :**

A smooth function $y = [x_0, x_1] \to \mathbb{R}$ such that $y(x_0) = y_0$ , $y(x_1) = y_1$ , and the functional

$$J(y) = \int_{x_0}^{x_1} f(x\,, y, y')dx \qquad (2.12)$$

Is an extremum . We assume that for any given extremal $y\,, f(x\,, y(x), y'(x))$ is smooth in a neighbourhood of $x_0$ and in a neighbourhood of $x_1$.

$$f(y,\hat{y},\hat{y}') = f(x\,,y\,,y')\epsilon\left(\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y\prime}\right) + \frac{\epsilon^2}{2}\left(\left\{\eta^2\frac{\partial^2 f}{\partial y} + 2\eta\eta'\frac{\partial^2 f}{\partial y\partial y'} + \eta'^2\frac{\partial^2 f}{\partial y'^2}\right\}\right) + O(\epsilon^3)\,,$$

We use the following notation

$$f_{yy} = \frac{\partial^2 f}{\partial y^2} \quad , f_{yy'} = \frac{\partial^2 f}{\partial y \partial y'} \quad , f_{y'y'} = \frac{\partial^2 f}{\partial y'^2}$$

Where , unless otherwise noted , the partial derivative are evaluated $(x\,, y\,, y')$

Thus ,

$$J(\hat{y}) - J(y) = \epsilon \delta J(\eta, y) + \frac{\epsilon^2}{2} \delta^2 J(\eta, y) + O(\epsilon^3) ,$$

Where $\delta J(\eta, y)$ is the first variation and

$$\delta^2 J(\eta, y) = \int_{x_0}^{x_1} \left( \eta^2 \frac{\partial^2 f}{\partial y} + 2\eta\eta' \frac{\partial^2 f}{\partial y \partial y'} + \eta'^2 \frac{\partial^2 f}{\partial y'^2} \right) dx .$$

The term $\delta^2 J(\eta, y)$ is called the second variation of $J$ . Let H denoted the set of function $\eta$ smooth on $[x_0, x_1]$ such that $\eta(x_0) = \eta(x_1) = 0$.

**Theorem 2.3**

Suppose that $J$ has a locale extremum in $S$ at $y$ . If $y$ is a local minimum , Then

$$\delta^2 J(\eta, y) \geq 0$$

For $\eta \in H$ , If $y$ is a local maximum then

$$\delta^2 J(\eta, y) \leq 0$$

For all $\eta \in H$.

## 2.7 Ordinary Differential Equations [67]

**Set**

$$J(y) = \int_{x_0}^{x_1} f(x, y(x), y'(x)) dx$$

And for given $y_0, y_1 \in \mathbb{R}$

$$V = \{y \in C^1[x_0, x_1]: y(x_0) = y_0 , y(x_1) = y_1\} ,$$

Where $-\infty < x_0 < x_1 < \infty$ and $f$ is sufficiently regular .One of the basic problems in the calculus of variation is

(P) $\min_{y \in V} J(y)$

### 2.7.1 Euler Equation

Let $y \in V$ be a solution of (P) and assume additionally $y \in C^2(x_0, x_1)$ , then

$$\frac{d}{dx} f_{\ y'}\big(x, y(x), y'(x)\big) = f_y(x, y(x), y'(x))$$

In $(x_0,x_1)$

**Basic lemma in the calculus of variations 2.1**

Let $h \in C(a,b)$ and $\int_a^b h(x)\, \emptyset(x)dx = 0$

For all $\emptyset(x) \in C_0^1(a,b)$.Then $h(x) = 0$ on (a,b).

**Proof**

Assume $h(x_0) > o$ for an $x_0 \in (a,b)$, then there is a $\delta > 0$ such that

$(x_0 - \delta, x_0 + \delta) \subset (a,b)$and $h(x) \geq h(x_0)/2$ on $(x_0 - \delta, x_0 + \delta)$ .Set

$$\emptyset(x) = \begin{cases} (\delta^2 - |x - x_0|^2)^2 \ if\ x \in (x_0 - \delta, x_0 + \delta) \\ 0 \qquad\qquad ifx\ \in\ (a,b)\backslash[x_0 - \delta, x_0 + \delta) \end{cases}$$

Thus $\emptyset \in C_0^1(a,b)$ and

$$\int_a^b h(x)\emptyset(x)dx \geq \frac{h(x_0)}{2} \int_{x_{0-\delta}}^{x_0+\delta} \emptyset(x)dx > 0$$

Which a contraction to the assumption of the lemma.

**2.7.2 Brachistochrone:**

In 1696 Johan Bernoulli studied the problem of a brachistochrone to find a curve connecting two points $P_1$ and $P_2$ such that a mass point moves from $P_1$to $P_2$ asfost as possible in a downward directed constant gravitional field see Figure 2.2 .The associated variation problem is here

$$\min_{(x,y)\in V} \int_{t_1}^{t_2} \frac{\sqrt{x'(t)^2 + y'(t)^2}}{\sqrt{y(t) - y_1 + k}} \, dt$$

Where V is the set of $C^1[t_1, t_2]$ curves defined by (x(t),y(t)) , $t_1 \leq t \leq t_2$ with $x'(t)^2 + y(t)^2 \neq 0$ , $(x(t_1), y(t_1) = P_1$ ,$x(t_2), y(t_2)) = P_2$ *and*

$k := v_1^2/2g$ *where* $v_1$ the absolute value of the initial velocity of the mass point and $y_1 = y(t_1)$. Solutions are cycloids

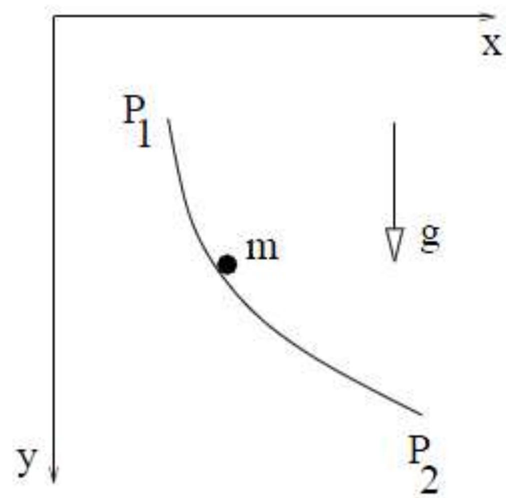



**Figure 2.1** problem of a brachistochronce

One arrives at the above functional which we have minimize since

$$v = \sqrt{2g(y - y_1) + v_1^2} \text{ , } v = ds/dt \text{ , } ds = \sqrt{x_1(t)^2 + y'(t)^2}dt$$

And $$T = \int_{t_1}^{t_2} dt = \int_{t_1}^{t_2} \frac{ds}{v}$$

Where $T$ is the time which the mass point needs to move form $P_1$ and $P_2$

A piecewise continuous solution $u$ is called optimal control and the solution $x$ of the associated system of boundary value problems is said to be optimal trajectory.

## 2.8 Partial Differential Equations[67]

The same procedure as above to the following multiple integral leads to a second-order quasilinear partial differential equation . Set

$$J(y) = \int_{\Omega}(x, y, \nabla y)dx$$

Where $\Omega \subset \mathbb{R}^n$ is a domain ,

$x = (x_1, \dots, x_n),\ y = y(x) : \Omega \to \mathbb{R} \qquad and \quad \nabla y = (y_{x_1}, \dots, y_{x_n})$

It is assumed that the function F is sufficiently regular in its arguments .

For a given function h , defined on $\partial\Omega$

$$y = \{y \in C^1(\overline{\Omega}) : y = h \ on \ \partial\Omega\}$$

### 2.8.1 Euler Equation

Of (P) and additionally $y \in C^2(\Omega)$

Then

$$\sum_{i=1}^{n} \frac{\partial}{\partial x_i} F_{y_{x_i}} = F_y$$

In $\Omega$.

## 2.9 Functions of $n$ Variables [67]

Let $f$ be a real- function defined on a nonempty subset $X \subseteq R^n$ . In the conditions below where derivative occur , we assume that $f \in C^1$ or $f \in C^2$ on open set $X \subseteq R^n$.

### 2.9.1 Optima, Tangent Cones

Let $f$ be a real –valued functional defined on a nonempty subset $V \subseteq X$.

**Definition 2.3**

A subset $V \subseteq X$ is said to be convex if for any tow vectors $x, y \in V$ That inclusion $\lambda x + (1-\lambda)y \in V$ holds for all $0 \leq \lambda \leq 1$ .

**Definition 2.4**

We say that a functional $f$ defined on a convex subset $V \subseteq X$ is convex if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

For all $x, y \in V$ and for all $0 \leq \lambda \leq 1$ , and $f$ is strictly convex if the strict inequality holds for all $x, y \in V$ , $x \neq y$ , and for all $\lambda, 0 < \lambda < 1$ .

**Theorem2.4**

If $f$ is q convex functional on a convex set $V \subseteq X$ , then any local minimum of $f$ in $V$ is a global minimum of $f$ in $V$ .

**Proof**

Suppose that $x$ is no global minimum , then there exists an $x^1 \in V$ such that $f(x_1) < f(x)$ . Set

$$y(\lambda) = \lambda x^1 + (1-\lambda)x \quad , \; 0 < \lambda < 1$$

Then

$$f(y(\lambda)) \leq \lambda f(x^1) + (1-\lambda)f(x) < \lambda f(x) + (1-\lambda)f(x) = f(x).$$

For each given $\rho > 0$ there exists a $\lambda = \lambda(\rho)$ such that $y(\lambda) \in B_\rho(x)$ and $f(y(\lambda)) < f(x)$ . This is a contradiction to the assumption.

**Theorem 2.5**

If $f$ is a strictly convex functional on a convex set $V \subseteq X$ , then a minimum (local or global) is unique.

**Proof**

Suppose that $x^1$ , $x^2 \in V$ define minima of $f$ , then $f(x^1) = f(x^2)$ ,see theorem 2.4 . Assume $x^1 \neq x^2$ , then for $0 < \lambda < 1$

$$f(\lambda x^1 + (1-\lambda)x^2) < \lambda f(\, x^1) + (1-\lambda)f(x^2) = f(x_1) + f(x_2)$$

This is a contradiction to the assumption that $x^1$ , $x^2$ define global minima

**Theorem 2.6**

- **(a)** If $f$is a convex function and $V \subset X$ a convex set , then the set of minimizers is convex.
- **(b)** If $f$ is concave, $V \subset X$ convex , then the set of maximizers is convex .

**Theorem 2.7**

Suppose the $V \subset X$ is convex .Then $f$ is convex of $V$ if and only if

$$f(y) - f(x) \geq \langle\, f'(x), y - x\rangle \text{ for all } x, y \in V$$

**Remark 2.1**

The inequality of the theorem says that the surface S defined by $z = f(y)$ is above of the tangent plane $T_x$ defined by $z = \langle f'(x). y - x\rangle + f(x)$ ,see Figure 2.1 for the case n = 1.

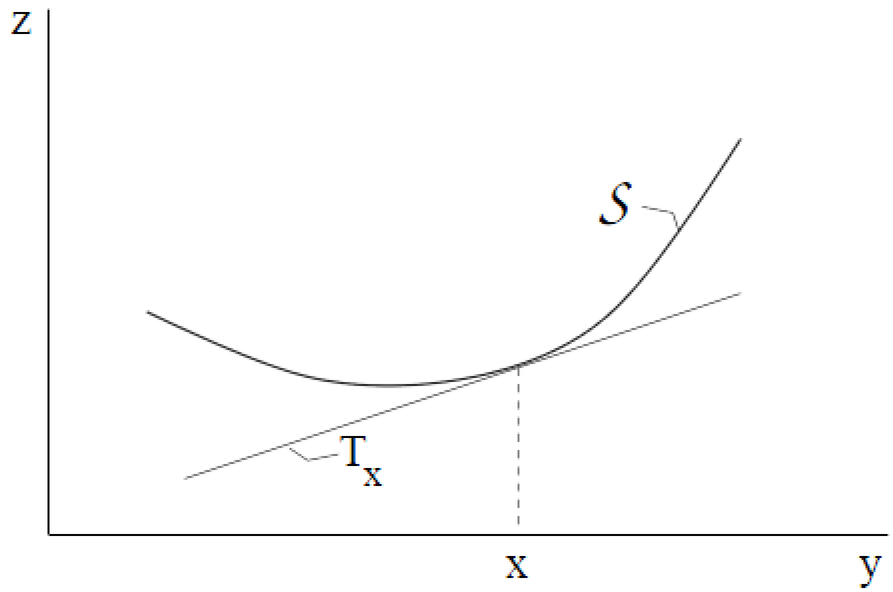



**Figure2.2**: figure the theorem 2.7

**Definition 2.5**

A nonempty subset $C \subseteq \mathbb{R}^n$is said to be a cone with vertex at $z \in \mathbb{R}^n$, if $f \in C$ implies that $z + t(y - z) \in C$ for each $t > 0$.

Let $V$ be a nonempty subset of $X$ .

**Definition2.6**

For given $x \in V$ we define the local tangent cone of $V$ at $x$ by $T(V, x) = \{w \in \mathbb{R}^n$: there exist sequences $x^k \in V$ , $t_k \in \mathbb{R}, t_k > 0$, such that $x^k \to x$ and $t_k(x^k - x) \to w$ as $k \to \infty\}$.

The definition implies immediately

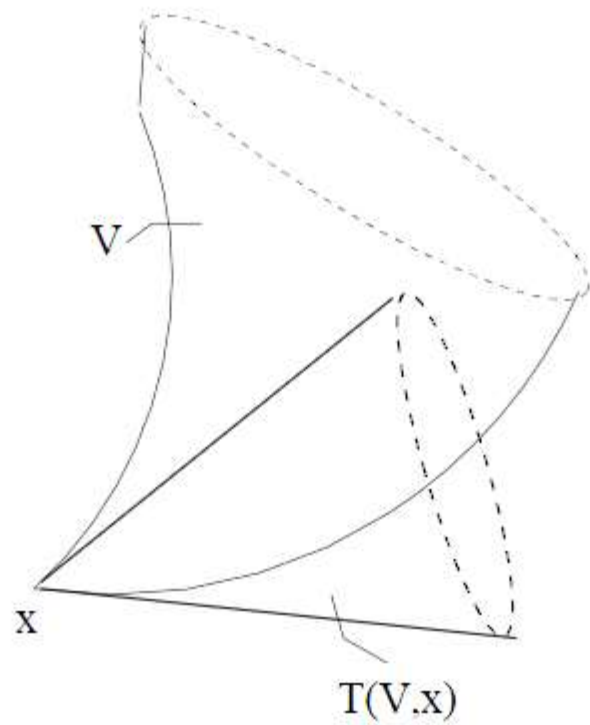



**Figure 2.3** : Tangent Cone

**Corollaries 2.1**

(i) the set $T(V,x)$ is a cone with vertex at zero

(ii)A vector $x \in V$ is not isolated if and only if $T(V,x) \neq \{0\}$.

(iii)Suppose that $w \neq 0$ then $t_k \to \infty$ .

(iv)$T(V,x)$ is closed

(v)$T(V,x)$ is convex if $V$ is convex.

**2.10 Isoperimetric Problems[27]**

Here we consider the problem of finding an extremum to a functional subject to an equality constraint involving a second functional . Historically , the first problems of this type involved finding an optimal curve whose total length (perimeter ) was fixed –hence the name. It turns out that standared method used in the optimization of functions in $\mathbb{R}^n$ under equality constraints- Lagrange multipliers- also applies here:

Problem IP: Minimise the functional

$$J(y) = \int_{x_0}^{x_1} f(x, y, y')dx$$

With $y(x_0) = y_0$ , $y(x_1) = y_1$ , subject to the integral constraint:

$$I(y) = \int_{x_0}^{x_1} g(x, y, y')dx = c$$

Where $c$ is constant

**Theorem 2.8**

In order that $y = y^*(t)$ is a solution of problem IP (isoperimetric problem) it is necessary that it should be an extremal of

$$\int_{x_0}^{x_1} f(x, y, y') + \lambda g(x, y, y'))dx$$

For a certain $\lambda$ (Lagrange multiplier ).

## 2.11 Hamilton's Conical Equation[27]

Define the momentum $P$ and Hamiltonian $H$ as

$$p = p(x, y, y') = L_{y'}(x, y, y')$$

And

$$H = H(x, y, y', p) \coloneqq py' - L(x, y, y')$$

The variable $y$ and $p$ are called the canonical variable . Let $y$ be extremal , that is $y$ satisfies E-L equation .Then, it follows that $y$ and $p$ satisfies

$$\left.\begin{aligned} \frac{dy}{dx} &= H_p \\ \frac{dp}{dx} &= -H_y \end{aligned}\right\}$$

The above system is known as Hamilton's canonical system of equations.

# Chapter Three

# Mathematical Optimal Control

## 3.1 Introduction

Control system design is generally a trial and error process in classical which various methods of analysis are used iteratively to determine the design parameters of an "acceptable" system . Acceptable performance is generally defined in terms of time and frequency domain criteria such as rise time, settling time, peak overshoot, gain and phase margin, and bandwidth. Radically different performance criteria must be satisfied, however, by the complex, multiple-input, multiple-output systems required to meet the demands of modern technology. For example, the design of a spacecraft attitude control system that minimizes fuel expenditure is not amenable to solution by classical methods. A new and direct approach to the synthesis of these complex systems, called optimal control theory, has been made feasible by the development of the digital computer. The objective of optimal control theory is to determine the control signals that will cause a process to satisfy the physical constraints and at the same time minimize (or maximize) some performance criterion. Later, we shall give a more explicit mathematical statement of “ the optimal control problem ” ,but first let us consider the matter of problem formulation.

## 3.2 The Basic problem[22]

### 3.2.1 Controlled Dynamics

We open our discussion by considering on ordinary differential equation having the from

$$\begin{cases} \dot{x}(t) = f\big(x(t)\big) \\ x(0) = x^0 \end{cases} \qquad (t > 0) \qquad (3.1)$$

We are here given the initial point $x^0 \in \mathbb{R}^n$ and the function $f: \mathbb{R}^n \to \mathbb{R}^n$. The unknown is the curve $x: [0, \infty) \to \mathbb{R}^n$ , which we interpret as the dynamical evolution of the state of some "system". We generalize a bit and suppose now that $f$ depends also upon some "control" parameters belonging to a set $A \subset \mathbb{R}^m$ so that $f: \mathbb{R}^n \times A \to \mathbb{R}^n$ Then if we select some value $a \in A$ and consider the corresponding dynamics :

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

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

suppose we define the function $u: [0, \infty) \to A$ this way:

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

For time $0 < t_1 < t_2 < t_3$ ... and parameter value $a_1$ , $a_2$ , $a_3$ ... $\in A$ ; and we then solve the dynamical equation

$$\begin{cases} \dot{x}(t) = f(x(t), u(t)) \\ x(0) = x^0 \end{cases} \qquad (t > 0) \qquad (3.2)$$

The point is that the system may behave quite differently as we change the control parameters .More general function $u: [0, \infty) \to A$ a control . Corresponding to each control , we consider the ODE (3.2) and regard the trajectory $x(\cdot)$ as the corresponding response of the system .

### 3.2.2 Payoffs

Let us define the payoff functional

$$P[u(\cdot)] := \int_0^T r\big(x(t), u(t)\big)dt + g(x(t)) \,, \qquad (3.3)$$

Where $x(\cdot)$ solves (ODE) for the control $u(\cdot)$ . 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.

## 3.3 Controllability[22]

### 3.3.1 Controllability Question

**Definition 3.1**

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 $x(t) = 0$ 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 that our ODE is linear in both the sate $x(\cdot)$ and the control $u(\cdot)$ , and consequently has the form

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ x(0) = x^0 \end{cases} \qquad (t > 0) \qquad \text{(ODE)}$$

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

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

**Definition3.2**

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

$$\begin{cases} \dot{x}(t) = Ax(t) \\ x(0) = \mathrm{I} \end{cases} \qquad (t \in \mathbb{R})$$

We call $x(\cdot)$ a fundamental solution , and sometime write

$$x(t) = e^{tA} := \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}$$

Observe that $x^{-1}(t) = x(-1)$

**Theorem 3.1 (Solving Linear Systems of ODE)**

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

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

Is

$$x(t) = X(t)x^0 = e^{tA}x^0$$

(ii) The unique solution of the nonhomogeneous system

$$\begin{cases} \dot{x}(t) = Ax(t) + f(t) \\ x(0) = x^0 \end{cases}$$

Is

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

## 3.3.2 Controllability of Linear Equations[22]

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

$$x(t) = X(t)x^0 + X(t) \int_0^t X^{-1}(s) Bu(s) ds$$

Where $X(t) = e^{tM}$ Furthermore , observe that

$$x^0 \in C(t)$$

If and only if

There exists a control $u(\cdot) \in A$ such that $x(t) = 0$ ( 3.4)

If and only if

$$0 = X(t)x^0 + X(t) \int_0^t X^{-1}(s) Bu(s) ds \qquad (3.5)$$

For some control $u(\cdot) \in A$

If and only if

$$x^0 = -\int_0^t X^{-1}(s)Bu(s)ds \qquad (3.6)$$

For some control $u(\cdot) \in A$.

**A simple Example 3.1**

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

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

This is a system of the form $\dot{x} = Ax + Bu$ , for

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

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

**Definition 3.3**

The controllability matrix is

$$G = G(A , B) := [B , AB , A^2B , \dots , A^{n-1}B]$$

**Theorem 3.2**

$$\text{rank } G = n \Leftrightarrow \quad 0 \in C^0$$

**Proof**

1. Suppose that . Note that

$$\text{rank } G \le n$$

If rank $G < n$, then there exists a vector $b \in \mathbb{R}^n , b \neq 0$ such that

$$b^T G = 0$$

This yields

$$b^T B = b^T AB = \cdots = b^T A^{n-1} B = 0$$

By Cayley- Hamilton’s theorem , we also have

$$b^T A^K B = 0 \quad , k \geq 0 \quad , b^T X^{-1}(t) B = 0$$

We now claim

$b$ is perpendicular to $C(t)$ , i.e. $C^0 = \emptyset$

If $x^0 \in C(t)$, then

$$x^0 = -\int_0^t X^{-1}(s)Bu(s)ds \ , \qquad u \in A$$

Therefore

$$b^T x^0 == -\int_0^t b^T X^{-1}(s)Bu(s)ds = 0$$

2. Suppose that $0 \notin C^0$. Thus $0 \notin C^0(t)$ ,$\forall t \geq 0$ .Since $C(t)$ convex , there exists $b \neq 0$ such that

$$b^T x^0 \leq 0 \ , \ \ x^0 \in C(t)$$

For $x^0 \ \in C(t)$

$$x^0 = -\int_0^t X^{-1}(s)Bu(s)ds$$

Thus

$$b^T x^0 = -\int_0^t b^T X^{-1}(s)Bu(s)ds \leq 0$$

This yield

$$b^T X^{-1}(s)B = 0$$

By differentiating the above relation, we have

$$b^T B = b^T AB = \cdots = b^T A^{n-1} B = 0 \text{ ,i.e., } b^T G = 0$$

Hence

$$\text{rank } G < n \ .$$

**Definition 3.4**

We say the liner system (ODE) is controllable if

$$C = \mathbb{R}^n$$

**Theorem 3.3**

Let λ be the eigenvalue of the matrix A

$$\text{rank}G = n \ \text{ and } \ \text{Re}\,(\lambda) \leq 0$$

then the system (ODE) is controllable .

**Example 3.2**

For which $n = 2\ , m = 1,\ \ A = [-1\ ,1]$ and

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u$$

Then

$$G = [B, AB] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Therefore

$$\text{rank}G = 2 = n$$

Also ,the characteristic polynomial of the matrix M is

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

**3.4 Uncontrollable System[16]**

An uncontrollable system has a subsystem that is physically disconnected from the, Input

**3.5 Observability[16]**

We discuss the observability of linear systems. Consider the unforced System described by the following equations

$$\dot{x} = Ax \tag{3.7}$$

$$Y = CX \tag{3.8}$$

Where **x** = state vector (n-vector)

$Y = output\ control$ (m-vector)

A=n×n matrix

C =m× $n$ matrix

The system is said to be completely observable if every state **x** $(t_0)$can be determined from the observation of **y**(t) over a finite time interval,$t_0 \leq t \leq t_1$ The system is therefore completely observable if every transition of the state eventually affects every element, of the output vector. The concept of observability is useful in solving the problem of reconstructingun measurable state variables from measurable variables in the minimum possible length of time we treat only linear, time-invariant systems. Therefore, without loss of generality, we can assume that $t_0$=0 the concept of observability is very important because, in practice, the difficulty. Encountered with state feedback control is that some of the state variables are not. Accessible for direct measurement, with the result that it becomes necessary to estimate the un measurable state variables in order to construct the control signals. It will be. Shown in that such estimate of state variables are possible if and only if the system is completely observable. In discussing observability conditions, we consider the unforced system as given by Equations (3.7) and (3.8).The reason for this is as follows: If the system is described

$$\dot{x} = Ax + Bu$$

$$Y = CX + DU$$

Then

$$X(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)} Bu(\tau)d\tau$$

And y (t) is

$$Y(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)} Bu(\tau)d\tau + DU$$

Since the matrices A, B, C, and Dare known and u (t) is also known, the last two terms, on the right-hand side of this last equation are known quantities. Therefore they may be subtracted from the observed value of **y** (t). Hence, for investigating a necessary and sufficient condition for complete observability, it suffices to consider the system described by Equations (3.7) and (3.8)

### 3.5.1Complete Observability of Continuous Time Systems[16]

Consider the system described by Equations (3.7) and (3.8).The output vectorY (t) is

$$Y(t) = Ce^{At}X(0).$$

We have

$$e^{At} = \sum_{k=0}^{n-1} \alpha_k(\tau) A^k$$

Where n is the degree of the characteristic polynomial. Hence we obtain.

$$y(t) = \sum_{k=0}^{n-1} \alpha_k(\tau)\, CA^k X(0)$$

$$y(t = \alpha_0(t)CX(0) + \alpha_1(t)CAX(0) + \cdots + \alpha_{n-1}(t)CA^{n-1}X(0 \qquad (3.9)$$

If the system is completely observable, then, given the output **y** (t) over a time interval $t_0 \leq t \leq t_1$ **,X** (0) is uniquely determined from Equation (3.18). It can be shown that this requires the rank of the nm×n matrix

$$\begin{bmatrix} C \\ CA \\ \vdots \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

From this analysis, we can state the condition for complete observability as follows the system described by Equations (3.7) and (3.8) is completely observable if and only if the n×nm matrix

$$[C^T \vdots A^T C^T \vdots \cdots\cdots \vdots (A^T)^{n-1} C^T]$$

Is of rank n or has n linearly independent column vectors. This matrix is called the observability matrix

**Example 3.3** considers the system .Is this system controllable and observable. Since the rank of the matrix

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$Y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$[B \vdots AB] = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$$

Is 2, the system is completely state controllable. For output controllability, let us find the rank of the matrix $[CB \vdots CAB]$ since

$$[CB \vdots CAB] = [1 \quad 0]$$

The rank of this matrix is 1. Hence, the system is completely output controllable. To test the observability condition, examine the rank of

$$[C^T \vdots A^T C^T]$$

Since

$$[C^T \vdots A^T C^T] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The rank of $[C^T \vdots A^T C^T]$ is 2. Hence, the system is completely observable

## 3.6 Bang-Bang principle[22]

We will again take $A$ to be the cube $[-1, 1]^m$ in $\mathbb{R}^m$

**Definition 3.5**

A control $u(\cdot) \in A$is called Bang-Bang if for each time $t \geq 0$ and each index $i = 1, \ldots, m$ ,we have $|u^i(t)| = 1$ where

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

**Theorem 3.4 (Bang-Bang Principle)**

Let $t > 0$ and suppose $x^0 \in C(t)$ , for The system

$$\dot{x}(t) = Ax(t) + Bu(t)$$

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

**Proof**

Let $x^0 \in C(t)$. Then , we set

$$K = \{u(\cdot) \in A : u(\cdot) \; steers \; x^0 \; to \; 0 \; at \; time \; t\}$$

## 3.7 Detectability[16]

For a partially observable system, if the unobservable modes are, stable and the observable modes are unstable, the system is said to be detectable. Note, that the concept of detectability is dual to the concept of stabilizability.

## 3.8 Unconstrained Optimization[27]

Let $\Omega$ be an open set in $\mathbb{R}^n$ , $f$ be $C^1$ function and assume $x^*$ is called a local minimum . Note that $x^*$ is called a local minimum of $f$ if $\exists$ a neighborhood of $U$ of $x_0$ in $\Omega$ such that $f(x^*) \leq f(x)$ for all $x \in U$. One can derive a first order necessary condition as $\nabla f(x^*) = 0$ and assuming that $f \in C^2$ , a second order necessary condition can also be derived as $\nabla^2 f(x^*) \geq 0$ that is $\langle \nabla^2 f(x^*) h , h \rangle \geq 0$ for all $h \in \mathbb{R}^n$. A sufficient condition for optimality can be obtained by strengthening the second order condition ; that is if $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) > 0$ ,Then $x^*$ is a strict local minimum of $f$ .

**Remark 3.1**

The idea of the proof is to consider any fixed vector $d \in \mathbb{R}^n$ , Then for any $\alpha$ sufficiently close to 0 , we have $x^* + \alpha d \in \Omega$ (unconstrained case ) and consider the one dimensional function $g\alpha = f(x^* + \alpha d)$ . Then $g$ has a minimum at $\alpha = 0$ And $g'(\alpha) = \nabla f(x^* + \alpha d) \cdot d$ , $g'(0) = \nabla f(x^*) \cdot d$ . To derive second order conditions , one has to expand $g$ up to order 2 .

**Remark 3.2**

Convex functions and convex sets plays an important role in optimization problems and quite often , it is easily tractable .The first order necessary condition is also sufficient , a local minimum is automatically a global one and there is uniqueness of the minimum. The main fact is that the function f lies above the linear function $L(x) = f(x^*) + \nabla f(x^*) \cdot (x - x^*)$ which satisfies $L(x) = f(x^*)$

. Further the numerical algorithms like steepest descent (gradient) method converges to $x^*$ satisfying $\nabla f(x^*) = 0$ and leads to global minimum in convex problems. A point $x^*$ satisfying $\nabla f(x^*) = 0$ is called a stationary point.

## 3.9 Constrained Optimization and Lagrange Multipliers[27]

Quite often one may not minimize over all points in a full neighborhood , but may be minimizing with a constraint say on a surface (manifold) of dimension less than $n$ .

**Example 3.4**. **(Linear Constraint**).

Consider a simple problem of minimizing $f(x,y) = x^2 + y^2$ with the constraint $y = x + 1$ . In other words, minimizing along the straight line $y = x + 1$, not on the whole space $\mathbb{R}^2$.

**Method 1 (Reduction Method)** The idea is to reduce the number of variables and minimize over $\mathbb{R}$ . Define $g(x) = f(x, x + 1)$. It is easy to see that a minimization of $g$ over $\mathbb{R}$ yields the minimum is achieved at $\left(-\frac{1}{2}, \frac{1}{2}\right)$ and the minimum value is $\frac{1}{2}$ .

**Method 2:** One can add the constraint with the minimizing function with the help of a new variable λ . Define

$$F(x,y,\lambda) = x^2 + y^2 - \lambda(y - x - 1)$$

Now treat $F$ as an unconstrained minimization problem over $\mathbb{R}^3$ instead of $\mathbb{R}^2$ . By applying $\frac{\partial F}{\partial x} = 0$ , $\frac{\partial F}{\partial y} = 0$, $\frac{\partial F}{\partial \lambda} = 0$ one obtains $\lambda = -1$ and $x = -1/2$ , $y = 1/2$ as obtained earlier $y = x + 1$.

### 3.9.1 Linear Constraints

Let $f: \mathbb{R}^n \to \mathbb{R}$ to be minimized over a plane

$$L = \left\{x \in \mathbb{R}^n : x = x_0 + \sum_{i=1}^{k} \alpha_i u_i \, , \alpha_i \in \mathbb{R}\right\}$$

Of dimension $k$ , where $u_1, \dots, u_k$ are $k$-linearly independent vectors in $\mathbb{R}^n$. The space $N$ normal to $L$ will be spanned by $n-k$ vectors , say $b_1, \dots, b_{n-k}$ . Then we can write

$$L = \{x \in \mathbb{R}^n : \langle b_i , x - x_0 \rangle = 0, 1 \le i \le n-k\}$$

**Theorem 3.5**

Let $f \in C^2(\mathbb{R})$ and $x^*$ minimizes $f$ subject to $k$ linear constraints

$$g_i(x) = c_i + \sum_{j=1}^{k} b_{ij} x_j = c_i + \langle b_i, x \rangle = 0, \quad 1 \le i \le k \tag{3.10}$$

Then there exists $k$ multipliers $\lambda_1, \dots, \lambda_k$ such that

$$\nabla F(x^*) = 0, \ \langle \nabla^2 F(x^*) h, h \rangle \ge 0 \tag{3.11}$$

For $h \ne 0$ and orthogonal to $b_1, \dots, b_k$ . Here $F$ is the augmented function

$$F(x) = f(x) + \sum_{i=1}^{k} \lambda_i \cdot g_i(x)$$

Conversely if $\lambda_1, \dots, \lambda_k$ and $x^*$satisfying $\nabla F(x^*) = 0$ and $\langle \nabla^2 F(x^*) h, h \rangle \ge 0$ for all $h \ne 0$ orthogonal to $b_1, \dots, b_k$ , then $x^*$ is a local minimum of $f$ satisfying the constraints .

**3.9.2 Non- Linear Constraints**

Let $\Omega$ be a surface in $\mathbb{R}^n$ with the equality constraints

$$h_1(x) = \cdots = h_k(x) = 0$$

Where $h_i \in C^1(\mathbb{R}^n, \mathbb{R})$ . Let $x^* \in \Omega$ be a local minimum of $f$ over $\Omega$ . Assume $x^*$ is a regular point , that is the set $\{\nabla h_i(x^*), 1 \le i \le k\}$ is an independent set in $\mathbb{R}^n$. First order necessary condition (Lagrange multipliers ): Since it is surface one need to consider curves lying in $\Omega$ passing through $x^*$ then the line segments .Let $x(\alpha)$ be a curve in $\Omega$ such that $x(0) = x^*$ and let $g(\alpha) \coloneqq f(x(\alpha))$ , where $\alpha$ is small parameter varying in the neighborhood of 0 . A simple calculation will give us

$$g'(\alpha) = \nabla f(x(\alpha)) \cdot x'(\alpha)$$

And at $\alpha = 0$,we have

$$g'(0) = \nabla f(x^*) \cdot x'(0) = 0 \qquad (3.12)$$

Since $x^*$ is a minimum over $\Omega$ . Note that , the vector $x'(0)$ is a tangent vector in the tangent space $T_{x^*}\Omega$ . Since $h_i(x(\alpha)) = 0$ for all $\alpha$ and $i$ , a further calculation will show that

$$\nabla h_i(x^*) \cdot x'(0) = 0 \qquad (3.13)$$

In fact , the converse is also true . That is if $d \in \mathbb{R}^n$ satisfying $\nabla h_i(x^*) \cdot d = 0$ $i = 1, \dots, k$ , then $d$ is a tangent vector to $\Omega$ at $x^*$ corresponding to some curve . In other words , tangent vectors to $\Omega$ are exactly the vector $d$ satisfying (3.11) , that is $T_{x^*}\Omega = \{d \in \mathbb{R}^n : \nabla h_i(x^*) \cdot d = 0\}$ .In view of this characterization ,we can rewrite the condition (3.11) as

$$\nabla f(x^*) \in \text{span}\{\nabla h_i(x^*) : 1 \le i \le k\} \qquad (3.14)$$

The condition (3.13) ,implies the existence of Lagrange multiplier $\lambda_1, \dots, \lambda_k$ such that

$$\nabla f(x^*) + \sum_{i=1}^{k} \lambda_i \nabla h_i(x^*) = 0 \qquad (3.15)$$

## 3.10 Optimal Control problem

### 3.10.1 Cost Functional [77]

A control system consists a system of constraints given by

$$\dot{x}(t) = f\big(t, x(t), u(t)\big), x(t_0) = x_0 \qquad (3.16)$$

Here $x = x(t) \in \mathbb{R}^n$ is called the state , $u(t) \in U \subset \mathbb{R}^m$ is the control , $t \in \mathbb{R}$is the time variable , $t_0$ is the initial time with initial state $x_0 \in \mathbb{R}^n$.

Together with control system , there will be an associated cost functional. The optimal control problem is to minimize this cost according to the dynamic (3.16).

We Will consider cost functional of the form

$$J(u) = \int_{t_0}^{t_f} L(t, x(t), u(t)dt + K(t_f, x_f) \qquad (3.17)$$

Here $t_f$ ,and $x_f \coloneqq x(t_f)$ are the final (or terminal ) time and state , $L$: $\mathbb{R}\times \mathbb{R}^n \times U \rightarrow \mathbb{R}$ is the running cost (or Lagrangian) and $K: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$ is the terminal cost .Since the cost depends on the final time as well as on the control , it would be more accurate to write $J(t_0, x_0, t_f, u)$ but we write $J(u)$ .For simplicity and to reflect the fact that the cost is being minimized over the space control functions. Optimal control problems in which the cost are known or problems in the Bolza form , or collectively as the Bolza problem there are two important special cases of the first one is the Lagrange problem , in which there is no terminal cost : $L \equiv 0$ . Indeed, given a problem with a terminal cost K ,we can write

$$K\big(t_f, t_f\big) = K(t_0, x_0) + \int_{t_0}^{t_f} \frac{d}{dt} K\big(t, x(t)\big)dt$$

$$=K(t_0,x_0) + \int_{t_0}^{t_f} K_t\big(t, x(t)\big), K_x\,(t, x(t)) \cdot f\big(t, x(t), u(t)\big)dt$$

Since K$(t_0, x_0)$is constant independent of u ,we arrive at equivalent problem in the Lagrange from with

$K_t(t,x), K_x(t,x) \cdot f(t,x,u)$ added to the original running cost . One the other hand , given a problem with running cost L satisfying the same regularity conditions as $f$, we can introduce an extra state variable $x^0$

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

This yields

$$\int_{t_0}^{t_f} L(t, x(t), u(t)dt = x^0\,(t_f)$$

Thus converting the problem to the Mayer form. Note that the similar trick of introducing the additional state variable $x_{n+1} \coloneqq t$ eliminat

S the dependence of L and / or K on time ; for the Bolza problem this gives

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

## 3.11 Necessary Conditions[10]

If $u^*(t)$ , $x^*(t)$ are optimal ,then the following conditions hold ..

For now, let us derive the necessary conditions. Express our objective functional in terms of the control:

$$J(u) = \int_{t_0}^{t_1} f(t, x(t), u(t))dt$$

where$x = x(u)$ is the corresponding state.

Assume a (piecewise continuous) optimal control exists, and that $u^*$ is such a control ,with $x^*$ the corresponding state .Namely , $J(u) \leq J(u^*) < \infty$ for all controls $u$ . Let $h(t)$ be a piecewise continuous variation function and $\epsilon \in \mathbb{R}$a constant .then

$$u^\epsilon(t) = u^*(t) + \epsilon h(t)$$

Is anther piecewise continuous control

Let $x^\epsilon$by the state corresponding to the control $u^\epsilon$, namely , $x^\epsilon$ satisfies

$$\frac{d}{dt} x^\epsilon(t) = g\big(t, x^\epsilon(t), u^\epsilon(t)\big) \qquad (3.18)$$

It easily seen that , $u^\epsilon(t) \to u^*(t)$ for all $t$ as $\epsilon \to 0$

$$\left.\frac{du^{\epsilon}(t)}{d\epsilon}\right|_{\epsilon=0} = h(t)$$

In fact , something similar is true for $x^{\epsilon}$ . Because of the assumptions mode on $g$ , it follows that

$$x^{\epsilon}(t) \to x^{*}(t)$$

For each $t$. Further, the derivative

$$\left.\frac{\partial}{\partial\epsilon}x^{\epsilon}(t)\right|_{\epsilon=0}$$

Exists for each $t$. Then actual value of quantity will prove unimportant . we need only to know that it exists .

The objective functional at $u^{\epsilon}$ is

$$J(u^{\epsilon}) = \int_{t_0}^{t_1} f(t, x^{\epsilon}(t), u^{\epsilon}(t))dt$$

To introduce the adjoint function or variable or variable $\lambda$ . Let $\lambda(t)$ be a piecewise differentiable function on $[t_0, t_1]$ to be determined . By the fundamental theorem of calculus ,

$$\int_{t_0}^{t_1} \frac{d}{dt}[\lambda(t)x^{\epsilon}(t)]dt = \lambda(t_1)x^{\epsilon}(t_1) - \lambda(t_0)x^{\epsilon}(t_0)$$

$$\int_{t_0}^{t_1} \frac{d}{dt}[\lambda(t)x^{\epsilon}(t)]dt + \lambda(t_0)x_0 - \lambda(t_1)x^{\epsilon}(t_1) = 0$$

Adding this 0 expression to our $J(u^{\epsilon})$ gives

$$J(u^{\epsilon}) = \int_{t_0}^{t_1} \left[f(t, x^{\epsilon}(t), u^{\epsilon}(t)) + \frac{d}{dt}(\lambda(t)x^{\epsilon}(t)\right] dt + \lambda(t_0)x_0 - \lambda(t_1)x^{\epsilon}(t_1)$$

$=\int_{t_0}^{t_1}\left[f\big(t, x^\epsilon(t), u^\epsilon(t)\big) + \lambda'(t)x^\epsilon(t) + \lambda(t)g\big(t, x^\epsilon(t), u^\epsilon(t)\big)\right] dt +$
$\lambda(t_0)x_0 - \lambda(t_1)x^\epsilon(t_1)$

We used $g(t, x^\epsilon, , u^\epsilon) = \frac{d}{dt} x^\epsilon$.Since the maximum of $J$ with reset to the control $u$ occurs at $, u^*$, the derivative of $J(u^\epsilon)$ with respect to $\epsilon$ is zero , i.e.,

$$0 = \frac{d}{d\epsilon}J(u^\epsilon)\Big|_{\epsilon=0} = \lim_{\epsilon=0}\frac{J(u^\epsilon) - J(u^*)}{\epsilon}$$

Therefore,

$$0 = \frac{d}{d\epsilon}J(u^\epsilon)\Big|_{\epsilon=0}$$

$$=\int_{t_0}^{t_1}\frac{\partial}{\partial\epsilon}\left[f\big(t, x^\epsilon(t), u^\epsilon(t)\big) + \lambda'(t)x^\epsilon(t) + \lambda g\big(t, x^\epsilon(t), u^\epsilon(t)\big)dt\right]\Big|_{\epsilon=0} - \frac{\partial}{\partial\epsilon}\lambda(t_1)x^\epsilon(t_1)\Big|_{\epsilon=0}$$

Applying the chain rule to $f$and $g$ it follows

$$0 = \int_{t_0}^{t_1}\left[f_x\frac{\partial x^\epsilon}{\partial\epsilon} + f_u\frac{\partial u^\epsilon}{\partial\epsilon} + \lambda'(t)\frac{\partial x^\epsilon}{\partial\epsilon} + \lambda(t)(g_x\frac{\partial x^\epsilon}{\partial\epsilon} + g_u\frac{\partial u^\epsilon}{\partial\epsilon}\right]\Big|_{\epsilon=0} - \lambda(t_1)\frac{\partial x^\epsilon}{\partial\epsilon}(t_1) \qquad (3.19)$$

where the arguments of the $f_x$ ,$f_u$ ,$g_x$ and $g_u$ terms are $\big(t, x^*(t), u^*(t)\big)$ in (3.19) gives

$$0 = \int_{t_0}^{t_1}\left[\big(f_x + \lambda(t)g_x + \lambda'^{(t)}\big)\frac{\partial x^\epsilon}{\partial\epsilon}\Big|_{\epsilon=0} + (f_u + \lambda(t)g_u)h(t)\right] - \lambda(t_1)\frac{\partial x^\epsilon}{\partial\epsilon}(t_1)\Big|_{\epsilon=0} \qquad (3.20)$$

We want to choose the adjoint function to simplify (3.20) by making the coefficients of

$$\frac{\partial x^{\epsilon}}{\partial \epsilon}(t_1)\Big|_{\epsilon=0}$$

Thus, we choose the adjoint function , $\lambda(t)$ to satisfy

$\lambda'(t) = -\left[f_x\left(t, x^*(t), u^*(t) + \lambda(t) g_x(t, x^*(t), u^*(t))\right)\right]$ (adjoint equation) ,

And the boundary condition

$\lambda'(t_1) = 0$ ( transversality condition)

Now(3.19) reduces to

$$o = \int_{t_0}^{t_1} f_u\left(t, x^*(t), u^*(t) + \lambda(t) g_u\left(t, x^*(t), u^*(t)\right)\right) h(t)dt$$

As this for any continuous variation function $h(t)$ it holds for

$$h = f_u\left(t, x^*(t), u^*(t) + \lambda(t) g_u\left(t, x^*(t), u^*(t)\right)\right)$$

In this case

$$0 = \int_{t_0}^{t_1} f_u\left(t, x^*(t), u^*(t) + \lambda(t) g_u\left(t, x^*(t), u^*(t)\right)\right)^2 dt$$

Which implies the optimality condition

$f_u\left(t, x^*(t), u^*(t) + \lambda(t) g_u\left(t, x^*(t), u^*(t)\right)\right) = 0$

for all $t_0 \leq t \leq t_1$.

These equations form a set of necessary conditions that an optimal control and state must satisfy. In practice, one does not need to rederive the above equations in this way for a particular problem. In fact, we can generate the above necessary conditions from the Hamiltonian $H$, which is defined as follows ,

$$H(t,x,u,\lambda) = f(t,x,u) + \lambda g(t,x,u)$$

We are maximizing $H$ with respect to $u$ to $u^*$ and the above condition can be written in terms of the Hamilton:

$$\frac{\partial H}{\partial u} = 0 \text{ at } u^* \Rightarrow f_u + \lambda g_u = 0 \qquad \text{( optimality condition)}$$

$$\lambda' = -\frac{\partial H}{\partial x} \Rightarrow \lambda' = -(f_x + \lambda g_x) \qquad \text{(adjoint equation)}$$

$$\lambda(t_1) = 0 \qquad \text{(transversality equation)}$$

We are given the dynamic of the state equation :

$$x' = g(t,x,u) = \frac{\partial H}{\partial \lambda}, \; x(t_0) = x_0$$

## 3.12 Pontryagin's Maximum Principle (PMP)[10]

Let $u^*: [t_0, t_1] \to U$ be an optimal control and $x^*: [t_0, t_1] \to \mathbb{R}^n$ be the corresponding optimal state trajectory .

**Theorem 3.6**

If $u^*(t)$and $x^*(t)$ are optimal for problem

$$\max_u \int_{t_0}^{t_1} f\big(t, x(t), u(t)\big) dt$$

$$\text{Subject to } x'(t) = g\big(t, x(t), u(t)\big) \qquad (3.21)$$

$$x(t_0) = x_0 \text{ and } x(t_1) \text{ free .} \qquad (3.22)$$

Then There exists a piecewise differentiable adjoint variable $\lambda(t)$ such that

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

For all controls $u$ at each time $t$ ,where the Hamiltonian H is

$$H = f\big(t, x(t), u(t)\big) + \lambda g(t, x(t), u(t)) \text{ and}$$

$$\lambda'(t) = -\frac{\partial(t, x^*(t), u^*(t), \lambda(t))}{\partial x}$$

$$\lambda'(t_1) = 0$$

**Theorem 3.7**

Suppose that $f(t, x, u)$ and $g(t, x, u)$ are both continuously differentiable functions in their three arguments and concave in $u$ . suppose $u^*$is an optimal control for problem (3.20) ,with associated state $x^*$ , and $\lambda$ a piecewise differentiable function with $\lambda(t) \geq 0$ for all $t$ . Suppose for all $t_0 \leq t \leq t_1$ .

$$0 = H_u((t, x^*(t), u^*(t), \lambda(t))$$

Then for all controls $u$ and each $t_0 \leq t \leq t_1$ , we have

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

**Proof**

Fix a control $u$ and a point in time $t_0 \leq t \leq t_1$ .Then

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

$$= \big[f\big(t, x^*(t), u^*(t)\big) + \lambda g(t, x^*(t), u^*(t)\big] -$$

$$\big[f\big(t, x^*(t), u(t)\big) + \lambda g(t, x^*(t), u(t))\big]$$

$$= f\big(t, x^*(t), u^*(t)\big) - f\big(t, x^*(t), u(t)\big)$$
$$+ \lambda(t)\big[g\big(t, x^*(t), u^*(t)\big) - g(t, x^*(t), u(t)\big]$$

$$\geq \big(u^*(t) - u(t)\big) f_u\big(t, x^*(t), u^*(t)\big) + \lambda(t)\big(u^*(t) - u(t)\big) g_u\big(t, x^*(t), u^*(t)\big)$$

$$= \big(u^*(t) - u(t)\big) H_u(\big(t, x^*(t), u^*(t), \lambda(t)\big) = 0 \ .$$

We can also check concavity conditions to distinguish between controls that maximize and those that minimize the objective functional . If

$$\frac{\partial^2 H}{\partial u^2} < 0 \text{ at } \ u^*$$

Then the problem is maximization ,while

$$\frac{\partial^2 H}{\partial u^2} > 0 \text{ at } \ \ u^*,$$

Goes with minimization .

When we are able to solve for the optimal control in terms of $x^*$ and $\lambda$, we will call that formula for $u^*$ the characterization of the optimal control. The state equations and the adjoint equations together with the characterization of the optimal control and the boundary conditions are called the optimality system.

**Example 3.5**

$$\min_u \int_0^1 u(t)^2 dt$$

Subject to $x'(t) = x(t) + u(t)$ , $x(0) = 1, x(1)$ free

Can we see what the optimal control should be? The goal of the problem is to minimize this integral, which does not involve the state. Only the integral of control (squared) is to be minimized. Therefore, we expect the optimal control is 0. We verify with the necessary conditions.

We begin by forming the Hamiltonian

$$H = u^2 + \lambda(x + u)$$

The optimality condition is

$$0 = \frac{\partial H}{\partial u} = 2u + \lambda \text{ at } u^* \Rightarrow u^* = -\frac{1}{2}\lambda$$

We see problem is indeed minimization as

$$\frac{\partial^2 H}{\partial u^2} = 2 > 0$$

The adjoint equation is given by

$$\lambda'(t) = -\frac{\partial H}{\partial x} = -\lambda \Rightarrow \lambda(t) = Ce^{-t} = 0 \Rightarrow C = 0$$

So thus $\lambda \equiv 0$ so that $u^* = -\lambda/2 = 0$ so $x^*$ satisfies $x' = x$ and $x(0) = 1$

Hence, the optimal solutions are

$$\lambda \equiv 0, \ u^* = 0 , \ x^* = e^t .$$

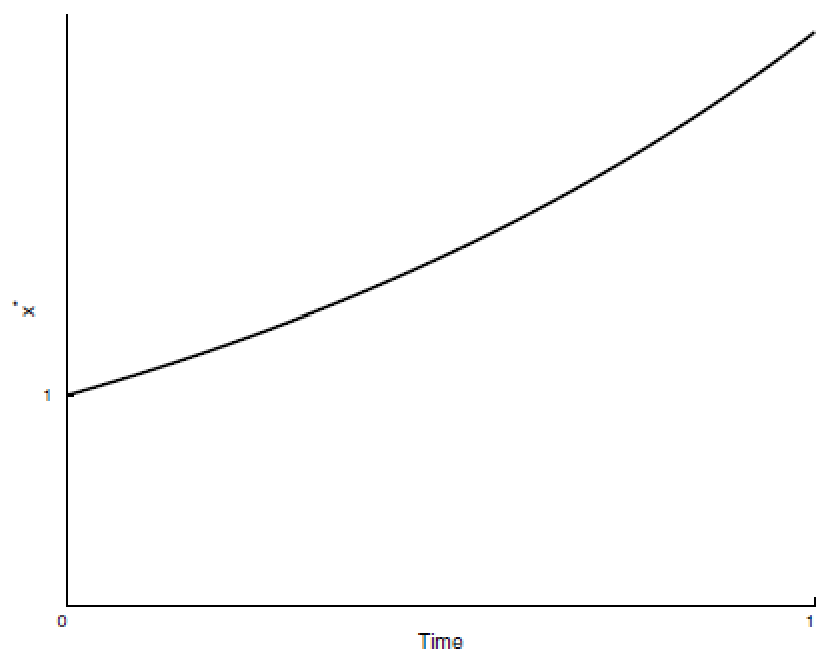



**Figure 3.1** Optimal state for Example 3.5 plotted as a function of time

**Example 3.6**

$$\min_{u} \frac{1}{2}\int_{0}^{1} 3x(t)^2 + u(t)^2 dt$$

Subject to $x'(t) = x(t) + u(t)\,, x(0) = 1$

1

The $\frac{1}{2}$ which appears before the integral will have no effect on the minimizing control and, thus, no effect on the problem. It is inserted in order to make the computations slightly neater. You will see how shortly. Also, note we have omitted the phrase “$x(1)$ free” from the statement of the problem .This is standard notation, in that a term which is unrestricted is simply not mentioned. We adopt this convention from now on.

Form the Hamiltonian of the problem

$$H = \frac{3}{2}x^2 + \frac{1}{2}u^2 + x\lambda + u\lambda$$

The optimality condition gives

$$0 = \frac{\partial H}{\partial u} + u + \lambda \text{ at } u^* \Longrightarrow u^* = -\lambda$$

Notice $\frac{1}{2}$ cancels with the 2 which comes from the square on the control $u$

Also, the problem is a minimization problem as

$$\frac{\partial^2 H}{\partial u^2} = 1 > 0$$

We use the Hamiltonian to find a differential equation of the adjoint $\lambda$ ,

$$\lambda'(t) = -\frac{\partial H}{\partial x} = -3x - \lambda\,, \lambda(1) = 0$$

Substituting the derived characterization for the control variable $u$ in the equation for $x'$ , we arrive at

$$\begin{pmatrix} x \\ \lambda \end{pmatrix}' = \begin{pmatrix} 1 & -1 \\ -3 & -1 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix}$$

The eigenvalues of the coefficient matrix are 2 and -2. Finding the eigenvectors

, the equations for $x$ and $\lambda$ ¸ are

$$\binom{x}{\lambda}(t) = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{2t} + c_2 \begin{pmatrix} 1 \\ 3 \end{pmatrix} e^{-2t}$$

Using $x(0) = 1$ and $\lambda(1) = 0$, we find $c_1 = 3c_2e^{-4}$ and $c_2 = \frac{1}{3e^{-4}+1}$

Thus using the optimality equation, the optimal solutions are

$$u^*(t) = \frac{3e^{-4}}{3e^{-4}+1}e^{2t} - \frac{3}{3e^{-4}+1}e^{-2t}$$

$$x^*(t) = \frac{3e^{-4}}{3e^{-4}+1}e^{2t} + \frac{3}{3e^{-4}+1}e^{-2t}$$

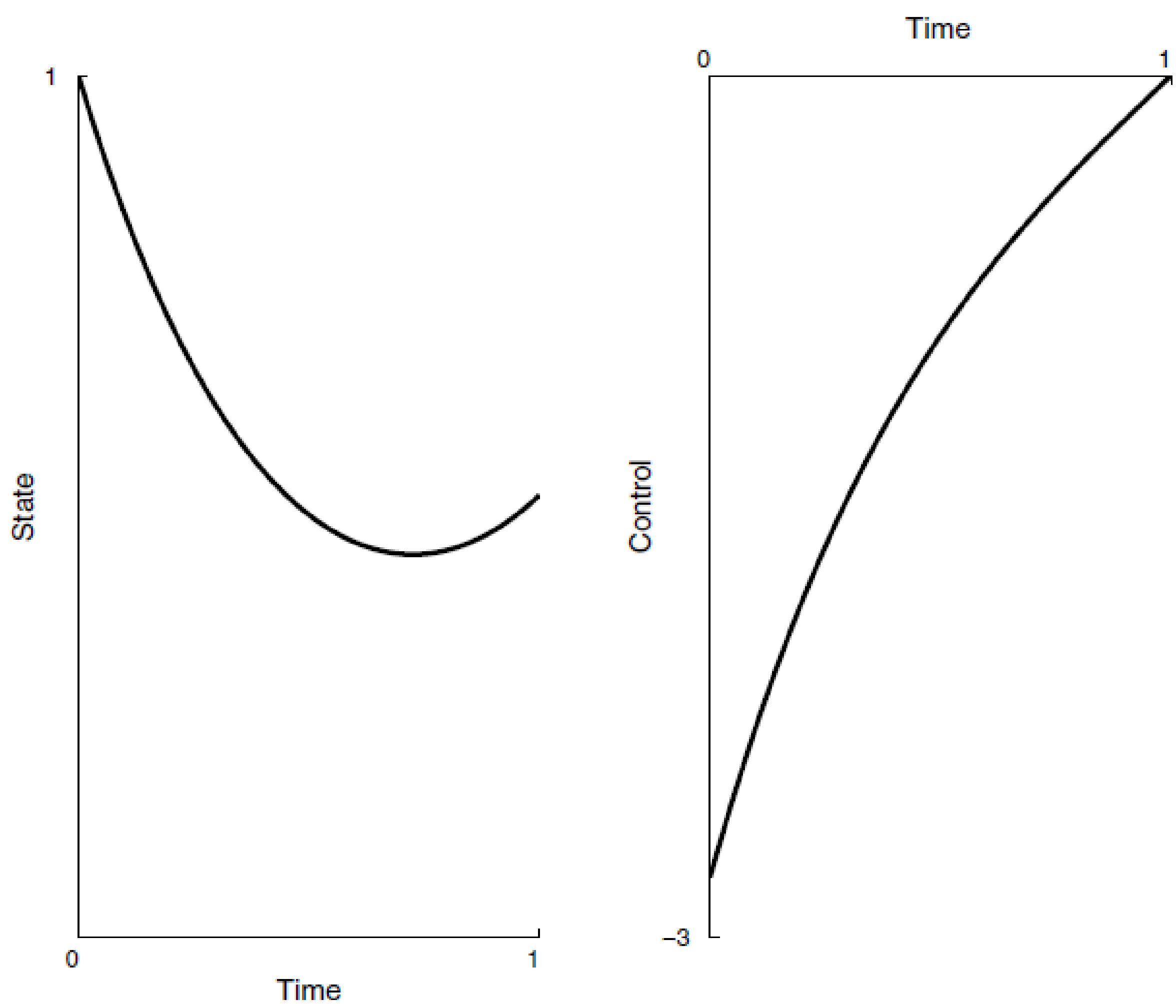



**Figure 3.2**: optimal control and state for Example 3.6

## 3.13 Existence and Uniqueness Results[10]

**Theorem 3.8**

Consider

$$J(u) = \int_{t_0}^{t_1} f\big(t, x(t), u(t)\big)dt$$

Subject to $x'(t) = g\big(t, x(t), u(t)\big),\ x(t_0) = x_0$

Suppose that $f(t, x, u)$ and $g(t, x, u)$ are both continuously differentiable functions In their three arguments and concave in $x$ and $u$ . Suppose $u^*$is a control , with associated state $x^*$ and λ a piecewise differentiable function , such that $u^*$ , $x^*$ and λ together satisfy on $t_0 \le t \le t_1$ :

$$f_u + \lambda g_u = 0$$

$$\lambda' = -(f_x + \lambda g_x)$$

$$\lambda(t_1) = 0$$

$$\lambda(t) \ge 0$$

Then for all controls $u$ , we have

$$J(u^*) > J(u)$$

**Proof**

Let $u$ be any control , and $x$ its associated state . Note ,as $f(t, x, u)$ is concave in both the $x$ and $u$ variable , we have by the tangent line property

$$f(t, x^*, u^*) - f(t, x, u) \ge (x^* - x)f_x(t, x^*, u^*) + (u^* - u)f_u(t, x^*, u^*)$$

This gives

$$J(u^*) - J(u) = \int_{t_0}^{t_1} f(\ t, x^*, u^*) - f(t, x, u)dt$$

$$\ge \int_{t_0}^{t_1} (x^*(t) - x(t))\ f_x(t, x^*, u^*) \qquad (3.23)$$

$$+(u^*(t) - u(t))f_u(t, x^*, u^*)dt$$

Substituting

$$f_x(t, x^*, u^*) = -\lambda'(t) - \lambda(t) g_x(t, x^*, u^*) \text{ and}$$

$$f_u(t, x^*, u^*) = -\lambda(t) g_u(t, x^*, u^*)$$

As given the hypothesis , the last term in (3.23) becomes

$$\int_{t_0}^{t_1} \big(x^*(t) - x(t)(-\lambda(t) g_x(t, x^*, u^*) - \lambda'(t)\big)$$

$$+ \big(u^*(t) - u(t)\big)(-\lambda(t) g_u(t, x^*, u^*) dt.$$

Using integration by parts , and recalling $\lambda(t_1) = 0$ and $x(t_0) = x^*(t_0)$ we see

$$\int_{t_0}^{t_1} -\lambda'(t)\big(x^*(t) - x(t)\big) dt = \int_{t_0}^{t_1} \lambda(t)\big(x^*(t) - x(t)\big)' dt$$

$$= \int_{t_0}^{t_1} \lambda(t) g\big(t, x^*(t), u^*(t)\big) - g(t, x(t), u(t) dt.$$

Making this substitution,

$$J(u^*) - J(u \geq \int_{t_0}^{t_1} \lambda(t)[g(t, x^*, u^*) - g(t, x, u) - (x^* - x) g_x(t, x^*, u^*)$$

$$- (u^* - u) g_u(t, x^*, u^*)]\, dt.$$

Taking into account $\lambda(t) \geq 0$ and that $g$is concave in both $x$ and $u$ , this gives the desired result $J(u^*) - J(u) \geq 0$.

**Theorem 3.9**

Let the set of controls for problem (3.20)be Lebesgue integrable functions (instead of just piecewise continuous functions ) on $t_0 \leq t \leq t_1$ with values in $\mathbb{R}$ . suppose that $f(t, x, u)$ is convex in $u$ and there exist constants $C_4$ and $C_1, C_2, C_3 > 0$and $\beta > 1$ such that

$$g(t, x, u) = \alpha(t, x) + \beta(t, x) u$$

$$|g(t, x_1, u) - g(t, x, u)| \leq C_2 |x_1 - x|(1 + |u|)$$

$$f(t, x, u) \geq C_3 |u|^\beta - C_4$$

For all $t$ with $t_0 \leq t \leq t_1$ , $x$ , $x_1$ , $u$ in $\mathbb{R}$ .then there exists an optimal control $u^*$ maximizing $J(u)$with $J(u^*)$ finite .

Also of interest is the idea of uniqueness. Suppose an optimal control exists, i.e, there is $u^*$ such that $J(u) \leq J(u^*) < \infty$ for all controls $u$ (in the maximization case). We say $u^*$ is unique if whenever $J(u^*) = J(u)$, then $u^* = u$ at all but finitely many points. In this case, the associated states will be identical. We call this state, $x^*$ , the unique optimal state. Clearly, uniqueness of solutions of the optimality system implies uniqueness of the optimal control, if one exists. We can frequently prove uniqueness of the solutions of the optimality system, but only for a small time interval. This small time condition is due to opposite time orientations of the state equation and adjoint equation, meaning the state equation has an initial time condition and the adjoint equation has a final time condition. However, in general, uniqueness of the optimal control does not necessarily guarantee uniqueness of the optimality system. To prove uniqueness of the optimal control directly, strict concavity of the objective functional $J(u, x(u))$ must be established. Direct uniqueness results tend to be cumbersome and difficult to state, and, as they will not be needed here, they will not be treated. If $f$ , $g$, and the right hand side of the adjoint equation are Lipschitz in the state and adjoint variables, then the uniqueness of solutions of the optimality system holds for small time.

## 3.14 Dynamic Programming and the HJB Equation[10]

### 3.14.1 Principle of Optimality

An important result in both optimal control and dynamic programming is the Principle of Optimality. It concerns optimizing a system over a subinterval of the original time span, and in particular, how the optimal control over this smaller interval relates to the optimal control on the full time period.

**Theorem 3.10**

Let $u^*$ be an optima control , and $x^*$ the resulting stat , for the problem

$$\max_u J(u) = \max_u \int_{t_0}^{t_1} f(t, x(t), u(t))dt$$

$$\text{Subject to } x'(t) = g\big(t, x(t), u(t)\big),\ x(t_0) = x_0 \qquad (3.24)$$

Let $\hat{t}$ be a fixed point in time such that $t_0 < \hat{t} < t_1$ . Then , the restricted functions $\hat{u}^* = u^*|_{[\hat{t},t_1]}$ , $\hat{x}^* = x^*|_{[\hat{t},t_1]}$ , form an optimal pair for the restricted

$$\max_u \hat{J}(u) = \max_u \int_{t_0}^{t_1} f(t, x(t), u(t))dt$$

$$\text{Subject to } x'(t) = g\big(t, x(t), u(t)\big),\ x(\hat{t}) = x^*(\hat{t}) \qquad (3.25)$$

Further, if $u^*$ is the unique optimal control for (3.23), then $\hat{u}^*$ is the unique optimal control for (3.24).

**Proof**

This proof is done by contradiction. Suppose, to the contrary, that $\hat{u}^*$ is not optimal, i.e., there exists a control $\hat{u}_1$ on the interval $[\hat{t}, t_1]$ such that $\hat{J}(\hat{u}_1) > \hat{J}(\hat{u}^*)$ Construct a new control u1 on the whole interval $[t_0, t_1]$ as follows

$$u_1(t) = \begin{cases} u^*(t) \text{ for } t_0 \le t \le \hat{t} \\ \hat{u}_1(t) \text{ for } \hat{t} < t \le t_1 \end{cases}$$

Let $x_1$ be the state associated with control $u_1$. Notice that $u_1$ and $u^*$ agree on $[t_0\ , \hat{t}]$, so that $x_1$ and $x^*$ will also agree there. Hence,

$$J(u_1) - J(u^*) = \left(\int_{t_0}^{\hat{t}} f(t, x_1, u_1)dt + \hat{J}(\hat{u}_1)\right) - \left(\int_{t_0}^{\hat{t}} f(t, x^*, u^*)dt + \hat{J}(\hat{u}^*)\right)$$

$$= \hat{J}(\hat{u}_1) - \hat{J}(\hat{u}^*)$$

$$> 0.$$

However, this contradicts our initial assumption that $u^*$ was optimal for (3.24). Thus, no such control $\hat{u}_1$ exists, and $\hat{u}^*$ is optimal for (3.25).

**Example 3.7**

$$\min_u \int_0^2 x(t) + \frac{1}{2}u(t)^2 dt$$

$$\text{Subject to } x'(t) = x(t) + u(t)\,, x(0) = \tfrac{1}{2}e^2 - 1$$

First, we will solve this example on [0, 2] , then solve the same problem on a smaller interval [1, 2] . The Hamiltonian in this example is

$$H = x + \frac{1}{2}u^2 + x\lambda + u\lambda$$

The adjoint equation and transversality condition give

$$\lambda' = -\tfrac{\partial H}{\partial x} = -1 - \lambda\,,\ \lambda(2) = 0 \Longrightarrow \lambda(t) = e^{2-t} - 1$$

And the optimality condition leads to

$$0 = \tfrac{\partial H}{\partial u} = u + \lambda \Rightarrow\ u^*(t) = -\lambda(t) = 1 - e^{2-t}$$

Finally , from the state equation , the associated state is

$$x^*(t) = \frac{1}{2}e^{2-t} - 1$$

Now , consider the same problem ,except on the interval [1,2] ,i.e.,

$$\min_u \int_1^2 x(t) + \frac{1}{2}u(t)^2 dt$$

$$\text{Subject to } x'(t) = x(t) + u(t), x(1) = \tfrac{1}{2}e - 1$$

Clearly, the Principle of Optimality can be applied to find an optimal pair immediately, namely, the pair found above. The original problem on the interval [0, 2] has the same optimal control as the above problem on [1, 2]. Let us solve this example by hand, though, to reinforce the power of the theorem. The Hamiltonian will be the same, regardless of interval. Because the end point remains fixed , the adjoint equation and transversality also remain the same:

$$\lambda' = -\frac{\partial H}{\partial x} = -1 - \lambda\,, \lambda(2) = 0 \Rightarrow \lambda(t) = e^{2-t} - 1$$

While the optimality is also unchanged

$$0 = \frac{\partial H}{\partial u} = u + \lambda \Rightarrow u^*(t) = -\lambda(t) = 1 - e^{2-t}$$

Using the new initial condition $x(1) = \frac{1}{2}e - 1$, we find the corresponding state

$$x^*(t) = \frac{1}{2}e^{2-1} - 1.$$

**3.14.2 The Value Function and Dynamic Programming Principle[27]**

We consider the cost functional

$$J(u) = J(t_0, x_0, u) = \int_{t_0}^{t_1} L\big(t, x(t), u(t)\big)dt + K(x(t_1)) \qquad (3.26)$$

The idea is to vary $t_0$ and $x_0$ and introduce the family of minimization problems

$$J(t, x, u) = \int_{t_0}^{t_1} L(s\ , x(s), u(s)dt + K(x(t_1) \qquad (3.27)$$

Where $t \in [t_0\ , t_1)$ , $x \in \mathbb{R}^n$ and $x(s)$ is the solution to the ODE with the initial condition $x(t) = x$ solved for $s \in [t_0\ , t_1]$ . Introduce , the value function $V: [t_0\ , t_1) \times \mathbb{R}^n \to \mathbb{R}$

$$V(t, x) = \inf_u J(t, x, u) \qquad (3.28)$$

Which is optimal cost to go from state $x$ at time $t$ to the state $x_1$ at time $t_1$. Indeed $V$ satisfies the final condition

$$V(t_1, x) = K(x), \text{ for all } x \in \mathbb{R}^n$$

**Theorem 3.11**

For every $(t, x) \in [t_0, t_1) \times \mathbb{R}^n$ and every $\tau \in (t\ , t_1)$, the value function $V$ satisfies

$$V(t, x) = \inf_u \left\{ \int_t^{\tau} L\big(s, x(s), u(s)\big)ds + V(\tau\ , \mathrm{x}(\tau)) \right\} \qquad (3.29)$$

Where $x(s)$ is the trajectory corresponding to the control $u$ .

When an optimal control and trajectory exist , then the inf is achieved at the optimal solution and it will become a minimum . The infinitesimal version of DPP the Hamilton –Jacobi-Bellman equation (HJB) which is a first order PDE satisfied by $V$ and is given in $[t_0\,, t_1) \times \mathbb{R}^n$ by

$$V_t(t\,, x) = \sup_{u\in U}\{-L(t, x, u) - V_x(t, x)\cdot f(t, x, u)\}\,, V(t_1, x) = K(x) \qquad (3.30)$$

**Theorem 3.12**

Let $(\bar{x}\,, \bar{u})$ be a solution to the ODE of the associated optimal control problem . Suppose $\emptyset \in C^1$ is a solution to (3.29) that satisfies :$\emptyset(t_1\,, x) = K(x)$ and

$$\emptyset_x(t\,, \bar{x}(t))\cdot f\big(t, \bar{x}(t), \bar{u}(t)\big) + L\big(t, \bar{x}(t), \bar{u}(t)\big)$$
$$= \min_{u\in U}\{\emptyset_x(t\,, \bar{x}(t))\cdot f(t, \bar{x}(t), u) + L(t, \bar{x}(t), u)\} \qquad (3.31)$$

Which is equivalent to the Hamiltonian maximization . Then

1. $(\bar{x}\,, \bar{u})$ is an optimal solution

2. $\emptyset(t_0\,, x_0)$ is minimum cost

## 3.15 Hamilton Jacobi –Bellman (HJB) and Maximum Principle [27]

Recall That PMP is formulated using the canonical equations

$$\dot{x}^* = \mathcal{H}_\lambda \quad , \dot{\lambda}^* = -\mathcal{H}_x \qquad (3.32)$$

And says that at each time $t$ , the value $u^*(t)$ of the optimal control must maximize $\mathcal{H}\big(x^*(t), u, \lambda^*(t)\big)$ with respect to $u$ ; that

$$u^*(t) = \arg\ \max_{u\in U} \mathcal{H}\big(x^*(t), u, \lambda^*(t)\big) \qquad (3.33)$$

This is an open – loop control because to compute $u^*$ , we need not only the stat $x^*$ ,but need co –state $\lambda^*$as well which is given from the adjoint equation

Let us see the Hamiltonian associated with HJB . The inf in the equation (3.30) becomes a minimum for the optimal control . Thus we get

$$H(t, x^*(t), u^*(t) - V_x\big(t, x^*(t)\big) = \max_{u\in U} H(t, x^*(t), u - V_x\big(t, x^*(t)\big)\ , (3.34)$$

Where $H$ is given by $H(t,x,u,\lambda) \coloneqq \lambda \cdot f(t,x.u) - L(t,x,u)$. Thus , the optimal control must satisfy (assume $H$ is independent of $t$)

$$u^*(t) = \arg\ \max_{u\in U} H\big(x^*(t),u,\lambda^*(t)\big) \qquad (3.35)$$

This is a closed-loop (feed-back specification) quite useful in applications. If we know the value everywhere, $u^*$ is determined by the current state $x^*(t)$ . This feature of generating an optimal control policy from the state is the trade mark of Dynamic programming approach. However, the major drawback is that $V$ is obtained by solving HJB which is a very difficult task. Though this approach is more novel than PMP, the later one involves solving only ODEs and hence computationally more fruitful. Another question is the derivation of PMP from HJB. The equations (3.32) and (3.34) suggests that we should look for the co-state of the form $\lambda^*(t) = -V_x(t\ ,x^*(t))$ . The PMP follows if one proves that $\lambda^*$ satisfies the second equation in (3.31) which can be established as usual in the smooth case.

**Example 3.8 [77]**

Consider the standard integrator $\dot{x} = u$ (with $x,u\ \in \mathbb{R}$) and let $L(x,u) = x^4 + u^4$ .The corresponding HJB equation is

$$-V_t(t\ ,x) = \inf_{u\in\mathbb{R}}\{x^4 + u^4 + V_x(t,x)u\} \qquad (\ 3.36)$$

Since the expression inside the infimum is polynomial in $u$ , we can easily find the control that achieves the infimum : differentiating with respect to $u$ ,we have $4u^3 + V_x(t,x) = 0$ which yields $u = -(\frac{1}{4}V_x(t,x))^{1/3}$ . Plugging this control into the HJB equation (3.35) , we obtain

$$-V_t(t,x) = x^4 - 3(\tfrac{1}{4}V_x(t,x))^{4/3} \qquad (3.37)$$

Assuming that we can solve the PDE (3.36) for the value function $V$ , The optimal control is given in the state feedback from $u^*(t) = -(\frac{1}{4}V_x(t,x^*(t))^{1/3}$ .

## 3.16 The Linear Quadratic Regulator (LQR)[27]

The well –known LQR problem is the optimal control problem with linear constraint

$\dot{x}(t) = A(t)x + B(t)u$ , $x(t) \in \mathbb{R}^n$ , $u \in \mathbb{R}^n$ with the quadratic cost

$$J(u) = \int_{t_0}^{t_1} [x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)]\,dt + x^T(t_1)Mx(t_1)$$

Where $Q(\cdot)$ , $R\ (\cdot), M$ are matrices with

$$M = M^T \geq 0\ , Q(t) = Q^T \geq 0 \text{ and } R(t) = R^T > 0.$$

Here $f(t,x,u) = A(t)x + B(t)u$ and $L(t,x,u) = x^T Q(t)x + u^T R(t)u$ . All the results defined above are applicable in this case , but one can get further information . The Hamiltonian is given by

$$H(t,x,u,\lambda) = \lambda^T A(t)x + \lambda^T B(t)u - x^T Q(t)x - u^T R(t)u\ .$$

The gradient of $H$ with respect to $u$ is $H_u = B^T(t)\lambda - 2R(t)u$

Some computation will yield that an optimal control must satisfy

$$u^*(t) = \tfrac{1}{2} R^{-1}(t)B^T(t)\lambda^*(t)$$

Further since $H_{uu} = -2R(t) < 0$ the above control maximizes the Hamiltonian globally . Interestingly , one can do a further analysis on the co-state equation to get

$$\lambda^*(t) = -2P(t)x^*(t)$$

Where $P(t)$ is some matrix valued function to be determined . But the crucial point is that $\lambda^*$ and $x^*$ is related by a linear relation and $u^*$ then take the from

$$u^*(t) = \frac{1}{2} R^{-1}(t)B^T(t)P(t)x(t)$$

Which is a linear state feedback law

One can also deduce that $P(t)$ will satisfy linear matrix differential equation

$$\dot{P}(t) = -P(t)A(t) - A^T(t)P(t) - Q(t) + P(t)B(t)R^{-1}(t)B^TP(t)$$

Which is Known as Riccati Differential Equation (RDE).

**Example 3.9[10]**

$$\frac{1}{2}\min_u \int_0^T x(t)^2 + u(t)^2 dt$$

Subject to $x'(t) = u(t), x(0) = x_0$

In this case, all the matrices are scalars (size $1 \times 1$) and $S(T) = M = 0$ ,$A = 0$ ,$B = Q = R = 1$ . The Riccati equation is

$$-P' = 1 - P^2, P(T) = 0$$

Solving as a separable equation , and using partial fractions ,

$$\frac{1}{2}\ln\left|\frac{P-1}{P+1}\right| = \int \frac{P'}{P^2-1}dt = \int 1dt = t + C$$

Which along with $P(T) = 0$ gives

$$P(t) = \frac{1 - e^{2(t-T)}}{1 + e^{2(t-T)}}$$

The optimal control satisfies $u = -Px$ , so that optimal state satisfies $x' = -Px$ . Using partial fractions we can find an antiderivative of $P$ , and solve the separable equationto see

$$x(t) = C(e^{t-T} + e^{T-t})$$

Taking into account $x(0) = x_0$,

$$x^*(t) = x_o \frac{e^t + e^{2T-t}}{1+e^{2T}} \quad \text{and} \quad u^*(t) = x_0 \frac{e^t - e^{2T-t}}{1+e^{2T}} .$$

## 3.17 Discrete Time Optimal Control Systems [80]

We discussed the optimal control of continuous-time systems described by differential equations. There, we minimized cost functionals which are essentially integrals of scalar functions. Now, we know that discrete-time systems are characterized by difference equations, and we focus on minimizing the cost functionals which are summations of some scalar functions.

We obtain the necessary conditions for optimization of cost functionals which are summation such as

$$J(x(k_0), k_0) = J = \sum_{k=k_0}^{k_f-1} f(x(k), x(k+1), k)$$

Where ,the discrete instant $k = k_0, k_1, \ldots, k_f - 1$

Consider a linear ,time –varying ,discrete –time control system described by

$$x(k+1) = A(k)x(k) + B(k)u(k) \tag{3.38}$$

Where , $k = k_0, k_1, \ldots, k_f - 1$, $x(k)$ is $n$th order state vector , $u(k)$ is $r$th order control vector , and $A(k)$ and $B(k)$ are matrices of $n \times n$ and $n \times r$ dimensions , respectively . Note that we used $A$ and $B$ for the state space representation for discrete-time case as well as for the continuous-time case as shown in the previous chapters. One can alternatively use, say G and E for the discrete-time case so that the case of discretization of a continuous-time system with A and B will result in G and E in the discrete-time representation. However, the present notation should not cause any confusion once we redefine the matrices in the discrete-time case. We are given the initial condition as

$$x(k = k_0) = x(k_0) \tag{3.39}$$

# Chapter Four

# Turnpike Theory and its Properties

## 4.1 About the Turnpike Theory[85]

Turnpike Theory refers to a set of economic theories about the optimal path of accumulation (often capital accumulation ) in a system , depending on the initial and final levels . In the context of a macroeconomic exogenous growth model .For example , it says that if an infinite optimal path is calculated , and an economic planner wishes to move an economy from one level of capital to another , as long as the planner has sufficient time , the most efficient path to quickly move the level of capital stock to a level close to the infinite optimal path , and to allow capital to develop along that path until it is nearly the end of the desired final level .The name of the theory refers of the idea that a turnpike is the fastest route between two points which are far apart , even if it is not most direct route .

## 4.2 Origins[1]

A turnpike theorem was first proposed, at least in a way that came to wide attention, by Dorfman, Samuelson, and Solow in their famous Chapter 12 of Linear Programming and Economic Analysis , entitled "Efficient Programs of Capital Accumulation." This was in the context of a von Neumann model in which labor is treated as an intermediate product. I would like to quote the critical passage: "Thus in this unexpected way, we have found a real normative significance for steady growth-not steady growth in general, but maximal von Neumann growth. It is, in a sense, the single most effective way for the system to grow, so that if we are planning long-run growth, no matter where we start, and where we desire to end up, it will pay in the inter-mediate stages to get into a growth phase of this kind. It is exactly like a turnpike paralleled by a network of minor roads. There is a fastest route between any two points; and if the origin and destination are close together and far from the turnpike, the best route may not touch the turnpike. But if origin and destination are far enough apart, it will always pay to get on to the turnpike and cover distance at the best rate of travel, even if this means adding a little mileage at either end. The best intermediate capital configuration is one which will grow most rapidly, even if it is not the desired one, it is temporarily optima . It is due to this reference, I believe, that theorems on asymptotic properties of efficient, or optimal., paths of capital accumulation came to be known as "turnpike theorems." For a long time the theory continued to be developed for the von Neumann model, in the strict sense of a model in which consumption appears as a necessary input to processes of production. In order to discuss efficient accumulation an objective must be introduced and in such a model the natural objective is to maximize the level of terminal stocks in some sense. The objective chosen by Dorfman, Samuelson, and Solow was to maximize the distance from the origin of the terminal stocks along a prescribed ray. The original proofs they used were only valid in a neighborhood of the turnpike. Also their arguments were incomplete and a slip occurred at one place.

## 4.3 A history of Turnpikes[85]

The turnpike theory is originated in two famous papers , one is the paper by John von Numann (1937) titled “A model of General Economic Equilibrium ” ,and the other is the one by Frank Ramsay (1928) titled “ A mathematical theory of Saving ” . Later the turnpike property has been demonstrated in either of these models. Mckenzie referred to them as the “ Samuelson turnpike ” and the “Ramsay turnpike ” respectively . We will explain both turnpike respectively . The reader who wants to quickly learn a brief history of turnpike theory should consult an excellent article by Mckenzie (1986).

## 4.4 Kinds of Turnpikes[1]

The first turnpike theorem due to Dorfman, Samuelson, and Solow , was concerned with a finite accumulation path that swung toward an efficient balanced path in the middle phase of its history. There is an assigned terminal capital stock and the objective is to maximize the sum of utility over the finite accumulation period. Then we show that if the accumulation period is long enough the optimal path will stay most of the time within an assigned small neighborhood of an infinite path that is optimal (using the term "optimal" vaguely at present).This kind of turnpike is illustrated in Figure 4.1, where the infinite path is balanced . It should be mentioned that the use of a balanced path as the turnpike is incidental to the stationarity of the model. The real ground for the result is the tendency for finite optimal paths to bunch together in the middle time, and this tendency is preserved even in models which are time-dependent. The second kind of turnpike theorem also concerns finite optimal paths but it compares them with an infinite path that is price-supported and starts from the same initial stocks. It asserts that a sufficiently long finite path will hug the infinite optimal path in its initial phase whatever terminal stocks are assigned . The second kind of turnpike is illustrated in Figure 4.2 The third kind of turnpike deals with infinite paths that are optimal. It is the basic result that optimal paths converge to each other in appropriate circumstances. However, in stationary models it is convenient to describe this situation as convergence of infinite optimal paths to the optimal balanced path. The critical property of optimal balanced paths in these models is that they can be supported by prices. This fact may be used to prove that infinite optimal paths exist from any initial stocks .The third kind of turnpike is illustrated in Figure 4.3. It is worthwhile describing the practical utility of the three kinds of turnpike theorem. If the initial steps of a finite program of length $T$ that is optimal must lie near the initial steps of the infinite optimal program from the same starting point, even though the target capital stock in period T ranges over a wide set of possibilities, it will not be necessary to know much about tastes and technology in periods beyond T in order to approximate an optimal program in the first period. Our models have a Markov property. The significance of facts beyond period T is fully allowed for in the choice of capital stocks for that period. To the degree that T period stocks can vary without substantial effect on choices in the first period, knowledge of tastes and technology beyond T is not needed. On the other hand, if the capital stocks of finite optimal programs of length T must lie near together in period $\tau < T$ for widely differing initial and terminal stocks, it becomes possible to plan for an infinite program that is approximately optimal by aiming at the stock of period $\tau$ for whatever program of the set is easiest to compute. Once more, it is not necessary to know tastes and technology beyond T and, in addition, planning can be concentrated on the first $\tau$ periods. This assumes, of course, that the T period stock of the infinite optimal program belongs to the set of terminal stocks for which the theorem holds, and that an infinite optimal program exists. Finally, the convergence to one another of the infinite optimal paths from different initial stocks means that infinite optimal paths may be approximated by computing finite optimal paths with the stock of any (within limits) optimal path in some period T as the target. This is useful if the infinite optimal path from a particular initial stock is easy to compute.

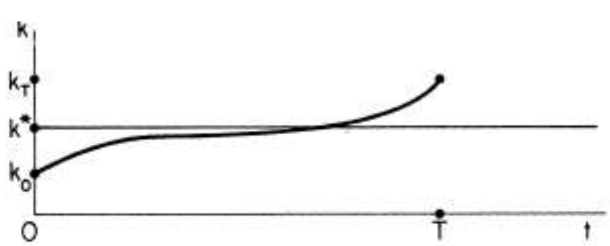


**Figure 4.1**:The middle turnpike


**Figure 4.2**:The early turnpike

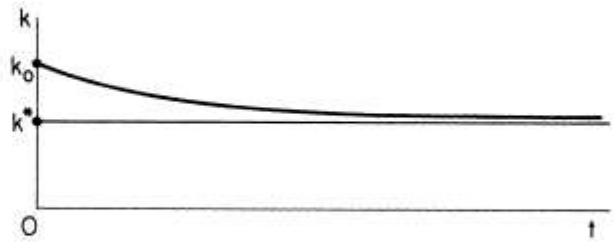



**Figure 4.3**:The late turnpike

## 4.5 A Turnpike in the von Neumann Model[84]

The von Neumann model may be defined by an input matrix $A = [a_{ij}]$ and an output matrix $B = [b_{ij}]$ . The term $a_{ij} \geq 0$ represents the input of the $i$th good needed at a unite level of the $j$th activity , and $b_{ij} \geq 0$ represents the output of the $i$th good achieved at a unit level of the $j$th activity . There are $n$ goods and $m$ activities so $A$ and $B$ are $n \times m$. Inputs occur at the start of a production period, which is uniform for all activities, and outputs appear at the end of the period. Goods at different levels of depreciation are treated as different goods but the number of goods may be as large as needed to achieve an adequate approximation

to reality. An equilibrium of the von Neumann model is defined by price vector $p \geq 0$ , a vector of activity levels $x \geq 0$ , and a rate of expansion $a > 0$ which satisfy the relations (1) $Bx \geq \alpha Ax$ , (2) $pB \leq \alpha pA$ , and (3) $pBx > 0$. Relation (1) provides that output is adequate to supply next period's input requirements . Relation (2) implies that no activity is profitable . Relation (3) implies that some good that is produced has a positive price . $x$ and $p$ may be chosen to satisfy $\sum_1^m x_j = 1$ and $\sum_1^n p_t = 1$ . If (1) is multiplied on the left by $p$ and (2) on the right by $x$ , we find that $pBx = \alpha pAx > 0$. Therefore , some activities are used and they earn zero profits. Assume the conditions (1) $a_{ij} > 0$ for some $i$ and any $j$ , (2) $b_{ij} > 0$ for some $j$ for any $i$ and (3) if $\alpha'$ is maximum value of a such that $Bx \geq Ax$ holds for some $x \geq 0$ ,then $Bx > 0$ . With these conditions (essentially) von Neumann proved that the model has a unique equilibrium, after normalizing x and p, and that the equilibrium value of a is the maximal rate of proportional expansion. The turnpike name was applied to an asymptotic result for the von Neumann model . They consider paths of accumulation starting from given initial stocks which maximize the size of terminal stocks at the end of the period of accumulation where the proportions of goods in the terminal stocks is specified in advance. They show that for sufficiently long paths which are maximal in this sense the configuration of stocks will be within an arbitrary neighbourhood of the von Neumann equilibrium for all but an arbitrary fraction of the time. This theorem gives the von Neumann equilibrium, which is called 'the turnpike', a general significance for efficient accumulation. An efficient path may be supported by prices just as the equilibrium path is finite list of processes. (4) if $(x, p, \alpha)$ is von Neumann equilibrium and $x^1$ is any other vector levels , $p(B - \alpha A)x^1 < 0$ . With this assumption Radner proved a 'value loss ' lemma which may be stated in this way , for any $\varepsilon > 0$ there is $\delta < 1$ such that $pBx^1 \leq (\alpha pAx^1)$ ,if $x^1$ and $|x^1/|x^1| - x/|x|| > \varepsilon$ .With this lemma it is easy to prove a turnpike theorem. Let a sequence of capital stock vectors, $(y_0 , y_1, \dots , y_T)$ be a path if there is a corresponding sequence of activity vectors $(x_0 , x_1, \dots , x_T)$ such that $y_t = Ax_{t+1}$ for $t = 0, \dots , T - 1$ ,and $y_t \geq Bx_t$ for $t = 0, \dots , T$. Assume that the vector $y_0$ of initial stocks satisfies $y_0 > 0$ . Then by disposal $y < y_0$ may be chosen so that

$y = Ax$ and $(x, p, \alpha)$ is a von Neumann equilibrium . then $(y, \alpha y, \dots, \alpha^T y)$ is a feasible path . Suppose $(y_0, \dots, y_T)$ is maximal path . then the value loss lemma implies that for any $\varepsilon > 0$ there is $\delta < 1$such that $\delta\alpha p y_t \geq p y_{t+1}$ where $|x_t/|x_t| - x/|x|| > \varepsilon$ . But the equilibrium conditions imply $\alpha p y_t \geq p y_{t+1}$. Thus if $x_t/|x_t|$ is outside the $\varepsilon$-neighbourhood of $x/|x|$ for $\tau$ periods then for $T > \tau$ it will be true that

$$\delta^T \alpha^T p y_0 \geq p y_T \qquad (4.1)$$

Let the desired configuration of terminal stocks be given by the vector y. Define a utility function on terminal stocks by $\rho(z) = \min z(i)/\bar{y}(i)$ over $i = 1, \dots, n$ .Then $(y_0, \dots, y_T)$ maximal implies

$$\rho(y_t) \geq \rho(\alpha^T y) = \alpha^T \rho(y) \qquad (4.2)$$

Since $y > 0, \rho(y) > 0$ . Now choose the length of the equilibrium price vector so that $p(i) \geq 1/\bar{y}(i)$ for some $i$ with $p(i) > 0$ . this implies that

$$pz \geq z(i)/\bar{y}(i) \geq \rho(z) \qquad (4.3)$$

Combining (4.1) ,(4.2) and (4.3) , gives the inequalities

$$\alpha^T \rho(y) \leq \rho(y_T) \leq p y_T \leq \delta^T \alpha^T p y_0 \qquad (4.4)$$

The first and last terms of (4) imply that $\delta^\tau$cannot exceed $\rho(y)/p y_0$which is a well defined positive number. Thus (4) implies that an integer $\bar{\tau}$ exists such that $x_t/|x_t|$ cannot lie outside the $\varepsilon$-neighbourhood of $x/|x|$ for more than $\bar{\tau}$ periods regardless of the length $T$ of the accumulation path. Since $y$ and $y_t$ are linear transforms of $x$ and $x_{t+1}$, an analogous statement holds for $y$and $y_t$ . This is a stronger form of the conclusion of the original Dosso theorem.

## 4.6 Relevant Results in Continuous Time [54]

Consider the following problem subsequently referred to as problem $\mathcal{P}_c$ .

$$\max\ J_T\left(x_0, u(\cdot)\right) = \int_0^T f_0\left(x(t), u(t)\right)dt$$

Subject to $\dot{x}(t) = f\left(x(t), u(t)\right)$ and $x(0) = x_0$

Where $f(\cdot,\cdot)\colon \mathbb{R}^n \times R^m \to \mathbb{R}^n$ is continuous mapping with respect to both arguments $x(t)$ and $u(t)$ , $U(x) \subset \mathbb{R}^m$ is compact , and that the mapping $x \to U(x)$ is upper semicontinuous .

We now turn to optimality criteria for infinite –horizon problems

**Definition 4.1**

A trajectory $x^*(\cdot)$ from $x_0$ is said to be maximal if for any other trajectory from $x_0$

$$\liminf_{T\to\infty} \int_{\mathrm{t}=0}^{\mathrm{T}} \mathrm{J_T}\left(x_0 - u(\cdot)\right) - J_T\left(x_0, u^*(\cdot)\right) \leq 0$$

A program is optimal if the lim sup operator above is substituted for lim inf .

For the problem $\mathcal{P}_\mathrm{c}$ , let $(\hat{x}, \hat{u})$ be the solution to the following associated static problem:

$$\max f_0(x, u) \text{ such that } f(x, u) = 0 \text{ and } u \in U(x). \tag{4.5}$$

The infinite –horizon problem we wish to investigate in general ,as an extension of the finite - horizon problem

$$\max \int_0^T L_0\left(x(t), \dot{x}(t)\right)dt \text{ subject to } x(0) = x_0, x(T) = x_T$$

Has the from

$$\max \int_0^\infty L_0\left(x(t), \dot{x}(t)\right)e^{-\rho}dt \text{ subject to } x(0) = x_0.$$

## 4.7 The Basic in Discrete Time[54]

We now turn to discrete-time setting, and consider the corresponding version of the Problem $\mathcal{P}_c$ We shall refer to it as the Problem $\mathcal{P}_d$ for emphasis.

$$V_T(x_0, \rho) = \max \sum_0^\infty \rho^t u(x_t, x_{t+1})$$

And where $u: \Omega \subset \mathbb{R}_+^{2n} \to \mathbb{R}$ and $0 < \rho < 1$.We shall also consider the case where $\rho = 1$ and refer to it as the undiscounted version the of the problem $\mathcal{P}_d$.

**Definition 4.2**

A program from xo is a sequence $\{x(t)\}$ such that $x(0) = x_0$, and for all $t \in \mathbb{N}$,$(x(t), x(t+1)) \in \Omega$ A program $\{x(t)\}$ is simply a program from .

A program $\{x(t)\}$ is called stationary if for all $t \in \mathbb{N}$, $x(t) = x(t+1)$ .

We now turn to optimality criteria for a program. In the discounted case, with assumptions on $\Omega$that ensure boundedness of programs, there is no issue of convergence of the performance functional.

**Definition4.3**

For all $0 < \rho < 1$ , a program $\{x^*(t)\}$ from $x_0$ is said to be optimal if

$$\sum_{t=0}^{\infty} \rho^t \big[u\big(x(t), x(t+1)\big) - u(x^*(t), x^*(t+1))\big] \leq 0$$

For every program $\{x(t)\}$ from $x_0$ . A stationary optimal program is a program that is stationary and optimal .

**Definition4.4**

A program $\{x^*(t)\}$ from $x_0$ is said to be maximal if for any other program from $\{x(t)\}$ from $x_0$

$$\lim_{T\to\infty} inf \sum_{t=0}^{T} u\big(x(t), x(t+1)\big) - u(x^*(t), x^*(t+1) \leq 0$$

A program is optimal if the lim sup operator above for lim inf .

We consider the discrete –time analogue of static problem considered above as (4.5)

$$\max u(x, x') \text{ such that } x' \geq x \text{ and } (x, x') \in \Omega . \qquad (4.6)$$

## 4.8 Support Prices[1]

The turnpike theorems will be proved by use of the support prices . When feasible path are infinite , the utility sum $\sum_t^T u_\tau\,(k_{\tau-1}, k_\tau)$ may diverge as $T \to \infty$. Let $\{k_t'\}$ ,$t = 0,1, ...$,be an infinite path from $k_0 = \bar{x}$. Then $(k_t\,, k_{t+1}) \in D_{t+1}$ for all $t$ .If $\{k_t'\}$ is the second path from $\bar{x}$ at time 0 , let us say that $\{k_t\}$ catches up to $\{k_t'\}$ if

$$\text{lim sup} \sum_{1}^{T} (u_t(k_{t-1}'\ , k_t') - u_t(k_{t-1}, k_t) \leqq 0$$

As $T \to \infty$ .On the other hand , let us say that $\{k_t'\}$ overtakes $\{k_t\}$ if

$$\text{lim inf} \sum_{1}^{T} (u_t(k_{t-1}'\ , k_t') - u_t(k_{t-1}, k_t) \geqq \varepsilon$$

for some $\varepsilon > 0$ as $T \to \infty$. An optimal path is a feasible path that catches up to every other feasible path from the same initial stocks. A weakly maximal path is a feasible path which is not overtaken by any other feasible path from the same initial stocks.

**Assumption 4.1**

For any given $t$ and $\xi < \infty$ , there is $\xi < \infty$ such that $|x| < \xi$ implies $u_t(x, y) < \xi$ and $|y| < \xi$.

The derivation of support prices requires the following assumption:

**Assumption 4.2**

The utility function $u_t(x, y)$ are concave and closed for all $t$ .the set $D_t$ is convex .

Assumption 4.2 implies that $V_t(x)$ is a concave function. Let $P_t$ be the set of $y$ such that $(x, y) \in D_t$ for some $x$. $P_t$ is convex from the convexity of $D_t$. Let $F_t$ be the smallest flat in $E_t$ that contains $P_t$ and $K_t$. Given initial stocks $\bar{x}$, we make the following assumption:

**Assumption 4.3**

$\bar{x} \in$ relative interior $K_0$ and : for $t \geqq 1$ ,interior $P_t \cap K_t \neq \emptyset$ relative to $F_t$ .

**Lemma4.1**

Let $\{k_t\}$ , $t = 0,1, \dots, K_0 = \bar{x}$ , be a weakly maximal path of accumulation . If Assumption 1,2 , and 3 are met , there exists a sequence of price vectors $P_t \in F_t$ , $P_t . F_t \neq 0, t = 0,1, \dots$,which satisfy

$$V_t(k_t) - P_t k_t \geqq V_t(y) - P_t y \text{ for all } y \in k_t \quad (4.7)$$

$$u_{t+1}(k_t, k_{t+1}) + P_{t+1} k_{t+1} - P_t k_t \geqq u_{t+1}(x, y) + P_{t+1} y - P_t x \quad (4.8)$$

For all $(x, y) \in D_{t+1}$ .

## 4.9 An Insignificant Future[1]

We assume the following

**Assumption 4.4**

The utility function $u_{t+1}(x, y)$ is strictly concave and closed .

First we define the notion of value loss .Given $(x, y) \in D_t$ , let $(p, q)$ satisfy

$$u_t(x, y) + qy - px = u_t(z, w) + qw - pz + \delta_t(z, w) \quad (4.9\ )$$

For any $(z, w) \in D_t$ , where $\delta_t(z, w) \geqq 0$ .Thus $(p, q)$ are support prices for $(x, y)$ in the $t$th period . Then $\delta_t(z, w)$ is the value loss associated with $(z, w)$ .

**Lemma 4.2**

If $u_t$ satisfies Assumption 4.4 , give $(x, y) \in D_t$ and $(p, q)$ satisfies (4.9) , there is $\delta > 0$ such that $|z - x| > \varepsilon$ implies $\delta_t(z, w) > \delta$ for any $(z, w) \in D_t$ .

**Theorem4.1**

Let $\{k_t\}$ be weakly maximal path . Suppose Assumptions 4.1, 4.3,and 4.4 are met .Let $\bar{y}$ be a stock vector that is freely reachable from $\{k_t\}$ and from which $\{k_t\}$ is freely reachable .for any $\varepsilon > 0$ and any $\tau_1$ there is $\tau_2$ such that if $\{k_t'\}$ , $t = 0, \dots, T, T \geqq \tau_2$ , $\bar{k}_0 = k_0$ , , $\bar{k}_T = \bar{y}$ . is an optimal path , then $|\bar{k}_t - k_t| < \varepsilon$ for $t \leqq \tau_1$ .

An example of a utility function that satisfies the condition of theorem 4.1 is a stationary current utility function that is strictly concave where the objective is to maximize a utility sum discounted at a positive rate , that is ,

$$\sum_{1}^{T} u_t(k_{t-1}, k_t) = \sum_{1}^{T} \rho^t u(k_{t-1}, k_t)$$

For $0 < \rho < 1$.

**Proposition4.1**

Let $\{k_t\}$ be a weakly maximal path and let Assumptions 4.1,4.3,and 4.4 hold . Suppose $\{k_t\}$ is freely reachable from any path from $k_0$ that it does not overtake . Then $\{k_t\}$ is optimal.

**Assumption4. 5**

The utility function $\{u_t\}$ are uniformly concave , that is ,the $\delta$ of lemma 4. 2 may be chosen independent of $(x, y)$ and $t$ .

## 4.10 Scheinkman's "Neighborhood'' Stability Theorem[54]

**Assumption4.6**

There exist $M > 0$ and $N < 1$ such that for all $(x, x') \in \Omega$ ,$\|x\| < M$ implies $\|x'\| < N\|x\|$ .

**Assumption 4. 7**

If $(x, x') \in \Omega$ then $(y, y') \in \Omega$ for all $y \geq x$, $0 \leq y' \leq x'$and $u(y, y') \geq u(x, x')$ holds.

**Assumption4.8**

There exists $(\bar{x}, \bar{x}') \in \Omega$ such that $\bar{x}' >> \bar{x}$ . In this case , $\bar{x}$ is said to be expansible.

**Assumption 4. 9**

The golden –rule stock $\hat{x}$ is unique and expansible.

**Assumption4. 10**

(i) $u(\cdot,\cdot)$ is $C^3$in lnt $\Omega$ and concave . (ii) The matrix

$$\begin{bmatrix} u_{xx}(\hat{x}_\rho, \hat{x}_\rho) & u_{xx'}(\hat{x}_\rho, \hat{x}_\rho) \\ u_{x'x}(\hat{x}_\rho, \hat{x}_\rho) & u_{x'x'}(\hat{x}_\rho, \hat{x}_\rho) \end{bmatrix} = \begin{bmatrix} A_\rho & B_\rho \\ B'_\rho & C_\rho \end{bmatrix}$$

Is negative at $(\hat{x}, \hat{x})$ . (iii) All programs are interior programs so that the Euler difference equation system

$$u_{x'}(t-1), x(t) + \rho u_x(x(t+1) = 0 \text{ for all } t \geq 1 \qquad (4.10)$$

Is satisfied for an optimal program (iv) The characteristic equation of the Euler difference equation system (4.10) does not have a zero root . (v) there exist $\rho_\ell < 1$

Such that for all $\rho \leq \rho_\ell$ , there exists an expansible stationary program $\underline{x}$ such that $\hat{x}^\rho >> \underline{x}$

**Theorem4.2 (Scheinkman).**

Under Assumption 4.6,4.7 ,4.8 ,4.9 and 4.10. For any expansible $x$ ,there exists $\bar{\rho} > 1$ such that $1 \geq \rho \geq \bar{\rho}$ and $x_0 \geq x$ implies that there exists an optimal program $x(t, x_0, \rho)$such that $\lim_{t\to\infty} x(t, x_0, \rho) = \hat{x}^\rho$ where $\hat{x}^\rho$ is the unique modified golden-rule stock associated with $\rho$.

## 4.11 The Theory in Non-Smooth Environments[54]

In the chapter on Competitive Equilibriumover Time, McKenzie presents results for a "generalized Ramsey model", and his work is distinguished by its reliance on a bounded assumption on the technology **Ω**, and a joint assumption on the pair $(u, \Omega)$ guaranteeing free-disposal and monotonicity. Under Assumptions 4.6 and 4.7, he presents theorems for both the discounted and undiscounted cases, seeing the latter as logically antecedent to the former.

### 4.11.1 Theorems Under Refined Interiority Assumptions

We can now present a classical turnpike theorem, Card(A) denoting the cardinality of a finite set A.

**Theorem 4.3**

Let Assumption 4.1 ,4,2 and 4,3 hold and that $\hat{x}$ is expansible. Then ,there exists $\varepsilon > 0$ such that for all $T > L$, and all optimal programs $\{x_t\}_{t=0}^{t=T}$ such that $x_0$and $x_T$ are expansible ,

$$\text{Gard}\{i = [0, \dots, T-1]: \|x(t) - (\hat{x}, \hat{p})\| > \varepsilon\} \leq L$$

Mckenzie introduces his other theorem for the undiscounted Ramsey
setting as “an asymptotic theorem for infinite optimal paths.” For this, he needs a particular set of initial stocks$K$ from which there start programs $\{x(t)\}$such that

$$\liminf_{T\to\infty} \sum_{t=0}^{T} u\big(x(t), x(t+1) - u(x^*(t), x^*(t+1)\big) \geq -\infty.$$

### 4.11.2 Mckenzie’s “Neighborhood Turnpike’’ Theorem

We shall now assumption on the discount factors . We shall need the following additional notation .Let $D = \{x \in \mathbb{R}_+^n : (x, x) \in \Omega\}$ and

$$D_\rho = \{x \in D: u(x, x) \geq u(\bar{x}, \bar{x}') \in \Omega \text{ such that } \rho\bar{x}' >> x\}$$

**Definition 4.5**

The utility function u is said to be uniformly strictly concave over $D_\rho$ if for any $bx \in D_\rho$ , and any $(w, z) \in \Omega$, the following holds: For any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|(x, x) - (w, z)| > \varepsilon \Longrightarrow u\left(\frac{1}{2}(x, x) + \frac{1}{2}(w, z)\right) - \frac{1}{2}(u(x, x) + u(w, z)) > \delta$$

**Assumption 4.11**

There exists $\rho_\ell < 1$ and $(\bar{x}, \bar{x}') \in \Omega$ such that $\rho_\ell \bar{x}' >> \bar{x}$ . We assume
$\rho_\ell \leq \rho \leq 1$, $u$ is uniformly strictly concave over $D_\rho$ and that $D_\rho$ is in the relative interior of $D$ .

**Assumption 4.12**

$\hat{x}$ is expansible which is to say that there exists $(\hat{x}, x') \in \Omega$ such that $x' >> \hat{x}$ .Furthermore $D_\rho$ is in the relation interior of $D$ .

**Definition 4.6**

We shall say that an initial stock $x_0 \in \mathbb{R}_+^n$ is sufficient if there exists a finite program $\{x_t\}_{t=0}^{t=T}$ such that $x_T$ is expansible .

We can now present McKenzie's so-called "neighborhood turnpike theorem" for any expansible $\hat{x}$.

**Theorem4.4**

Under Assumptions 4.6 and 4.7, for any sufficient $x_0 \in \mathbb{R}^n$, and any $\rho$ satisfying Assumptions 4.11 and 4.12, and any $\varepsilon > 0$, there exists

$1 > \bar{\rho} > 0$ such that $\hat{x}^\rho$ is asymptotically stable at $(\rho, x_0)$ for all $\bar{\rho} \leq \rho \leq 1$.

## 4.12 Particular Undiscounted Cases: the RSS Model

### 4.12.1 A Theorem on Asymptotic Stability

Assume the (i) $w$ is strictly concave ,or that (ii) $\xi_\sigma \neq 1$. Let $M_0, \epsilon > 0$. Then there exists $a$ natural number $T_0$such that for each optimal program

$\{x(t), y(t)\}_{t=0}^{\infty}$satisfying $x(0) \leq M_0 e$ each integer $t \geq T_0$

$$\|x(t) - \hat{x}\|, \|y(t) - \hat{x}\| \leq \epsilon.$$

### 4.12.2 Classical Turnpike Theorems

The basic results are as follows.

**Theorem 4.5**

Let $M, \epsilon$ be positive numbers and $\Gamma \in (0,1)$. Then there exists a natural number $L$ such that for each integer $T > L$ , each $z_0, z_1 \in R_+^n$ satisfying

$z_0 \leq Me$ and $az_1 \leq \Gamma d^{-1}$ and each program $\{x(t)\}_{t=0}^{T}, \{y(t)\}_{t=0}^{T-1})$ which satisfies $x(0) = z_0$ , $x(T) \geq z_1$, $\sum_{t=0}^{T-1} w(by(t) \geq U(z_0, z_1, 0, T) - M$

The following inequality holds :

$$\text{Gard}\{i \in 0, \dots, T-1\}: \max\{\|x(t) - \hat{x}\|, \|y(t) - \hat{x}\| > \epsilon\} \leq L$$

**Theorem 4.6**

Let $M, \epsilon$ be positive numbers and $\Gamma \in (0,1)$. Then there exists a natural number $L$ and a positive number $\gamma$ such that for each integer $T > 2L$, each $z_0, z_1 \in R_+^n$

satisfying $z_0 \leq Me$ and $az_1 \leq \Gamma d^{-1}$ and each program $\{x(t)\}_{t=0}^{T}, \{y(t)\}_{t=0}^{T-1})$ which satisfies

$$x(0) = z_0 \, , x(T) \geq z_1 \, , \ \sum_{t=0}^{T-1} w(by(t) \geq U(\, z_0, z_1, 0, T) - \gamma$$

There are integer $\tau_1, \tau_2$ such that $\tau_1 \in [0, L]$ , $\tau_2 \in [T-1, T]$,

$$\|x(t) - \hat{x}\|, \|y(t) - \hat{x}\| \leq \epsilon \text{ for all } t = \tau_1, \dots, \tau_2 - 1$$

and $\|x(t) - \hat{x}\| \leq \epsilon$, $\|y(t) - \hat{x}\| \leq \gamma$ then $\tau_1 = 0$.

## 4.13 Turnpike Properties[30]

A□" Turnpike " is not necessarily a singleton or a half-ray. It can be an absolutely continuous time dependent function (trajectory) or a compact subset of $\mathbb{R}^n$ . To establish a turnpike property we consider a space of cost functions equipped with a natural complete metric and show that a turnpike property holds for most elements of this space in the sense of Baire categories . We obtain a turnpike theorem in the following way . We consider an optimality criterion (a cost function $f$ )and show that for a problem with this criterion there exists an optimal trajectory, say $X_f$ , on an infinite time interval. Then we perturb our cost function by some nonnegative small perturbation which is zero only on $X_f$ . We show that for our new cost function $f$ the trajectory $X_f$ is a turnpike, and that optimal solutions of the problem with a cost function $g$ which is closed to $\bar{f}$ , are also most of the time close to $X_f$ .

### 4.13.1 The Turnpike Phenomenon[3]

Let $|\cdot|$ be the Euclidean norm in the $n$-dimensional Euclidean space $\mathbb{R}^n$ and let $< . , . >$ be the scalar product in $\mathbb{R}^n$ . We consider the variational problem

$$\int_0^T f(v(t), v'(\text{t}))dt \rightarrow \text{min} \qquad (P_0)$$

$$v: [0, T] \to \mathbb{R}^n \text{ is an absolutely continuous function}$$

$$\text{Such that } v(0) = y, v(T) = z$$

Here $T$ is a real number , $y$ and $z$ are points of the space $\mathbb{R}^n$ and an integrand $f: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^1$ is strictly convex and differentiable function such that

$$f(y,z)/(|y| + |z|) \to \infty \text{ as } |y| + |z| \to \infty.$$

In order to meet our goal let us consider the following auxiliary minimization problem:

$$f(y,0) \to \min , y \in \mathbb{R}^n \qquad (P_1)$$

By the strict convexity of $f$ and the growth condition , the problem $(P_1)$ has a unique solution $\bar{y}$ . It is easy to see that

$$\partial f/\partial y(\bar{y}, 0) = 0.$$

Define an integrand $L: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^1$ by

$$L(y,z) = f(y,z) - f(\bar{y},0) - < \nabla f(\bar{y},0), (y,z) - (\bar{y},0) >$$

$$= f(y,z) - f(\bar{y},0) - < (\partial f/\partial z)(\bar{y},0), z > .$$

Clearly $L$ is also differential and strictly convex and satisfies the same growth condition as $f$:

$$L(y,z)/(|y| + |z|) \to \infty \text{ as } |y| + |z| \to \infty.$$

Since the functions $f$ and $L$ are strictly convex we obtain that

$$L(y,z) \geq 0 \text{ for all } (y,z) \in \mathbb{R}^n \times \mathbb{R}^n$$

And

$$L(y,z) = 0 \text{ if and only if } y = \bar{y}\, , z = 0.$$

Consider an auxiliary variational problem

$$\int_0^T f(v(t), v'(\mathrm{t}))dt \to \min \qquad (P_2)$$

$$v: [0,T] \to \mathbb{R}^n \text{ is an absolutely continuous function}$$

$$\text{Such that } v(0) = y\, , v(T) = z\, ,$$

Where $T > 0$ and $y\, , z \in \mathbb{R}^n$. It is easy to see that for any absolutely continuous function $x: [0,T] \to \mathbb{R}^n$. with $T > 0$

$$\int_0^T L\big(x(t), x'(t)\big)\, dt$$

$$= \int_0^T \big[f\big(x(t), x'(t)\big) - f(\bar{y}, 0) - < (\partial f/\partial z)(\bar{y}, 0), x'(t) >\big]\, dt$$

$$= \int_0^T f\big(x(t), x'(t)\big) + Tf(\bar{y}, 0) - < (\partial f/\partial z)(\bar{y})\, , x(T) - x(0) >.$$

These equations imply that the problems $(P_0)$ and $(P_2)$ are equivalent :

A function $x: [0,T] \to \mathbb{R}^n$ is solution of the problem $(P_0)$ if and only if it is a solution of the problem $(P_2)$ .

The integrand $L: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^1$ has the following property:

**(C)** If $\{(y_i\, , z_i)\}_{i=1}^{\infty} \subset \mathbb{R}^n \times \mathbb{R}^n$ satisfies $\lim_{i\to\infty} L(y_i\, , z_i) = 0$,then $\lim_{i=\infty} y_i = \bar{y}$ and $\lim_{i=\infty} z_i = 0$.

Then

$$\lim_{i\to\infty} L(y_i\, , z_i) = (\bar{y}, 0).$$

Assume that a sequence $\{(y_i\,, z_i)\}_{i=1}^{\infty} \subset \mathbb{R}^n \times \mathbb{R}^n$ satisfies $\lim_{i\to\infty} L(y_i\,, z_i) = 0$ : the growth condition implies that the sequence $\{(y_i\,, z_i)\}_{i=1}^{\infty}$ is bounded . Let $(y\,, z)$ be its limit point . Then,

$$L(y, z) = \lim_{i\to\infty} L(y_i\,, z_i) = 0 \text{ and } (y, z) = (\bar{y}, 0)\ .$$

This implies that $(\bar{y}, 0) = \lim_{i\to\infty} L(y_i\,, z_i)$ , as claimed.

Assume that $y, z$ are points of the space $\mathbb{R}^n$ , $T < 2$ is real number and that an absolutely continuous function $\bar{x}: [0, T] \to \mathbb{R}^n$ is an optimal solution of the problem $(P_0)$ . Since the problem $(P_0)$and $(P_2)$ are equivalent the function $\bar{x}$ is also an optimal solution of the problem $(P_2)$ .

We claim that

$$\int_0^T L(\bar{x}(t), \bar{x}'(t))dt \leq 2c_0(|y|, |z|)$$

Where a positive constants $2c_0(|y|, |z|)$ depends only on $|y|$ and $|z|$.

Consider an absolutely continuous function $\bar{x}: [0, T] \to \mathbb{R}^n$ defined by

$$x(t) = y + t(\bar{y} - y), t \in [0,1]\ , x(t) = \bar{y}\ \ , t \in [1, T-1]\ ,$$

$$x(t) = \bar{y} + (t - (T-1)(z - \bar{y}), t \in [T-1\ , T].$$

By the definitions of the functions $\bar{x}$and $x$ , we have

$$\int_0^T L(\bar{x}(t), \bar{x}'(t))dt \leq \int_0^T L(x(t), x'(t))dt$$

$$= \int_0^T L(x(t), \bar{y} - y)dt + \int_1^{T-1} L(\bar{y}, 0)dt + \int_{T-1}^T L(x(t), z - \bar{y})dt$$

$$= \int_0^1 L(x(t), \bar{y} - y)dt + \int_{T-1}^T L(x(t), z - \bar{y})dt\ .$$

Clearly ,the integrals

$$\int_0^1 L(x(t), \bar{y} - y)dt \text{ and } \int_{T-1}^T L(x(t), z - \bar{y})dt$$

Do not exceed a positive constant $c_0(|y|, |z|)$ which depends only on the norms $|y|, |z|$ and does not depend on $T$. Therefore

$$\int_0^T L(\bar{x}(t), \bar{x}'(t))dt \leq 2c_0(|y|, |z|).$$

We denote by $\text{mes}(E)$ the Lebesgue measure of a Lebesgue measurable set $E \subset R^1$

Now let $\epsilon$ be positive number . By the property (C) there is $\delta > 0$ such that if $(y, z) \in \mathbb{R}^n \times \mathbb{R}^n$ and $L(y\ z) \leq \delta$ then $|y - \bar{y}| + |z| \leq \epsilon$ . in view of the choice of $\delta$ and inequality $\int_0^T L(\bar{x}(t), \bar{x}'(t))dt \leq 2c_0(|y|, |z|)$, we have

$$\text{mes}\{t \in [0, T]: |\bar{x}(t), \bar{x}'(t)) - (\bar{y}, 0)| > \epsilon\}$$

$$\leq \text{mes}\ \{t \in [0, T]: L(\bar{x}(t), \bar{x}'(t)) > \delta\}$$

$$\leq \delta^{-1} \int_0^T L(\bar{x}(t), \bar{x}'(t))dt \leq \delta^{-1} 2c_0(|y|, |z|)$$

And

$$\text{mes}\{t \in [0, T]: |\bar{x}(t) - \bar{y}| > \epsilon)\} \leq \delta^{-1} 2c_0(|y|, |z|\ .$$

Therefore the optimal solution $\bar{x}$ spends most of the time in an $\epsilon$-neighborhood of the point $\bar{y}$ . The Lebesgue measure of the set of all points $t$, for which $\bar{x}(t)$ does not belong to this $\epsilon$-neighborhood, does not exceed the constant $2\delta^{-1}c_0(|y|, |z|)$ which depends only on $|y|, |z|$ and $\epsilon$ and does not depend on $T$. Following the tradition, the point $\bar{y}$ is called the turnpike. Moreover we can show that the set

$$\{t \in [0, T]: |\bar{x}(t) - \bar{y}| > \epsilon\}$$

Is contained in the union of tow intervals

$$[0\,,\tau_1] \cup [T-\tau_1, T] \text{ where } 0<\tau_1\,,\tau_1 \le 2\delta^{-1}c_0(|y|,|z|).$$

## 4.13.2 Nonconvex Nonautonomous Integrands [30]

We showed in the previous section that the structure of optimal solutions of the problem $(P_0)$, under the assumptions posed on $f$, is rather simple and the turnpike is calculated easily as a solution of the problem $(P_1)$. Nevertheless, the convexity of the integrand $f$and its time independence are very essential for the proof of this turnpike result. In order to obtain a turnpike result for essentially larger classes of variational problems and optimal control problems we need other methods and ideas. The following example helps to understand what happens if the integrand $f$ is nonconvex and nonautonomous and what kind of turnpike we have for general nonconvex nonautonomous integrands.

**Example 4.1**

Let

$$f(t,x,u) = (x-\cos(t))^2 + (u+\sin(t))^2\,,(t,x,u) \in R^1 \times R^1 \times R^1$$

And consider the family of the variational problems

$$\int_{T_1}^{T_2} [(v(t)-\cos(t))^2 + (v'(t)+\sin(t))^2]\,dt \to \min \qquad (P_3)$$

$v:[T_1\,,T_2] \to R^1$ is an absolutely continuous function

Such that $v(T_1) = y\,,v(T_2) = z,$

Where $y\,,z\ ,\ T_1\ ,T_2 \in R^1$ and $T_2 < T_2$ .The integrand $f$ depends on $t$ ,for each $t \in R^1$ the function $f(t,.,.): R^2 \to R^1$ is convex ,and for each $x\,,u \in R^1/\{0\}$ the function $f(.,x,u): R^1 \to R^1$ is nonconvex . Thus the function $f: R^1 \times R^1 \times R^1 \to R^1$ is also nonconvex and depends on $t$.

Let $y\,,z\,,T_1\,,T_2$ be real number ,$T_2 >, T_1\ +2$ and let a function $\hat{v}:[T_1\,,T_2] \to R^1$ is an optimal solution of problem $(P_3)$ . Note that the problem $(P_3)$ possesses a solution since the integrand $f$is a continuous function and the function $f(t,x,.) : R^1 \times R^1$ is convex and grows superlinearly at infinity for each point $(t\,,x) \in [0\,,\infty) \times R^1$.

Consider a function $\upsilon(t) : [T_1\,,T_2] \to R^1$ defined by

$$\upsilon(t) = y + (\cos(1) - y)(t - T_1)\,, t \in [T_1\,,T_1 + 1]\,,$$

$$\upsilon(t) = cos(t)\,, t \in [T_1 + 1\,,T_2 - 1]\,,$$

$$\upsilon(t) = cos(T_2 - 1) + (t - T_2 + 1)\big(z - cos(T_2)\big)\,, t \in [T_2 - 1, T_2]\,.$$

Clearly ,

$$\int_{T_1+1}^{T_2-1} f\big(t\,,\upsilon(t),\upsilon'(t)\big)\,dt = 0$$

And

$$\int_{T_1}^{T_2} f\big(t,\hat{\upsilon}(t),\hat{\upsilon}'(t)\big)\,dt \le \int_{T_1}^{T_2} f\big(t\,,\upsilon(t),\upsilon'(t)\big)dt$$

$$= \int_{T_1}^{T_1+1} f\big(t\,,\upsilon(t),\upsilon'(t)\big)\,dt + \int_{T_2-1}^{T_2} f\big(t\,,\upsilon(t),\upsilon'(t)\big)\,dt$$

$$\le 2\sup\{|f(t,x,u)|: t,x,u \in R^1, |x|,|u| \le |y| + |z| + 1\}\,.$$

Thus

$$\int_{T_1}^{T_2} f\big(t,\hat{\upsilon}(t),\hat{\upsilon}'(t)\big)\,dt \le c_1(|y|,|z|),$$

Where

$$c_1(|y|,|z|) =\le 2\sup\{|f(t,x,u)|: t,x,u \in R^1, |x|,|u| \le |y| + |z| + 1\}.$$

For any real number $\epsilon \in (0,1)$ the following inequality holds :

$$mes\,\{t \in [T_1,T_2]: |\hat{\upsilon}(t) - \cos(t)| > \epsilon\}$$

$$\le \epsilon^{-2} \int_{T_1}^{T_2} f\big(t,\hat{\upsilon}(t),\hat{\upsilon}'(t)\big)\,dt \le \epsilon^{-2} c_1(|y|,|z|)\,.$$

Since the constant $c_1(|y|,|z|)$ does not depend on $T_2$and $T_1$ we conclude that if $T_2 - T_1$ is sufficiently large , then the optimal solution $\hat{\upsilon}(t)$ is equal to $\cos(t)$ up to $\epsilon$ for most $t \in [T_1,T_2]$ . Again , as in the case of convex time independent problems we can show that

$$\{t \in [T_1, T_2]: |x(t) - \cos(t)| > \epsilon\} \subset [T_1, T_1 + \tau] \cup [T_2 - \tau, T_2]$$

Where $\tau > 0$ is a constant which depends on $\epsilon$ , $|y|$ and $|z|$.

This example demonstrates that there exist nonconvex time dependent integrands which have the turnpike property with the same type of convergence as in the case of convex autonomous variational problems. The difference is that the turnpike is not a singleton but an absolutely continuous time dependent function defined on the infinite interval $[0\,,\infty)$.

**4.13.3 Definition of the Turnpike Property for General Integrands**[30]

Consider the problem of the calculus of variations

$$\int_{T_1}^{T_2} f\big(t\,, v(t), v'(t)\big)dt \to \min \qquad \text{(P)}$$

$v: [T_1, T_2] \to R^n$ is absolutely continuous function

Such that $v(T_1) = y\,, v(T_2) = z$.

Here $T_1 < T_2$ a real numbers, $y$ and $z$ are elements of the $n$-dimensional Euclidean space $R^n$ and an integrand $f: [0, \infty) \times R^n \times R^n \to R^n$ is a continuous function .

We say that the integrand $f$ possesses the turnpike property if there exists a locally absolutely continuous function $X_f: [0\,,\infty) \to R^n$ (called the "turnpike" ) which depends only on $f$ such that the following condition holds :

For each bounded subset $K$of the space $R^n$ and each positive number $\epsilon$ there exists a positive constant $T(K\,,\epsilon)$ such that for each pair of real numbers $T_1 \geq 0$ , each $T_2 \geq T_1 + 2T(K, \epsilon)$ ,each $y\,, z \in K$and each optimal solution $v: [T_1, T_2] \to R^n$ of variational problem (P) , the inequality $\left|v(t) - X_f\right| \leq \epsilon$ holds for all $t \in [T_1 + T(K\,,\epsilon), T_2 - T(K\,,\epsilon)]$.

The turnpike property is very important for applications. Suppose that the integrand $f$ has the turnpike property, $k$ and $\epsilon$ are given , and we know a finite number of "approximate" solutions of the problem (P). Then we know the turnpike $X_f$ , or at least its approximation, and the constant $T(K,\epsilon)$ which is an estimate for the time period required to reach the turnpike. This information can be useful if we need to find an "approximate" solution of the problem (P) with a new time interval $[T_1, T_2]$ and the new values $y, z \in K$at the end points $T_1$and $T_2$.

Namely instead of solving this new problem on the "large" interval $[T_1, T_2]$ we can find an "approximate" solution of problem (P) on the "small" interval $[T_1, T_1 + T(K,\epsilon)]$ with the values $y, X_f(T_1 + T(K,\epsilon)$ at the end points and an "approximate" solution of problem (P) on the "small" interval $[T_2 - T(K,\epsilon), T_2]$ with the values $X_f\big(T_2 - T(K,\epsilon)\big), z$ at the end points. Then the concatenation of the first solution, the function $X_f: [T_1 + T(K,\epsilon), T_2 - T(K,\epsilon)]$and the second solution is an "approximate" solution of problem (P) on the interval $[T_1, T_2]$ with the values $y$ , $z$ at the end points.

**Example 4.2**

Let

$$f(x_1, x_2, u_1, u_2) = (x_1^2 + x_2^2 - 1)^2 + (u_1 + x_2)^2 + (u_2 - x_1)^2$$
$$(x_1 , x_2, u_1, u_2) \in R^2 \times R^2$$

And consider the family of the variational problems

$$\int_0^T f(v_1\,(t), v_2(t), v'_1(t)\,, v'_1(t))dt \to \min \qquad (P_4)$$

$(v_1, v_2): [0, T] \to R^2$ is absolutely continuous function

Such that $(v_1, v_2)(0) = y$ , $(v_1, v_2)(T) = z$

Where $y = (y_1, y_2)\,, z = (z_1, z_2) \in R^2$ and $T > 0$ . The integrand $f$ does not depend on $t$ . Since $f$ is continuous and for each $x = (x_1, x_2) \in R^2$ the function $f(x\,,\,\cdot): R^2 \to R^1$ is convex and grows superlinearly at infinity , the problem $(P_4)$ Has solution for each $T > 0$ and each $y\,, z \in R^2$.

Cleary , if $T > 0$ , $y = (\cos(0)\,, \sin(0))$ and $z = (\cos(T)\,, \sin(T))$ , then the function

$$\hat{x}_1(t) = \cos(t)\ ,\ \hat{x}_2(t) = \sin(t)\ ,\ t \in [0, T]$$

Is a solution of the problem $(P_4)$ .Thus , if the integrand $f$ has a turnpike property , then the turnpike is not a singleton .

Let $T > 2$ ,$y, z \in R^2$ and let $\bar{v} = (\bar{v}_1, \bar{v}_2): [0, T] \to R^n$ by

$$v(t) = y + t((\cos(1), \sin(1)) - \mathrm{y})\ , \mathrm{t} \in [0,1]$$

$$v(t) = (\cos(t), \sin(t))\ , t \in [1, T-1]$$

$$v(t) = \cos(T-1), \sin(T-1) + (t - T + 1)(z - \cos(T-1), \sin(T-1))$$

$t \in [T-1, T]$ .

Then

$$\int_1^{T-1} f(v(t), v'(t))dt = 0$$

And

$$\int_0^T (\bar{v}_1(t)^2 + \bar{v}_2(t)^2 - 1)^2\, dt \le \int_0^T f\big(\bar{v}(t), \bar{v}'(t)\big)dt$$

$$\le \int_0^T f(v(t), v'(t))dt$$

$$= \int_0^T f(v(t), v'(t))dt + \int_{T-1}^T f(v(t), v'(t))dt$$

$$\le \sup\{f(x_1, x_2, u_1, u_2): x_1, x_2, u_1, u_2 \in R^1\ and\ |x_i|, |u_i| \le 2|y| + 2|z| + 2\ , \\ i = 1,2\}$$

Thus

$$\int_0^T (\bar{v}_1(t)^2 + \bar{v}_2(t)^2 - 1)^2 dt \le c_2\,(|y|, |z|)$$

With

$$c_2(|y|, |z|) = \sup\{f(x_1, x_2, u_1, u_2): x_1, x_2, u_1, u_2 \in R^1\ and\ |x_i|, |u_i| \\ \le 2|y| + 2|z| + 2\ \}$$

Here $c_2(|y|,|z|)$ depend only on $|y|,|z|$ and does not on $T$ . For any $\epsilon \in (0,1)$ we have

$$\text{mes}\left\{t \in [0,T]: \left||\bar{v}_1(t), \bar{v}_2(t)| - 1\right| > \epsilon\right\}$$
$$\leq \text{mes}\{t \in [0,T]: |\bar{v}_1(t)^2 + \bar{v}_2(t)^2 - 1| > \epsilon^2\}$$
$$\leq \epsilon^{-4} \int_0^T (\bar{v}_1(t)^2 + \bar{v}_2(t)^2 - 1)^2 dt$$
$$\leq \epsilon^{-4} c_2(|y|,|z|).$$

It means that for most $t \in [0,T]$ , $\bar{v}(t)$ belongs to the $\epsilon$-neighborhood of the set $\{x \in R^2: |x| = 1\}$. Thus we can say that the integrand $f$ has a weakened version of the turnpike property and the set $\{|x| = 1\}$can be considered as the turnpike for $f$.

# Chapter Five

# Turnpike Properties in Calculus of Variations and Optimal Control

## 5.1 Introduction

In this chapter we surevey our results of the turnpike property for some classes of variational and optimal control problems. To have property means the approximate solution of the problems are determined mainaly by objective functions and essentially independent of the choice of interval and endpoint conditions except in regions close to the endpoint .We discuss necessary and sufficient conditions for turnpike properties of approximate solutions for variational problems and discrete-time optimal control problems. Turnpike properties have been established long time ago in finite-dimensional optimal control problem arising in econometry.

## 5.2 Turnpike Properties for Variational Problems[83]

We consider the variational problems

$$\int_{T_1}^{T_2} f\big(t, z(t), z'(t)\big)dt \to \min , z(T_1) = x , z(T_2) = y \qquad \text{(P)}$$

$z: [T_1, T_2] \to R^n$ is an absolutely continuous function ,

Where $T_1 \geq 0, T_2 > T_1, x, y \in R^n$ and $f: [0, \infty) \times R^n \times R^n \to R^1$ belongs to space of integrands described.

Let $T_1 \geq 0, T_2 > T_1, x, y \in R^n$ and $f: [0, \infty) \times R^n \times R^n \to R^1$ be an integrand and let $\delta$be positive number . We say that an absolutely continuous (a.c.) function

$u: [T_1, T_2] \to R^n$ satisfying $u(T_1) = x , u(T_2) = y$ is a $\delta -$ approximate solution of the problem (P) if

$$\int_{T_1}^{T_2} f\big(t, u(t), u'(t)\big)dt \leq \int_{T_1}^{T_2} f\big(t, z(t), z'(t)\big)dt + \delta$$

For each a.c. function $z: [T_1, T_2] \to R^n$ satisfying $z(T_1) = x , z(T_2) = y$.

The main results of this section deals with the so-called turnpike property of the variational problems (P) . To have this property means, roughly speaking, that the solutions of the problems (P) are determined mainly by the integrand (cost function), and are essentially independent of the choice of interval and endpoint conditions, except in regions close to the endpoints .

Let us now define the space of integrands. Denote by $|\cdot|$ the Euclidean norm in $R^n$.

Let a positive constant and let $\psi: [0,\infty) \to [0,\infty)$ bean increasing function such that $\psi(t) \to +\infty$ as $t \to \infty$ . Denote by $\mathcal{M}$ the set of all continuous functions $f: [0,\infty) \times R^n \times R^n \to R^1$ which satisfy the following assumptions:

**A(i)** the function $f$ is bounded on $[0,\infty) \times E$ for any bounded set $E \subset R^n \times R^n$;

**A(ii)** $f(t,x,y) \geq \max\{\psi|x|, \psi(|u|)|u|\} - a$ for each $(t,x,u) \in [0,\infty) \times R^n \times R^n$;

**A(iii)** for each $M, \epsilon > 0$ there exists $\Gamma, \delta > 0$ such that

$$|f(t,x_1,u) - f(t,x_2,u)| \leq \epsilon \max\{f(t,x_1,u), f(t,x_2,u)\}$$

For each $t \in [0,\infty)$ and each $u, x_1, x_2 \in R^n$ which satisfy

$$|x_i| \leq M , i = 1,2, \ |u| \geq \Gamma , |x_1 - x_2| \leq \delta.$$

**A(iv)**for each $M, \epsilon > 0$ there exists $\delta > 0$ such that$|f(t,x_1,u_1) - f(t,x_2,u_2)| \leq \epsilon$

For each $t \in [0,\infty)$ and each $u_1, u_2, x_1, x_2 \in R^n$ which satisfy

$$|x_i|, |u_i| \leq M \ i = 1,2 , \max\{|x_1 - x_2|, |u_{1-}u_2|\} \leq \delta.$$

For each set $\mathcal{M}$ we consider the uniformity his is determined by the following base :

$$E(N,\epsilon,\lambda) = \{(f,g) \in \mathcal{M} \times \mathcal{M}\}: |f(t,x,u) - g(t,x,u)| \leq \epsilon$$

For each $t \in [0,\infty)$ and each $x, u \in R^n$ satisfying $|x| \leq N$ , where $N > 0, \epsilon > 0, \lambda < 1$ .

We consider functionals of the from

$$I^f(T_1, T_2, x) = \int_{T_1}^{T_2} f(t, x(t), x'(t))dt \qquad (5.1)$$

Where $f \in \mathcal{M}$ ,$0 \leq T_1 < T_2 < +\infty$ and $x: [T_1, T_2] \to R^n$ is an a.c. function .

For $f \in \mathcal{M}, y, z \in R^n$ and numbers $T_1, T_2$ satisfying $0 \leq T_1 < T_2$ we set

$U^f(T_1, T_2, y, z) =$
$\inf\{I^f(T_1, T_2, x): x: [T_1, T_2] \to R^n$ is an a. c. function satsfying $x(T_1) =$
$y, x(T_2) = z\}$ (5.2)

It is easy to see that $-\infty < U^f(T_1, T_2, y, z) < +\infty$ for each $f \in \mathcal{M}$, each $y, z \in R^n$
And all numbers $T_1, T_2$ satisfying $0 \le T_1 < T_2$ . Let $f \in \mathcal{M}$.Absolutely continuous function $x: [T_1, T_2] \to R^n$ is called an $(f)$- good function if for any a.c function a.c function $y: [T_1, T_2] \to R^n$ there is a number $M_y$ such that

$$I^f(0, T, y) \ge M_y + I^f(0, T, x) \text{ or each } T \in (0, \infty) .$$

**Proposition 5.1**

Let $f \in \mathcal{M}$ and let $x: [0, \infty) \to R^n$ be a bounded a.c. function then the $x$ is $(f)$- good function if and only if there is $M > 0$ such that

$$I^f(0, T, x) \le U^f(0, T, x(0), x(T)) + M \text{ for any } T > 0 .$$

We introduced two properties (P1) and (P2) and show that $f$ has turnpike property if and only if $f$ possesses the properties (P1) and (P2) . the Property (P2) means that all $(f)$- good functions , function $\upsilon: [0, T_2] \to R^n$ is an approximate solution and $T$ is large enough , then there is $\tau \in [0, T]$ such that $\upsilon(\tau)$ is close to $X(\tau)$ .

**Theorem 5.1**

Let $f \in \mathcal{M}$ and $X_f: [0, \infty) \to R^n$ be a bounded continuous function . Then $f$ has the turnpike property with $X_f$ being the turnpike if and only if the following tow properties holds:

**(P1)** For each $K, \epsilon > 0$ there exist $\gamma, \iota > 0$ such that for each $T \ge 0$ and each a.c. function $w: [T, T + l] \to R^n$ which satisfies

$$|w(T)|, |\mathrm{w}(T + l| \le K, I^f(T, T + l, w) \le U^f(T, T + l, w(T), w(T + l) + \gamma$$

There is $\tau \in [T, T + l]$ for which $|X_f(\tau) - \upsilon(\tau)| \le \epsilon$.

**(P2)** For each $(f)$- good function $v: [0,\infty) \to R^n$,

$$\left|v(t) - X_f(t)\right| \to 0 \text{ as } t \to \infty.$$

**Theorem 5.2**

Let $f \in \mathcal{M}$ and let $X_f: [0,\infty) \to R^n$ be an $(f)$- good function . Assume that the properties (P1) ,(P2) hold . then for each $K, \epsilon > 0$ there exist $\delta, L > 0$ and a neighborhood $\mathcal{U}$ of $f$ in $M$ such that for each $g \in \mathcal{U}$ each $T_1 \geq 0, T_2 \geq T_1 + 2L$ and each a.c. function $v: [T_1, T_2] \to R^n$ which satisfies

$$|v(T_1)|, |v(T_2)| \leq K \ , I^g(T_1, T_2, v) \leq U^g\big(T_1, T_2, v(T_1), v(T_2)\big) + \delta$$

The inequality $\left|v(t) - X_f(t)\right| \leq \epsilon$ holds for all $t \in [T_1 + L, T_2 - L]$ .

## 5.3 Strong Turnpike Property for Variational Problems[48]

**Proposition 5.2**

Let $f \in \mathcal{M}$ and let for each $(t, x) \in [0,\infty) \times R^n$ the function $f(t, x, \cdot): R^n \to R^1$ be convex .Then for each $z \in R^n$ there is a bounded $(f)$- good function $Z: [0,\infty) \to R^n$ such that $Z(0) = z$ and that for each $T > 0$ ,

$$I^f(0, T, Z) = U^f(0, T, Z(0), Z(T))$$

We say that $f$ has the strong turnpike property , or briefly (STP) , if there exists a bounded a.c. function $X_f$ Which satisfies the following condition :

For each $K, \epsilon > 0$ there exist constants $\delta, L > 0$ such that f each $T_1 \geq 0$ , $T_2 \geq T_1 + 2L$ or and each a.c. function $v: [T_1, T_2] \to R^n$ which satisfies

$$|v(T_1)|, |v(T_2)| \leq K \ , I^f(T_1, T_2, v) \leq U^f\big(T_1, T_2, v(T_1), v(T_2)\big) + \delta$$

**(i)** there are $\tau_1 \in [T_1, T_1 + L]$ and $\tau_2 \in [T_2 - L, T_2]$ for which

$$\left|v(t) - X_f(t)\right| \leq \epsilon \ , t \in [\tau_1, \tau_2]$$

**(ii)** if $\left|v(T_1) - X_f(T_1)\right| \leq \delta$ then $\tau_1 = T_1$ and if $\left|v(T_2) - X_f(T_2)\right| \leq \delta$ then $\tau_2 = T_2$

The function $X_f$ is called turnpike of $f$ .

Let $f \in \mathcal{M}$. We say that an a.c. function $x: [0, \infty) \to R^n$ is $(f)$-overtaking optimal if for each a.c. function $y: [0, \infty) \to R^n$ satisfying $y(0) = x(0)$ ,

$$\lim_{T\to\infty} sup\left[ I^f(0, T, x) - I^f(0, T, y)\right]$$

**Theorem 5.3**

Let $f \in \mathcal{M}$ , for each $(t, x) \in [0, \infty) \times R^n$ the function $f(t, x, \cdot): R^n \to R^1$be convex and let $X_f: [0, \infty) \to R^n$ be a bounded a.c. function . Then $f$ has the strong turnpike property with $X_f$ being the turnpike if and only if the following there properties hold :

**(P1)** For each pair of $(f)$-good functions $v_1, v_2: [0, \infty) \to R^n$

$$|v_1(t) - v_2(t)| \to 0 \; as \; t \to \infty$$

**(P2)** $X_f$ is an $(f)$-overtaking optimal function and if an $(f)$-overtaking optimal function $v: [0, \infty) \to R^n$ satisfies $v(0) = X_f(0)$ then $v = X_f$ .

**(P3)** For each $K, \epsilon > 0$ there exist $\gamma, l > 0$ such for each $T \geq 0$ and each a.c. function $w: [T, T + l] \to R^n$ which satisfies

$$|w(T)|, |\mathrm{w}(T + l| \leq K, I^f(T, T + l, w) \leq U^f(T, T + l, w(T), w(T + l) + \gamma$$

There is $\tau \in [T, T + l]$ for which $\left|X_f(\tau) - v(\tau)\right| \leq \epsilon$.

## 5.4 Discrete-Time Unconstrained Problem[48]

In this section we analyze the structure of solutions of the optimization problem

$$\sum_{\boldsymbol{i=m_1}}^{\boldsymbol{m_{2-1}}} v_i\,(z_{\boldsymbol{i}}, z_{i+1}) \to \min \;, \{z_i\}_{i=m_1}^{m_2} \subset X \text{ and } z_{m_1} = x \,, z_{m_2} = y \qquad \text{(p)}$$

Let $Z = \{0, \pm1, \pm2, \dots\}$ be the set of all integers $(X, \rho)$ be compact matric space and let $v_i: X \times X \to R^1, i = 0, \pm1, \pm2, \dots$ be a sequence of continuous functions such that

$$sup\{|v_i(x,y)|: x,y \in X, i \in Z\} < \infty$$

And which satisfy the following assumption:

(**A**) For each $\epsilon > 0$there exists $\delta > 0$ such that if $i \in Z$and if

$$x_1, x_2, y_1, y_2 \in X$$

Satisfy $\rho(x_j, y_j) \leq \delta, j = 1,2$, then $|v_i(x_1, x_2) - v_i(y_1, y_2)| \leq \epsilon$

(**B**) For each $y, z \in X$and each pair of integers $n_1, n_2 > n_1$ set

$$\sigma(n_1, n_2, y, z) = \inf\left\{\sum_{i=n_1}^{n_2-1} v_i(x_i, x_{i+1}) : \{x_i\}_{i=n_1}^{n_2} \subset X, x_{n_1} = y, x_{n_2} = z\right\}$$

$$\sigma(n_1, n_2,) = \inf\left\{\sum_{i=n_1}^{n_2-1} v_i(x_i, x_{i+1}) : \{x_i\}_{i=n_1}^{n_2} \subset X\right\}$$

Choose a positive number $d_0$ such that

$$|v_i(x,y)| \leq d_0, x, y \in X, i \in Z$$

A sequence $\{y_i\}_{i=-\infty}^{\infty} \subset X$ is called good if there is $c > 0$ such that for each pair of integers $m_1, m_2 > m_1$

$$\sum_{m_1}^{m_2-1} v_i(y_i, y_{i+1}) \leq \sigma(m_1, m_2, y_{m_1}, y_{m_2}) + c$$

We say that the sequence $\{v_i\}_{i=-\infty}^{\infty}$ has turnpike property (TP) if there exists a sequence $\{\hat{x}_i\}_{i=-\infty}^{\infty} \subset X$ which satisfies the following condition:

For each $\epsilon > 0$ there are $\delta > 0$ and a natural number $N$ such that for each pair of integers $T_1, T_2 \geq T_1 + 2N$ and each sequence $\{y_i\}_{i=T_1}^{T_2} \subset X$ which satisfies

$$\sum_{i=T_1}^{T_2-1} \upsilon_i(y_i, y_{i+1}) \leq \sigma\big(T_1, T_2, y_{T_1}, y_{T_2}\big) + \delta$$

Where are integers $\tau_1 \in \{T_1, \dots, T_1 + N\}$ $\tau_2 \in \{T_2 - N, \dots, T_2\}$ such that :

**(i)** $\rho(y_i, \hat{x}_i) \leq \epsilon$ , $i = \tau_1, \dots, \tau_2$ ;

**(ii)** If $\rho\big(y_{T_1}, \hat{x}_{T_1}\big) \leq \delta$ Then $\tau_1 = T_1$ and if $\rho\big(y_{T_2}, \hat{x}_{T_2}\big) \leq \delta$ , Then $\tau_2 = T_2$

The sequence $\{\hat{x}_i\}_{i=-\infty}^{\infty} \subset X$ is called the turnpike of $\{\upsilon_i\}_{i=-\infty}^{\infty}$

The turnpike property is very important for applications .Suppose our sequence of cost functions $\{\upsilon_i\}_{i=-\infty}^{\infty}$ has the turnpike property and we know a finite number of "approximate " solution of the problem (P). Then we know the turnpike $\{\hat{x}_i\}_{i=-\infty}^{\infty}$ , or at least its approximation ,and the constant N which is an estimate for the time period to reach the turnpike.

**Theorem 5.4**

Let $\{\hat{x}_i\}_{i=-\infty}^{\infty} \subset X$ . Then the sequence $\{\upsilon_i\}_{i=-\infty}^{\infty}$ has turnpike property and $\{\hat{x}_i\}_{i=-\infty}^{\infty}$ is its turnpike if and only if the properties hold:

(**p1**) if $\{y_i\}_{i=-\infty}^{\infty} \subset X$ is good ,then

$$\lim_{i\to\infty} \rho(y_i, \hat{x}_i) = 0 \, , \lim_{i\to-\infty} \rho(y_i, \hat{x}_i) = 0$$

(**P2**) For each pair of integers $m_1, m_2 > m_1$

$$\textstyle\sum_{m_1}^{m_2-1} \upsilon_i(\hat{x}_i, \hat{x}_{i+1}) \leq \sigma\big(m_1, m_2, \hat{x}_{m_1}, \hat{x}_{m_2}\big)$$

And sequence $\{y_i\}_{i=-\infty}^{\infty} \subset X$ satisfies

$$\textstyle\sum_{m_1}^{m_2-1} \upsilon_i(y_i, y_{i+1}) \leq \sigma\big(m_1, m_2, y_{m_1}, y_{m_2}\big)$$

For each pair of integers $m_1, m_2 > m_1$ , then $y_i = \hat{x}_i$ , $i \in Z$

(p3) For each $\epsilon > 0$ there are $\delta > 0$ and a natural number $L$ such that for each integer $m$ and each sequence $\{y_i\}_{i=m}^{m+L} \subset X$ satisfies

$$\sum_{m_1}^{m_2-1} v_i(y_i, y_{i+1}) \leq \sigma(m, m+L, y_m, y_{m+L}) + \delta$$

There is $j \in \{m, \dots, m+L\}$ for which $\rho\left(y_{T_1}, \hat{x}_{T_1}\right) \leq \epsilon$

## 5.5 Turnpike Properties for Finite and Infinite Optimal Control Problem

The turnpike phenomenon is a property of trajectories of optimally controlled systems, that has long been observed in optimal control, even back to early work by von Neumann . The turnpike property describes the fact that an optimal trajectory "most of the time" stays close to an equilibrium point, as illustrated in Figure 5.2, below, for finite horizon optimal trajectories. This property attracted significant interest, particularly in the field of mathematical economics , because it directly leads to the concept of optimal economic equilibria and thus provides a natural economic interpretation of optimality.

### 5.5.1 Setting and Preliminaries [82]

We consider possibly discounted discrete time optimal control Problem

$$\text{minimize}_{u \in U^n(x_0)} J_N(x_0, u) \qquad (5.3)$$

Where

$$J_N(x_0, u) := \sum_{k=0}^{N-1} \beta^k \, \ell(x(k), u(k))$$

$$x(k+1) = f\left(x(k), u(k)\right), x(0) = x_0 \qquad (5.4)$$

$N = \infty, \beta \in ]0,1], f: X \times U \to X$ , and $\ell: X \times U \to \mathbb{R}$ for matric spaces $X$ and $U$ , state and input constraints $\mathbb{X} \subset X, \mathbb{U} \subset U$ and admissible control sets

$$\mathbb{U}^N(x_0) := \left\{ u(\cdot) \in U^N : \begin{matrix} x(n) \in \mathbb{X} \forall_n = 0, \dots, Nand \\ x(n) \in \mathbb{X} \forall_n = 0, \dots, N-1 \end{matrix} \right\}$$

## 5.5.2 The Undiscounted Case [82]

We consider the undiscounted case , the case $\beta = 1$. In this case, it is not guaranteed that$\sum_{k=0}^{\infty} \ell(x(k), u(k)) = \lim_{k\to\infty} J_k(x_0, u)$ exists for all trajectories and controls. In order to avoid the introduction of complicated constraints on the set of controls over which we minimize, we use $J_\infty(x_0, u) = \limsup_{k\to\infty} J_k(x_0, u)$in this case. Still, in order to obtain a meaningful optimal control problem, we need to ensure that $|V_\infty(x)|$ is finite for all $x \in \mathbb{X}$. A class of optimal control problems for which this can be achieved are dissipative optimal control problems satisfying a certain controllability property

**Definition 5.1[82]**

The optimal control problem is called strictly dissipative at an equilibrium $(x^e, u^e)$ if there exists a so-called storage function $\lambda: \mathbb{X} \to \mathbb{R}$ which is bounded from below, and a function $\rho \in \kappa$ such that for all $x \in \mathbb{X}$ and $u \in \mathbb{U}$ with the $f(x, u) \in \mathbb{X}$ inequality

$$\ell(x, u) - \ell(x^e, u^e) + \lambda(x) - \lambda(f(x, u) \geq \rho(d(x, x^e)) \quad (5.5)$$

holds. The optimal control problem is called dissipative if the same condition holds with $\rho \equiv 0$ .

**Definition5.2 : (finite horizon turnpike property)[82]**

The optimal control problem (2.1) has the finite horizon robust turnpike property at an equilibrium $x^e \in \mathbb{X}$ , if for each $\delta > 0$ , $\epsilon > 0$ ,and each bounded set $\mathbb{X}_b \subset \mathbb{X}$

there is a constant $C^{fin}_{\delta,\epsilon,\mathbb{X}_b} \in \mathbb{N}$ such that all trajectories $x(k), u(k))$ satisfying $J_N(x_0, u) \leq V_N(x_0) + \delta$ satisfy

$$\#\{k \in \{0, \dots, N\}: d(x(k), x^e) \geq \varepsilon\} \leq C^{fin}_{\delta,\epsilon,\mathbb{X}_b} \qquad (5.6)$$

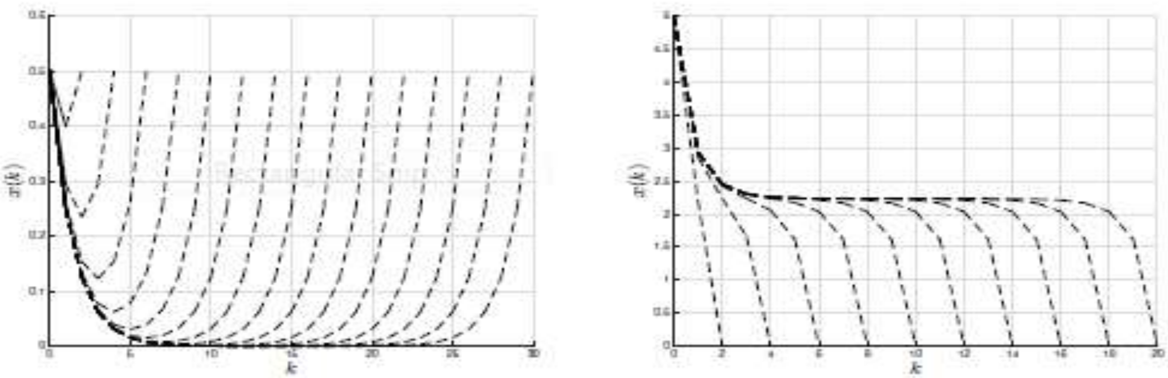


**Figure 5.1:** Finite horizon optimal control trajectories $x(\cdot)$ for different optimization horizons $N = 2,4, \dots 30$ (left) and $N = 2,4, \dots ,20$ (right)

**Definition 5.3:( infinite horizon turnpike property**)[82]

Consider an optimal control problem (5.3) with $N = \infty$ and $|V_\infty(x)| < \infty$ for all $x \in \mathbb{X}$ .Then the problem (5.3) has the infinite horizon robust turnpike property at an equilibrium $x^e \in \mathbb{X}$ , if for each $\delta > 0$ , $\epsilon > 0$ ,and each bounded set $\mathbb{X}_b \subset \mathbb{X}$ there is a constant $C^{fin}_{\delta,\epsilon,\mathbb{X}_b} \in \mathbb{N}$ such that all trajectories $(x(k), u(k))$ satisfying $J_N(x_0, u) \leq V_\infty(x_0) + \delta$ satisfy

$$\#\{k \in N: d(x(k), x^e) \geq \varepsilon\} \leq C^{\infty}_{\delta,\epsilon,\mathbb{X}_b} \qquad (5.7)$$

## 5.5.3 The Discounted Case[82]

We now turn our attention to the discounted case with $\beta \in ]0,1[$. For our analysis, the decisive difference in the discounted case is that the discount factor $\beta^K$tends to

0 as $k$ tends to infinity. This means that if a trajectory has a large deviation from the optimal trajectory, then this large deviation may nevertheless be barely visible in the cost functional, provided it happens sufficiently late. For this reason, it is unreasonable to expect that one can see the turnpike behavior for trajectories satisfying $J_N(x,u) \leq V_N(x) + \delta$. In order to fix this problem, we need make two changes to the robust turnpike Definitions 5.1 and 5.2. First, we need to restrict the time interval on which we can expect to see the turnpike phenomenon and second, we need to limit the difference $\delta$ between the value of the trajectory under consideration and the optimal value.

**Definition5.4 : (infinite horizon turnpike property)[82]**

The optimal control problem (2.1) has the infinite horizon near optimal approximate turnpike property, if for each $\epsilon > 0$ and each bounded set $\mathbb{X}_b \subset \mathbb{X}$ there is a constant $C_{\epsilon,\mathbb{X}_b}^{fin} > 0$ such that for each $M \in \mathbb{N}$ there is a constant

$\delta = \delta_{\epsilon,M,\mathbb{X}_b}^{fin} > 0$ such that for all $N \in \mathbb{N}$ with $N \geq M$, all trajectories $\big(x(k), u(k)\big)$ with $x_0 \in \mathbb{X}_b$ ,$u(\cdot) \in U^N(x_0)$ and $J_N(x_0,u) \leq V_N(x_0) + \delta$ satisfy

$$\#\{k \in \{0,\dots,M\}: d(x(k),x^e) \geq \varepsilon\} \leq C_{\epsilon,\mathbb{X}_b}^{fin}$$

**Definition 5.5: (finite horizon turnpike property)[82]**

The optimal control problem (2.1) has the infinite horizon near optimal approximate turnpike property, if for each $\epsilon > 0$ and each bounded set $\mathbb{X}_b \subset \mathbb{X}$ there is a constant $C_{\epsilon,\mathbb{X}_b}^{fin} > 0$ such that for each $M \in \mathbb{N}$ there is a constant

$\delta = \delta_{\epsilon,M,\mathbb{X}_b}^{fin} > 0$ such that all trajectories $\big(x(k), u(k)\big)$ with $x_0 \in \mathbb{X}_b$ ,$u(\cdot) \in U^\infty(x_0)$ and $J_\infty(x_0,u) \leq V_\infty(x_0) + \delta$ satisfy

$$\#\{k \in \{0,\dots,M\}: d(x(k),x^e) \geq \varepsilon\} \leq C_{\epsilon,\mathbb{X}_b}^{\infty}$$

## 5.6 The Turnpike Properties in Finite Dimensional Nonlinear Optimal Control

### 5.6.1 Dynamic Optimal Control Problem [12]

Consider the nonlinear control system

$$\dot{x}(t) = f(x(t), u(t)) \tag{5.8}$$

Where $f: R^n \times R^m \to R^n$ is of class $C^2$.Let $R = R^n \times R^n \to R^k$ be mapping of class $C^2$ , and let $f^0: R^n \times R^m \to R^n$ be a function of class $C^2$.For a given $T > 0$ we consider the optimal control problem of determining a control $u_T(\cdot) \in L^\infty(0, T; R^m)$ minimizing the cost functional

$$C_T(u) = \int_0^T f^0\big(x(t), u(t)\big)dt \tag{5.9}$$

Over all controls $u_T(\cdot) \in L^\infty(0, T; R^m)$ ,where $x(\cdot)$ is the solution of (5.8) corresponding to the control $u(\cdot)$ and such that

$$R\big(x(0), x(T)\big) = 0$$

Optimal solution :$(x_T(\cdot), u_T(\cdot))$ .

### 5.6.2 Pontryagin Maximum Principle [46]

There must exist an absolutely continuous mapping $\lambda_T(\cdot): [0, T] \to \mathbb{R}^n$, called adjoint vector and a real number $\lambda_T^0 \leq 0$ ,with $(\lambda_T(\cdot), \lambda_T^0) \neq (0,0)$

Such that

$$\dot{x}_T(t) = \frac{\partial H}{\partial \lambda}(x_T(t), \lambda_T(t), \lambda_T^0, u_T(t))$$

$$\dot{x}_T(t) = -\frac{\partial H}{\partial x}(x_T(t)\lambda_T(t), \lambda_T(t), \lambda_T^0, u_T(t)) \tag{5.10}$$

$$\frac{\partial H}{\partial u}(x_T(t), \lambda_T(t), \lambda_T^0, u_T(t)) = 0$$

Where

$$H(x, \lambda, \lambda^0, u) = \langle \lambda, f(x, u) \rangle + \lambda^0 f^0 \langle x, u \rangle \tag{5.11}$$

Moreover transversality condition s

$$\begin{pmatrix} -\lambda_T(0) \\ \lambda_T(T) \end{pmatrix} = \sum_{i=1}^{k} \gamma_i \nabla R^i(\, x_T(0), x_T(T))$$

Assumption made throughout:no abnormal⇒ $\lambda_T^0 = -1$

### 5.6.3 Static Optimal Control Problem[12]

We consider the static optimal control problem

$$\min_{\substack{(x,u)\in R^n \\ f(x,u)=0}} f^0(x,u) \tag{5.12}$$

We assume that this minimization problem has at least one solution $(\bar{x}, \bar{u})$ .

Note that the minimize exists and unique whenever $f$ is linear in $x$ and $u$ for instance , and $f^0$ is positive definite quadratic from in $(x, u)$ . According to the Lagrange multipliers rule , there exists $(\bar{\lambda}, \bar{\lambda}^0) \in \mathbb{R}^n \times \mathbb{R} \setminus \{(0,0)\}$ ,with $\bar{\lambda}^0 \le 0$ such that

$$f(\bar{x}, \bar{u}) = 0$$

$$\bar{\lambda}^0 \frac{\partial f^0}{\partial x}(\bar{x}\,, \bar{u}) + \langle \bar{\lambda}, \frac{\partial \mathrm{f}}{\partial \mathrm{x}}(\bar{x}\,, \bar{u}) \rangle = 0$$

$$\bar{\lambda}^0 \frac{\partial f^0}{\partial u}(\bar{x}\,, \bar{u}) + \langle \bar{\lambda}, \frac{\partial \mathrm{f}}{\partial \mathrm{u}}(\bar{x}\,, \bar{u}) \rangle = 0$$

Using Hamiltonian $H$ defined by (5.11)

$$\frac{\partial H}{\partial \lambda}(\bar{x}, \bar{\lambda}, \bar{\lambda}^0, \bar{u}) = 0$$

$$-\frac{\partial H}{\partial x}(\bar{x}, \bar{\lambda}, \bar{\lambda}^0, \bar{u}) = 0 \tag{5.13}$$

$$\frac{\partial H}{\partial \lambda}(\bar{x}, \bar{\lambda}, \bar{\lambda}^0, \bar{u}) = 0$$

This is the optimality system of the static optimal control problem .

### 5.6.4 The Turnpike Property [6]

Since $\left(\bar{x}\,,\bar{\lambda}\,,\bar{u}\right)$ is an equilibrium point of the extremal equations(5.10) , it is natural expected that ,under appropriate assumptions if $T$ is large time then the optimal extremal solution $(x_T(\cdot),\lambda_T(\cdot),u_T(\cdot))$ of optimal control problem remains most of the time close to the static extremal point $\left(\bar{x}\,,\bar{\lambda}\,,\bar{u}\right)$ .More precisely , it is expected that if $T$ is large time then the optimal extremal is approximately consists of three pieces:

1-short –time : $(x_T(0),\lambda_T(0),u_T(0))\rightarrow\left(\bar{x}\,,\bar{\lambda}\,,\bar{u}\right)$ (arcs transient)

2-longe-time, stationary : $\left(\bar{x}\,,\bar{\lambda}\,,\bar{u}\right)$

3-short –time : $\left(\bar{x}\,,\bar{\lambda}\,,\bar{u}\right)\rightarrow(x_T(T),\lambda_T(T),u_T(T))$ (arcs transient)

The first and the third is arcs are seen as transient. The solution of an optimal control problem in large time should spend most of its time near a steady –state . In infinite horizon the solution should converge to that steady –state . In econometry such such stady-states are known is as Von Neumann points .The turnpike property means then , in this context , that large time optimal trajectories are expected to converge ,in some sense , to Von Neumann points.

## 5.7 The Linear Quadratic Case[12]

We treat a special case in the optimal control of systems, in which the state differential equations are linear in $x$ and $u$ and the objective functional is quadratic. A solution can be found in a slightly different way in this case and has a very nice format. In particular, we are able to eliminate the adjoint variable in the necessary conditions.

We consider linear quadratic optimal control problems

$$\min_{\substack{(x,u)\in R^n \\ f(x,u)=0}} f^0\,(x,u) \tag{5.14}$$

$$f(x,u)=\dot{x}=Ax+Bu \tag{5.15}$$

With $A$ a matrix of size $n \times n$ and $B$ a matrix of size $n \times m$. The objective functional is

$$f^0(x,u) = \frac{1}{2}(x - x^d)^* Q(x - x^d) + \frac{1}{2}(u + u^d)^* \mathrm{U}(u - u^d) \qquad (5.16)$$

$$x(0) = x_0 , \qquad x(T) = x_1$$

Where $Q$ is a $n \times n$ symmetric positive definite matrix and $U$ is a $m \times m$ symmetric positive definite matrix ,and where $x^d \in R^n$, $u^d \in R^m$are arbitrary . It is assumed that the pair $(A, B)$ satisfies the Kalman condition.

Besides ,it is clear that $(OPC)_T$ has a unique solution $(x_T(\cdot), u_T(\cdot))$ ,having

a normal extremal lift$(x_T(\cdot), \lambda_T, -1, u_T(\cdot))$ , and the control has the simple expression

$$u_T = u^d + U^{-1} B^* \lambda_T(t)$$

$$\dot{x}_T(t) = A x_T(t) + B U^{-1} B^* \lambda_T(t) + B u^d , \; x_T(0) = x_0 \qquad (5.17)$$

$$\dot{\lambda}_T(t) = Q x_T(t) - A^* \lambda_T(t) - Q x^d$$

The static optimal control problem

$$\min_{\substack{(x,u) \in R^n \times R^m \\ Ax + bu = 0}} \frac{1}{2}((x - x^d)^* Q(x - x^d) + (u + u^d)^* \mathrm{U}(u - u^d)) \qquad (5.18)$$

There exists $\bar{\lambda} \in R^n \backslash \{0\}$such that $\bar{u} = u^d + U^{-1} B^* \bar{\lambda}$ and

$$A\bar{x} + B U^{-1} B^* \bar{\lambda} + B u^d = 0 \qquad (5.19)$$

$$Q\bar{x} - A^* \bar{\lambda} - Q x^d = 0$$

**Theorem 5.5[44]**

- $U > 0 , Q > 0$
- $rank(B, BA, \dots, A^{n-1}B) = n$ (Kalman contition)

There exist constants $C_1 > 0$ and $C_2 > 0$ such that for every time $T > 0$ the optimal control problem $(\mathbf{OCP})_\mathbf{T}$ has a unique solution $\left(x_T(\cdot), \lambda_T(\cdot), u_T(\cdot)\right)$ , which satisfies

$$\|x_T(t) - \bar{x}\| + \left\|\lambda_T(t) - \bar{\lambda}\right\| + \|u_T(t) - \bar{u}\| \leq C_1(e^{-c_2 t} + e^{-c_2(2T-t)})$$

For every $t \in [0, T]$.

**Theorem 5.6 [49] :** Consider the LQ-problem with $(A, B)$ stabilizable, $Q = C^T C$ and state and control constraint sets $\mathbb{X} = \mathbb{R}^n$ and $U \subseteq \mathbb{R}^n$ . Then the following properties are equivalent.

(i) The problem is strictly dissipative at an equilibrium $(x^e, u^e) \in int(\mathbb{X} \times \mathbb{U})$ .

(ii) The problem has the turnpike property at an equilibrium $(x^e, u^e) \in int(\mathbb{X} \times \mathbb{U})$

(iii) The pair $(A, C)$ is detectable, i.e., all unobservable eigenvalues $\mu$ of $A$ satisfy $\mathrm{Re}(\mu) < 0$.

Moreover, if one of these properties holds, then the equilibria in (i) and (ii) coincide.

If, in addition, $\mathbb{U} = \mathbb{R}^m$ holds, then the exponential turnpike property holds

## 5.8 Control Affine Systems with Quadratic Cost [12]

We consider the class of control –affine system with quadratic cost, that is ,

$$f(x, u) = f_0(x) + \sum_{i=1}^{m} u_i f_i(x)$$

Where $f_i$ is a $C^2$ vector field in $\mathbb{R}^n$ , for every $i \in \{0, \dots, m\}$ and

$$f^0(x, u) = \frac{1}{2}(x - x^d)^* Q(x - x^d) + \frac{1}{2}(u - u^d)^* U(u - u^d)$$

With $Q$ a $n \times n$ symmetric positive definite matrix, and $U$ a $m \times m$ symmetric positive definite matrix. The matrices $Q$ and $U$ are weight matrices, as in the LQ case. In this framework, we have

$$H = \langle \lambda, f_0(x) \rangle + \sum_{i=1}^{m} u_i \langle \lambda, f_i(x) \rangle - \frac{1}{2}(x - x^d)^* Q(x - x^d) + \frac{1}{2}(u - u^d)^* U(u - u^d)$$

$$H_{ux} = \begin{pmatrix} \langle \bar{\lambda}, df_1(\bar{x}) \rangle \\ \vdots \\ \langle \bar{\lambda}, df_m(\bar{x}) \rangle \end{pmatrix}$$

$$H_{xx} = -Q + \langle \bar{\lambda}, d^2 f_0(\bar{x}) \rangle + \sum_{i=1}^{m} u_i \langle \bar{\lambda}, d^2 f_i(\bar{x}) \rangle$$

$$H_{uu} = -U$$

And hence

$$W = -H_{xx} + H_{xu} H_{uu}^{-1} H_{ux} = Q - H_{ux}^{-1} U^{-1} H_{ux} - \langle \bar{\lambda}, d^2 f_0(\bar{x}) \rangle + \sum_{i=1}^{m} u_i \langle \bar{\lambda}, d^2 f_i(\bar{x}) \rangle$$

Intuitively, the requirement that $W > 0$ says that the positive weight represented by Q has to be large enough in order to compensate possible distortion by the vector fields.

**Theorem5.6 [44]**

Suppose that the $H_{uu}$ matrix is symmetric negative-definite ,the matrix is $W = -H_{xx} + H_{xu} H_{uu}^{-1} \mathrm{H_{ux}}$ is symmetric positive –definite ,and the pair of matrix $A = H_{\lambda x} - H_{\lambda u} H_{uu}^{-1} \mathrm{H_{ux}}$ and $B = H_{\lambda u}$ satisfies the Kalman condition:

$\mathrm{Rank}(B, AB, \dots, A^{n-1}B) = n$

$(\bar{x}, \bar{\lambda})$ almost satisfies” the terminal + transversality conditions

Then for$T < 0$ large enough:

$$\|x_T(t) - \bar{x}\| + \|\lambda_T(t) - \bar{\lambda}\| + \|u_T(t) - \bar{u}\| \leq C_1 (e^{-c_2 t} + e^{-c_2 (2T - t)})$$

For every $t \in [0, T]$

# Chapter Six

# Results and Conclusions

## Model 1: Linear Quadratic Case [12]

Let us provide a simple example in order to illustrate the turnpike phenomenon in the linear quadratic case. Consider the two-dimensional control system

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = -x_1(t) + u(t) \qquad (6.1)$$

With fixed initial point $(x_1(0), x_2(0)) = (0,0)$ and the problem of minimizing the cost functional

$$\frac{1}{2}\int_0^T (x_1(t) - 2)^2 + (x_2(t) - 7)^2 + u(t)^2)dt$$

$$T = 30$$

Optimal solution of the static problem

$$f(\bar{x}, \bar{u}) = 0$$

$$\bar{x}_2 = 0\,, \bar{x}_1 = \bar{u}$$

$$\min_{\substack{x_{2=0} \\ x_1 = u}} \frac{1}{2}((x_1(t) - 2)^2 + (x_2(t) - 7)^2 + u(\mathrm{t})^2))$$

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,; B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

This system satisfies the condition of the theorem 5.5.

First, we have to apply the Maximum Principle, and we get:

Hamiltonian : $H = \lambda_1 x_2 - \lambda_2 x_1 + \lambda_2 u - \frac{1}{2}((x_1 - 2)^2 + (x_2 - 7)^2 + \mathrm{u}^2)$

Adjoin vector:

$$\dot{\lambda}_1 = -\frac{\partial H}{\partial x_1} = \lambda_2 + x_1 - 2$$
$$\dot{\lambda}_2 = -\frac{\partial H}{\partial x_2} = -\lambda_1 + x_2 - 7$$

Hamiltonian maximization condition:

$$0 = \frac{\partial H}{\partial u} = \lambda_2 - u \text{ at } \bar{u} \Longrightarrow \bar{u} = \lambda_2$$

Use optimality system of static optimal control(5.13)

$$\bar{\lambda}_2 + \bar{x}_1 - 2 = 0$$

$$-\bar{\lambda}_1 + \bar{x}_2 - 7 = 0 \qquad (6.2)$$

Substation ($\bar{x}_2 = 0$ , $\bar{x}_1 = \bar{u}$,) in system (6.2)

$$2\bar{\lambda}_2 - 2 = 0 \Rightarrow \bar{\lambda}_2 = 1 = \bar{u} = \bar{x}_1$$

$$-\bar{\lambda}_1 - 7 = 0 \Rightarrow \bar{\lambda}_1 = -7$$

The optimal solution of static problem is

$$\bar{x} = (1,0)\,, \bar{u} = 1 \text{ and } \bar{\lambda} = (-7,1)$$

$$\left(\bar{x}_1, \bar{x}_2, \bar{\lambda}_1, \bar{\lambda}_1, \bar{u}\right) = (1,0,-7,1,1)$$

Such solutions is the point called turnpike of the problem .

The turnpike property can be observed on Figure 6.1.As expected , except transient initial and final arcs , the extremal $(x_1(\cdot), x_2(\cdot),\lambda_1(\cdot), \lambda_2(\cdot), u(\cdot))$ remains close to the steady-state $(1,0,1,-7,1))$ .It can be noted that , along the interval of time $[0,30]$ ,the curves $x_1(\cdot\ )$ , $x_2(\cdot\ )$,$\lambda_1(\cdot), \lambda_2(\cdot), u(\cdot)$ oscillate around their steady –state value (with an exponential damping ) . This oscillation is visible on Figure 6.2, where one can see the successive (exponential small ) loop that ( $x_1(\cdot\ )$ , $x_2(\cdot\ )$) makes around the point $(1,0)$ . the number of loops tend to $+\infty$ as the final time $T$ tends to $+\infty$.

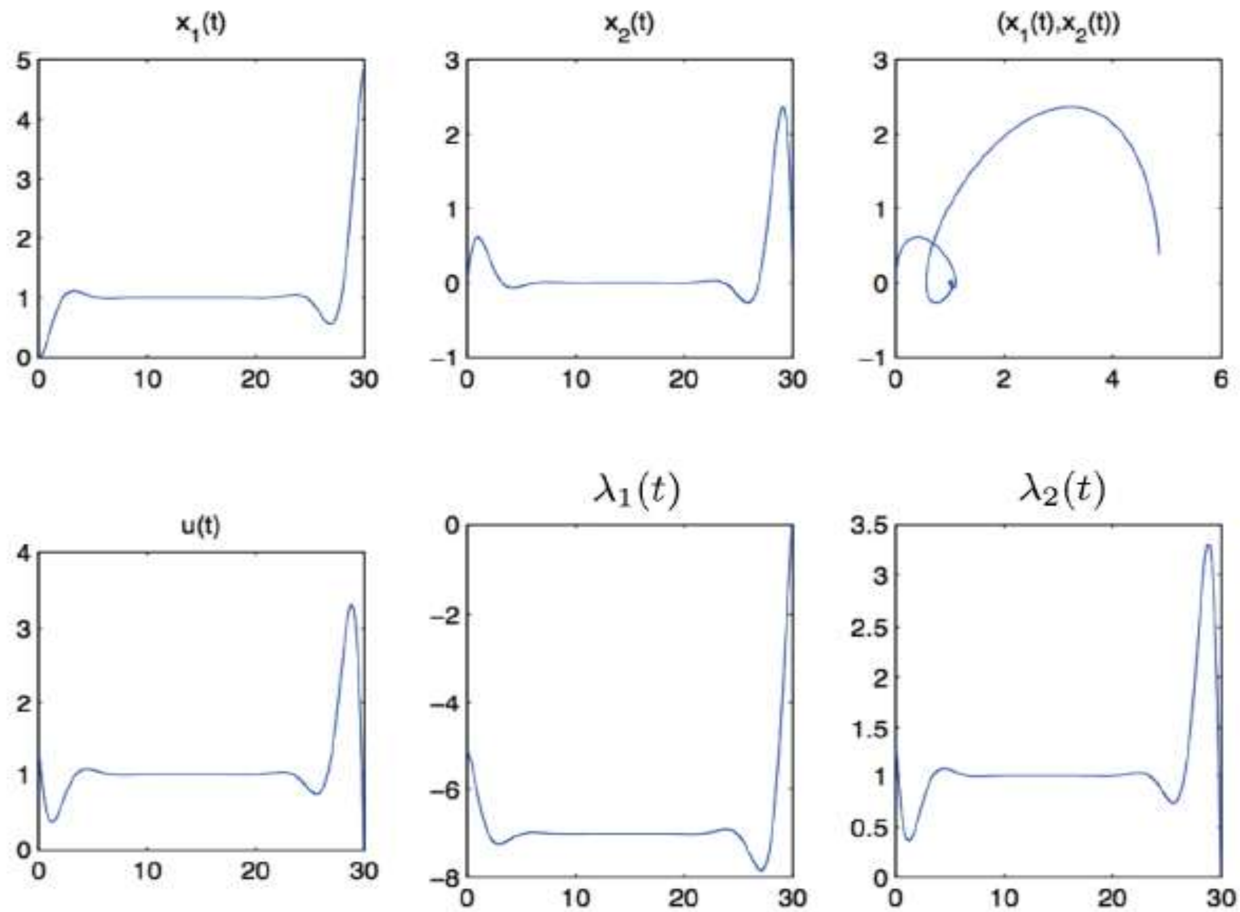



**Figure 6.1:** Turnpike in the LQ Case

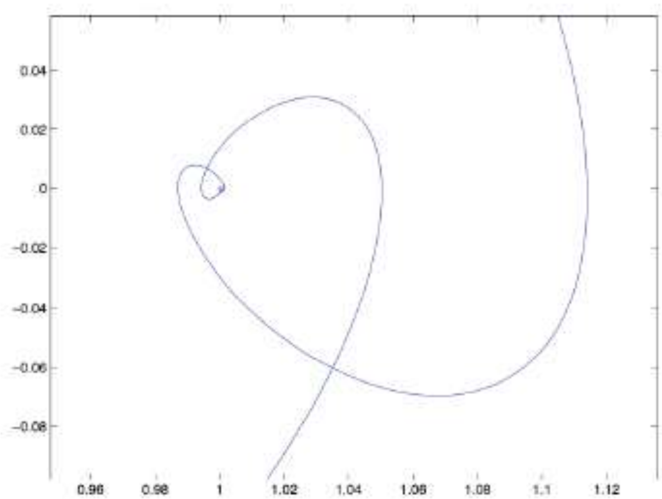


**Figure 6.2 :** Oscillation of $(x_1(\cdot), x_2(\cdot))$ around the steady-state (1,0).

**Model 2 : State Constraints in Turnpike Phenomena for LQ[49]**

Consider the LQ problem

$$\dot{x}(t) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t) , \quad t \in (0,T)$$

With $\mathbb{X} = \mathbb{R}^2$, $\mathbb{U} = [-10,10]$ and stage cost

$$\ell(x,u) = x_2^2 + 0.005u^2 , \qquad x = (x_1, x_2) .$$

With reference to the previous notations, $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , thus the pair $(A,B)$ is stabilizable (but not controllable) . Moreover, since $C = Q = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ ,the unobservable space, spanned by the eigenvector $x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ , corresponds to the eigenvalue $\lambda = -1$ of A, which has negative real part, and thus the pair $(A,C)$ is detectable. Thus the turnpike property holds for the system.

**Model 3 : Control Affine System with a Quadratic Cost [12]**

Let us provide a simple example in order to illustrate the turnpike phenomenon for control –affine system with a quadratic cost . Consider the optimal control problem of steering the two-dimensional control system

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = 1 - x_1(t) + x_2(t)^3 + u(t)$$

From the initial point $(x_1(0), x_2(0)) = (1,1)$ to the final point $(x_1(T), x_2(T)) = (3,0)$ , by minimizing the cost functional

$$\frac{1}{2}\int_0^T (x_1(t) - 1)^2 + (x_2(t) - 1)^2 + (u(t) - 2)^2)dt$$

An easy computation shows that the optimal solution of the static problem is given by

$$\bar{x}_2 = 0 \qquad , 1 - \bar{x}_1 + \bar{x}_2^3 + \bar{u} = 0$$

$$\min_{\substack{x_2 = 0 \\ 1 - x_1 + x_2^3 + u = 0}} \frac{1}{2}((x_1 - 1)^2 + (x_2 - 1)^2 + (u - 2)^2)$$

First, we have to apply the Maximum Principle, and we get:

Hamiltonian : $H = \lambda_1 x_2 + \lambda_2(1 - x_1 + \bar{x}_2^3 + u) - \frac{1}{2}((x_1 - 1)^2 + (x_2 - 1)^2 + (u - 2)^2)$

Adjoin vector:

$$\dot{\lambda}_1 = -\frac{\partial H}{\partial x_1} = \lambda_2 + x_1 - 1$$
$$\dot{\lambda}_2 = -\frac{\partial H}{\partial x_2} = -\lambda_1 + -3x_2^2 + x_2 - 1$$

Hamiltonian maximization condition:

$$0 = \frac{\partial H}{\partial u} = -\lambda_2 + u - 2 \text{ at } \bar{u} \Longrightarrow \bar{u} = \lambda_2 + 2$$

Use optimality system of static optimal control(5.13)

$$\lambda_2 + x_1 - 1 = 0$$
$$-\lambda_1 + -3x_2^2 + x_2 - 1 = 0 \qquad (6.3)$$

Substation ( $\bar{x}_2 = 0$ , $1 - \bar{x}_1 + \bar{x}_2^3 + \bar{u} = 0$ ) in system (6.3)

$$-\bar{\lambda}_1 - 1 = 0 \Rightarrow \bar{\lambda}_1 = -1$$
$$\bar{u} = \bar{x}_1 - 1 \qquad (6.4)$$
$$\bar{\lambda}_2 + \bar{u} = 0 \qquad (6.5)$$

Use maximization condition:

$$2\bar{\lambda}_2 + 2 = 0 \Rightarrow \bar{\lambda}_2 = -1$$

Substitute value $\bar{\lambda}_2$ in(6.5) Then $\bar{u} = 1$

$$\bar{x}_1 = \bar{u} + 1 \Rightarrow \bar{x}_1 = 2$$

The optimal solution of static problem is

$$\bar{x} = (2,0)\, , \bar{u} = 1 \text{ and } \bar{\lambda} = (-1,-1)$$
$$\left(\bar{x}_1, \bar{x}_2, \bar{\lambda}_1, \bar{\lambda}_1, \bar{u}\right) = (2,0,-1,-1,1)$$

Such solutions is the point called turnpike of the problem .

We compute the optimal solution $((x_1(\cdot), x_2(\cdot), \lambda_1(\cdot), \lambda_2(\cdot), u(\cdot))$ in time T = 20, by using a direct method. The result is drawn on Figure 6.3. Note that, according to the maximization condition of the Pontryagin maximum principle , we have $u(t) = 2 + \lambda_2(t)$ . The turnpike property can be observed on Figure 6.3 . As expected, except transient initial and final arcs, the extremal $((x_1(\cdot), x_2(\cdot), \lambda_1(\cdot), \lambda_2(\cdot), u(\cdot))$ remains close to the steady-state (2, 0, -1,-1, 1). It can be noted that, along the interval of time [0, 20], the curves $x_1(\cdot), x_2(\cdot), \lambda_1(\cdot), \lambda_2(\cdot)$, and $u(\cdot)$ oscillate around their steady-state value(with exponential damping). This oscillation can be seen on Figure 6.4, in the form of successive (exponentially small) heart-shaped loops that $(x_1(\cdot), x_2(\cdot))$ makes around the point (2, 0). The number of loops tends to $+\infty$ as the final time $T$ tends to $+\infty$. This turnpike property is shown for nonlinear control –affine systems where the vector fieled are assumed to by globally Lipschitz .It can be noted that ,due to the explosive term $x_2^3$ , the convergence of the above optimization problem may be difficult to ensure .We use here the particularly adequate initialization given by the solution of the static problem .Then the convergence is easily obtained. The convergence of an optimization solver with any other initialization would certainly not be ensured.

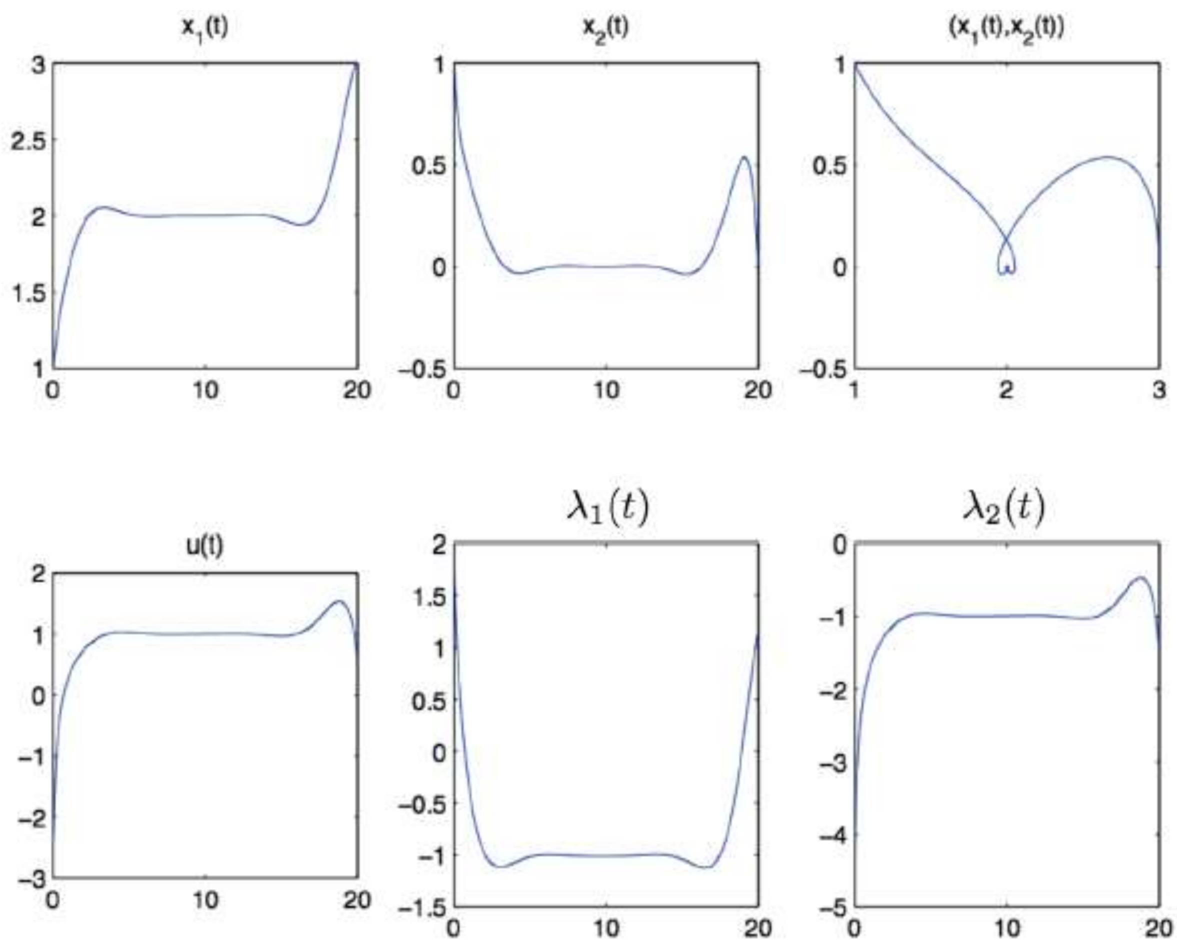



**Figure 6.3** : Example in the control- affine case

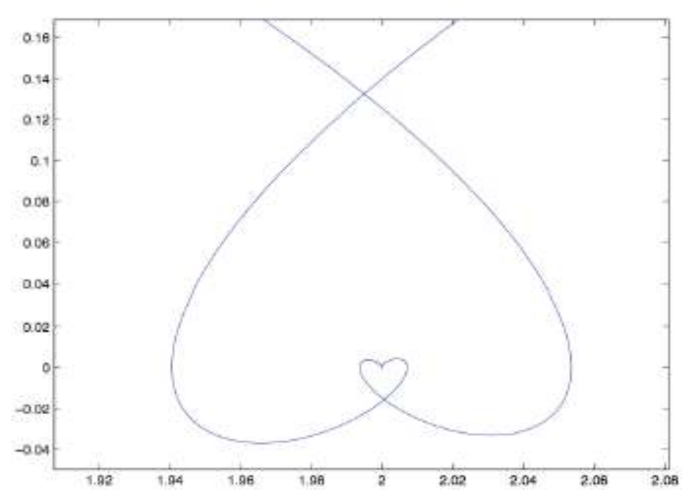


**Figure 6.4**: Oscillation of $(x_1(\cdot).x_2(\cdot))$ around the steady-state (2,0)

# Conclusions

Many real control problems must be considered in large time intervals, which requires precise numeric methods, since a small deviation in this kind of problem may have a very negative effect on the accuracy of the final solution .So for this problems ,the turnpike concept is particularly useful, because it allows performing a significant simplification of the analysis and computation of large-time horizon problems. Such application of turnpike property is very useful as it represent a good solution to the divergence and initialization problems of the classic method. necessary and sufficient condition for turnpike and near equilibrium turnpike properties in terms of spectral properties of the system matrices.

# References

[1] L.W. McKenzie, Turnpike Theory, Econometrica 44 (1976) ,841-865.

[2] Joseph,Disteefano,III,Ph.D,Department of Eletrical and computer Engineering University of Califonia.Feedback Control Systems,second Edition ,continuous (Analog)and Discrete(Digital).

[3]A.J Zaslavski, Turnpike Theory of Continuous-Time Linear Optimal Control Problems, Springer Optim. Appl. 104, Springer, New York, 2015.

[4] A.Ibanez, Optimal Control an Turnpike Properties of the Lotk- Vollterra model , Master Thesis, Universided del pais Vasco/Euskal Herriko Unibertsitatea . Available at http://www.bcamath.

[5] Vinter,R,B,Optimal Control ,Birkhauser ,Bostonn,2000.

[6] A.Ibanez, Optimal Control of Lotka-Voltrra system :turnpike property and numerical simulatioms ,(2016).

[7 ]A.porretta,E.Zuazua,Long time versus staeady state optimal control , SIAM J. Control optim. 51(2013), no.6,4242-4273.

[8]L.S.ponryagin,V.G.Boltyanskii,R.V.GramKreelidze,E.F.Miischenko,The mathematical theory of optimal processes ,Wiley Interscience ,1962.

[9] A.porretta,E.Zuazua , Remarks on long versus steady state optimal problem , Preprint (2014).

[10] S. Lenhart and J. T. Workman, Optimal Control Applied to Biological Models, Boca Raton, FL, Florida: Taylor & Francis Group, LLC, 2007.

[11]R. Dorfman, P. A. Samuelson, and R.M. Solow, Linear Programming and economic analysis, McGraw-Hill, New York, 1958.

[12] E. Tre'lat and E. Zuazua, The turnpike property in finite-dimensional nonlinear optimal control, J. Differential Equations, 258 (2015), pp. 81-114.
[13] Franklin,G. F, J.D. Powell, and A. Emami-Naeini, Feedback Control of Dynamic Systems,3rd ed. Reading, MA:Addison-Wesley Publishing Company, Inc., 1994.

[14] Dorf, R. C., and R. H. Bishop, Modern Control Systems, 9th ed. Upper Saddle River, NJ:Prentice Hall, 2001.

[15] Gilbert, E.G.,"Controllability and Observability in Multivariable Control Systems," *J.SIAMControl*, ser.A, **1** (1963) , pp. 128–51.

[16] U.A.Bakshi,M.V.Bakashi, Modern Control Theory. First Edition-2008.Tchnical publications pune.

[17]Herman,L.Russell, A course in Mathematical methods for physicists, 13: 978-1-4665-8468-6 ( (2014).

[18]Perko, Lawrence , Differential Equations and Dynamical Systems, Springer-Verlag, New York. Inc (2001).
[19]Kirk,Daniel.E,Optimal Control Theory an Introduction ,(1970)
[20]Liberzon Daniel , Calculus of Variation and Optimal Control Theory ,New York(2012).
[21] Katsuhiko,Ogata ,Modern Control Theory, Upper Saddle River, NJ:Prentice Hall
[22]L.C,E vans, An Introduction to Mathematical Optimal Control Theory ,Lecture Note.
[23] G. Fraser-Andrews. Numerical methods for singular optimal control.Journal of Optimization Theory and Applications, 61(3):377-401, 1989.
[24] John Gregory and Cantian Lin. Constrained Optimization in the Calculus of

Variations and Optimal Control. Van Nostrand Reinhold, New York, 1992.
[25] H. Halkin. A maximum principle of the Pontryagin type for systems described by nonlinear differential equations. J. SIAM Control, 4:90-111, 1966.
[26]Martino Bardial Italo Capuzzo-Dolcetta, Optimal Control and Viscosity Solutions of Hamilton-Jocobi -Bellman Equation, Birkhauser (1997).
[27]A.K, Nandakumaran , Calculus of Variations to Optimal Control :A short introduction , Lecture Note.
[28]Bruce,van Brunt , The Caculculus of Variations ,New York ,Lnc,2004.
[29] A. J. Zaslavski. Turnpike Phenomenon and Infinite Horizon Optimal Control. Springer International Publishing, 2014.
[30] A. J. Zaslavski. Turnpike Properties in the Calculus of Variations and Optimal Control. Springer, New York, 2006.
[31] E. D. Sontag. Mathematical Control Theory. Springer Verlag, New York, 2nd edition,1998.
[32] B.D. Anderson and P.V. Kokotovic, Optimal control problems over large time intervals, Automatica 23 (1987), pp. 355–363.
[33] P. A. Samuelson, “A catenary turnpike theorem involving consumption and the golden rule”, American Economic Review 55 (1965), 486-496
[34] A. J. Zaslavski, “Optimal programs on infinite horizon 1,2“, SIAM Journal on Control and Optimization 33 (1995), 1643-1686 .
[35] A. J. Zaslavski, “The existence and structure of extremals for a class of second order infinite horizon variational problems”, “Journal of Mathematical Analysis and Applications 194 (1995), 459-476.
[36] A. J. Zaslavski, “Structure of extremals for one-dimensional variational problems arising in continuum mechanics“, Journal of mathematical Analysis and Applications 198 (1996), 893-921.

[37] Aseev SM, Kryazhimskiy AV (2004) The Pontryagin maximum principle and transversality conditions for a class of optimal control problems with infinite time horizons. SIAM J Control Optim 43:1094–1119 .

[38] E. Tre'lat, Control optimal : the'orie & applications, Vuibert, Collection "Mathematiques Concretes" (2005), 246 pages.

[39] R.R. Wilde, P.V. Kokotovic, A dichotomy in linear control theory, IEEE Trans. Automatic Control 17 (1972), no. 3, 382-383.

[40] R.T. Rockafellar, Saddle points of Hamiltonian systems in convex problems of Lagrange, J. Optimization Theory Appl. 12 (1973), 367-390.

[41] H. Kwakernaak, R. Sivan, Linear optimal control systems, Wiley-Interscience, New York-London-Sydney, 1972, xxv+575 pp.

[42] A. Wachter, L.T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, Mathematical Programming 106 (2006), 25-57.

[43] J. Stoer and R. Bulirsch, Introduction to Numerical Analysis, New York, NY: Springer-Verlag, 1972.

[44]E.Trelat,Turnpike property in Optimal Control .Semminare Parisien d'optimisation ,2017.

[45] E.Trelat,Emmanuel, C. Zhang and Zuazua Enrique. (2018).Steady-state and periodic exponential turnpike property for optimal control problems in Hilbert Spaces, SIAM J. Control Optim, Vol. 56, No. 2, pp. 1222-1252.

[46]A. Ibanez, optimal control and turnpike properties of the Lotk- Vollterra model , Master Thesis, Universided del pais Vasco/Euskal Herriko Unibertsitatea . Available at http://www.bcamath.

[47] Kirsten Morris. Introduction to Feedback Control. Harcourt, San Diego, California, 2001.

[48] A. J. Zaslavski, Necessary and Sufficient Turnpike Conditions ,YoKohama Publishers,Volume 4, no 2,2019,463-476.

[49] L. Grune and M.A.Muller, On the relation between strict dissipativity and the turnpike property ,Syst.Contr.Lett.,90(2016),pp.45-53.

[50] A. J. Zaslavski, Turnpike properties of approximate solutions of autonomous variational problems, Control Cybernet., 37 (2008), pp. 491-512.

[51]T. Faulwasser, M. Korda, C. N. Jones, and D. Bonvin, Turnpike and dissipa- tivity properties in dynamic real-time optimization and economic MPC, in Proceedings of the 53rd IEEE Conference on Decision and Control , CDC 2014, 2014, pp. 2734-2739.

[52] A. J. Zaslavski, Necessary and sufficient of conditions for turnpike properties of solutions of optimal control systems arising in economic dynamics, Dynamics of Continuous ,Discrete and Impulsive Systems ,Ser.B Appl. Algorithms 20 (2013),391-420.

[53] A. J. Zaslavski,The turnpike property of discrete-time control problems arising in economic dynamics, Dynamics,Discrete Continuous Sys .B5(2005),861-880.

[54] M.A. Khan and A. Piazza. An overview of turnpike theory: towards the discounted deterministic case. Advances in Mathematical Economics, pages 39–67, 2011.

[55] A.J. Zaslavski. The turnpike property for approximate solutions of variational problems without convexity. Nonlinear Analysis, Theory, Methods and Applications, 58(5-6):547–569, 2004.

[56] J. C. Willems, Dissipative dynamical systems. II. Linear systems with quadratic supply rates, Arch. Ration. Mech. Anal., 45 (1972), pp. 352–393.

[57] B. D. O. Anderson and P. V. Kokotovi´c, Optimal control problems over large time intervals, Automatica J. IFAC, 23 (1987), pp. 355–363.

[58] L. C. Young. Lectures on the Calculus of Variations and Optimal Control Theory. Chelsea Pub. Co., New York, 2nd edition, 1980.
[59] B.D.O. Anderson and J. B. Moore. Optimal Control: Linear Quadratic Methods. Prentice Hall, New Jersey, 1990. Reprinted by Dover in 2007.
[60] V. I. Arnold. Ordinary Differential Equations. Springer, Berlin,3rd edition, 1992.
[61] R. W. Brockett. Finite Dimensional Linear Systems. Wiley, New York, 1970.
[62] I. M. Gelfand and S. V. Fomin. Calculus of Variations. Prentice Hall, New Jersey, 1963. Reprinted by Dover in 2000.
[63] R. Goebel and M. Subbotin. Continuous time linear quadratic regulator with control constraints via convex duality. IEEE Trans. Automat. Control, 52:886-892, 2007.
[64] J. P. Hespanha. Linear Systems Theory. Princeton University Press, 2009.
[65] A. Isidori. Nonlinear Control Systems. Springer, Berlin, 3rd edition, 1995.
[66]H. Kwakernaak and R. Sivan. Linear Optimal Control Systems. Wiley, New York, 1972.
[67] Erich Miersemann, Calculus of Variations.Lecture Note,2012.
[68] S. P. Sethi and G. L. Thompson. Optimal Control Theory: Applications to Management Science and Economics. Springer, New York, 2nd edition, 2005.
[69] Gelfand, I. M., Fomin, S. V., Calculus of variations, Prentice-Hall, 1963.
[70] N. I. Akhiezer. The Calculus of Variations. Blaisdell Publishing Company, Boston, MA, 1962.
[71] G. M. Ewing. Calculus of Variations with Applications. Dover Publications, Inc., New York, NY, 1985.
[72] A. J. Zaslavski: A turnpike result for autonomous variational problems, Optimization 53
(2004) 377-391.

[73] A. J. Zaslavski, A. Leizarowitz: Optimal solutions of linear control systems with nonperiodic integrands, Math. Oper. Res. 22 (1997) 726-746.

[74] Athans, M., and Falb, P. Optimal Control. 1966.

[75] Griffiths, D. F., and Higham, D. J. Numerical methods for ordinary differential equations. Springer, 2010.

[76] McKenzie, L. W. The turnpike theorem of morishima (1). Metroeconomica 15, 2-3 (1963), 146-154.

[77] Deniel Liberzon , Calculus of Variations and Optimal Control Theory,Princeton university press,2012.

[78] Morishima, M. Proof of a turnpike theorem: The" no joint production" case. The Review of Economic Studies (1961), 89-97.

[79] Leonov, G. Mathematical Problems of Control Theory: An Introduction.

[80] Naidu,D.s,Optimal control systems ,USA 2002.

[81] D. Hinrichsen, A. J. Pritchard, Mathematical systems theory I: modelling, state space analysis, stability and robustness, Vol. 1, Springer, 2005.

[82]Lars Grune, On the relation between Turnpike properties for finite and infinite horizon optimal control problems ,university of Bayreuth 95440 Bayreth,2017.

[83] Alexander J.Zaslavski ,A turnpike property for a class of variational problems ,Jounal of Convex Analyiss ,Volume 12 (2005),no .2,331-349.

[84] Lionel W.Mekenzie ,Turnpike Theory ,Dictionary of Economics 2nd edition 2008.

[85] Lionel W. Mekenzie, Turnpike Theory Website ,2011.