Sudan University of Science and Technology
College of Graduate Studies

كلية الدراسات العليا

جامعة السودان للعلوم والتكنولوجيا

The Maximum Principle Applications of Time- Optimality for Linear Control Systems

تطبيقات الأسس القصوى للزمن الامثل لانظمة التحكم الخطى

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

By:

Omer Bushra Elfadil Hassan

Supervisor:

Dr.Abdelrhim Bashir Hamid

2020

Sudan University of Science & Technology
College of Graduate Studies

بسم الله الرحمن الرحيم

كلية الدراسات العليا

جامعة السودان للعلوم والتكنولوجيا
كلية الدراسات العليا

Ref: SUST/ CGS/A11

Approval Page

(To be completed after the college council approval)

Name of Candidate: Omer Bushra ElFadil Hassan

Thesis title: The Maximum Principle Applications of Time - Optimality for Linear Control Systems

Degree Examined for: Ph. D.

Approved by:

1. External Examiner

Name: Osman Mohamed El Mekki

Signature: .... Date: 14/10/2020

2. Internal Examiner

Name: Prof. Shawgy Hussein AbdAlla

Signature: .... Date: 14/10/2020

3. Supervisor

Name: Abdelrahim Bashir Hamid

Signature: .... Date: 14-10-2020

Dedication

To my parents who always support me, to my wife who always encourage me, to my brothers and my sister, to my best friends who always stand beside and help me.

Acknowledgements

I am so grateful to my supervisor Dr.Abdelrhim Bashir Hamid, Associate professor, University of Gezira for continuous supervision and cooperation to achieve this work. I highly appreciate his sincerity, generosity and above all his humanitarian manner.

I would like to express my wholehearted thanks to my family and for their generous support they provided me throughout my entire life. They supported me to accomplish this work. Also, I would like to thank all my beloved friends for their effort. Thanks are extended to my colleagues. Finally, I hope this work may pave the way for others.

Abstract

Mathematical control theory is a branch of mathematics having as one of its main aims the establishment of a sound mathematical foundation for the control techniques employed in several different fields of applications, including engineering, economy, biology and so forth. The systems arising from these applied sciences are modeled using different types of mathematical formalism, primarily involving ordinary differential equations, or partial differential equations or functional differential equations. Optimal control theory-which is playing an increasingly important role in the design of modern systems-has as its objective the maximization or the minimization. In this research we consider a mathematical model of mosquito and insecticide, fish harvesting.The aim of these models is first, reduce the amount of mosquitoes in the ponds and swamps because Mosquitos are the main cause of malaria disease.We used the optimal spray strategies to minimize amount of mosquito. Second increase the profit to the maximum extent of the harvest during a specific time period. We used the strategies optimal control to maximize the profit of fish harvesting. We work optimal control framework by applying the Pontryagin's maximum principle. A characterization of the optimal control via adjoint variables was established. We obtained an optimality system that we sought to solve numerically by used MATLAB program.

الخلاصة

نظرية التحكم الرياضي هي فرع من فروع الرياضيات التي تهدف إلى إنشاء أساس رياضي سليم لتقنيات التحكم المستخدمة في العديد من المجالات التطبيقية المختلفة، الهندسة، الاقتصاد، علم الأحياء وما الى ذلك. تم تصميم الأنظمة الناشئة عن هذه العلوم التطبيقية باستخدام أنواع مختلفة من المعادلات الرياضية ، والتي تتضمن في المقام الأول المعادلات التفاضلية العادية ، أو المعادلات التفاضلية الجزئية أو المعادلات التفاضلية الدالية.

تهدف نظرية التحكم الأمثل - التي تلعب دورا متزايد الأهمية في تصميم الأنظمة الحديثة - إلى التعظيم أو التقليل. في هذا البحث نعتبر نموذجا رياضيا للبعوض والمبيدات الحشرية وحصاد الأسماك ، والهدف من هذه النماذج هو أولا تقليل كمية البعوض في البرك والمستنقعات لأن البعوض هو السبب الرئيسي لمرض الملاريا ، وقد استخدمنا استراتيجيات التحكم الأمثل لتقليل كمية البعوض. ثانيا زيادة الربح إلى أقصى حد للحصاد خلال فترة زمنية محددة. استخدمنا استراتيجيات التحكم الأمثل لتعظيم أرباح حصاد الأسماك. نحن نعمل فى إطار التحكم الامثل من خلال تطبيق مبدأ بونترياجن الأقصى. تم إنشاء توصيف للتحكم الأمثل عبر المتغيرات المجاورة. لقد حصلنا على نظام أمثل سعينا إلى حله عدديا باستخدام برنامج الماتلاب.

Introduction

Mathematics has always benefited from its involvement with developing sciences. Each successive interaction revitalizes and enhances the field. Biomedical science is clearly the premier science of the foreseeable future. For the continuing health of their subject, mathematicians must become involved with biology. With the example of how mathematics has benefited from and influenced physics, it is clear that if mathematicians do not become involved in the biosciences they will simply not be a part of what are likely to be the most important and exciting scientific discoveries of all time.

Mathematical biology is a fast-growing, well-recognized, albeit not clearly defined, subject and is, to my mind, the most exciting modern application of mathematics. The increasing use of mathematics in biology is inevitable as biology becomes more quantitative. The complexity of the biological sciences makes interdisciplinary involvement essential. For the mathematician, biology opens up new and exciting branches, while for the biologist, mathematical modeling offers another research tool commensurate with a new powerful laboratory technique but only if used appropriately and its limitations recognized. However, the use of esoteric mathematics arrogantly applied to biological problems by mathematicians who know little about the real biology, together with unsubstantiated claims as to how important such theories are, do little to promote the interdisciplinary involvement which is so essential.

A control system is an arrangement of physical components connected or relate in such a manner to command direct, or regulate itself or anther system. Control systems are classified into two general categories is open loop and closed loop system. Control system may have more than one input or output. Often all input and output are well defining by the system description. But sometimes they are not. To solve control systems problem, we must put the specification or description of the system configuration and its components into a form amenable to analysis or design. Three basic representations (model) of components and system are used extensively in the study of control systems is Mathematical models, block diagrams and signal flow graphs.

The calculus of variations a branch of mathematics that is extremely useful in solving optimization problems is the calculus of variations. A huge amount of problems in the calculus of variations have their origin in physics where one has to minimize the energy associated to the problem under consideration.

Optimal control theory is about controlling the given system in some 'best' way. The optimal control strategy will depend on what is defined as the best way. This is usually specified in terms of a performance index functional. The maximum principle is stated as a general assertion involving terms that are not yet precisely defined, and without a detailed specification of technical assumptions., where the terms are precisely defined and the appropriate technical requirements are completely specified, is stated for problems where all the basic objects-the dynamics, the Lagrangian and the cost functions for the switching's and the end-point constraints-are differentiable along the reference arc. We introduce optimal control theory for discrete-time systems. We begins with unconstrained optimization of a cost function and then generalizeto optimization with equality constraints.

MATLAB is needed to run the provided programs, it is certainly not needed to solve optimal control problems in general. Any mathematical programming language, such as FORTRAN or C++, is capable of the calculations needed. For each problem, there is a user-friendly interface that will guide you through. Each lab consists of two different MATLAB programs, *lab*

.m and *code .m*. For example, there are two programs associated with Lab 1, *lab1.m* and *code1.m*.

At the end of this research we presented a reducing and maximizing model, the reduction model represented by mosquitoes and insecticide, and the reduction model represented in fish harvesting.

The Contents

Chapter 1

Introduction and Basic Concept

1.1 Introduction

Mathematical control theory is a branch of mathematics having as one of its main aims the establishment of a sound mathematical foundation for the control Techniques employed in several different fields of applications, including engineering, economy, biology and so forth. The systems arising from these applied Sciences are modeled using different types of mathematical formalism, primarily involving ordinary differential equations, or partial differential equations or functional differential equations. Optimal control theory-which is playing an increasingly important role in the design of modern systems-has as its objective the maximization of the return from, or the minimization of the cost of, the operation of physical, social, and economic processes.

A control system is an arrangement of physical components connected or relate in such a manner to command direct, or regulate itself or anther system. Control systems are classified into two general categories is open loop and closed loop system. Control system may have more than one input or output. Often all input and output are well defining by the system description. But sometimes they are not. To solve control systems problem, we must put the specification or description of the system configuration and its components into a form amenable to analysis or design. Three basic representations (model) of components and system are used extensively in the study of control systems is mathematical models, block diagrams and signal flow graphs.

We talk about the basic concepts for state model of linear system, mean value, concave and convex function and maximum and minima. The primary components of a dynamic mathematical model correspond to the molecular species involved in the system (which are represented in the corresponding interaction diagram). The abundance of each species is assigned to a state variable within the model. The collection of all of these state variables is called the state of the system. It provides a complete description of the system's condition at any given time. The model's dynamic behavior is the time-course for the collection of state variables. Besides variables of state, models also include parameters, whose values are fixed. Model parameters characterize interactions among system components and with the environment.

1.2 Control System:

Definition (1.2.1)

A system is an arrangement, set or collection of things connected or related in such a manner as to form an entirety or whole.

A system is an arrangement of physical components connected or related in such a manner as to from and/or act entire unit.

The word control is usually taken to mean regulate, direct or command.

Definition (1.2.2)

A control system is an arrangement of physical components connected or relate in such a manner to command direct, or regulate itself or anther system.

In most abs track sense it is possible to consider every physical object a control system.

Examples of control systems:

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

The **input** is the stimulus excitation or command applied to a control system typically from an external energy source usually in order to produce a specified response from the control system.

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.

Input and output can have many different forms. Input, for example, may be physical variables or more abstract quantities such as reference set point, or desired values for the output of the control system.

The purpose of the control system usually identifies or defines the output and input. If the output and input are given it is possible to identify delineate or define the nature of the system components.

Control system may have more than one input or output. Often all input and output are well defining by the system description. But sometimes they are not. For example an atmospheric electrical storm may intermittently interfere with radio reception, producing an unwanted output from a loudspeaker in the form of static. This “noise” output is part of the total output as define above, but for the purpose of simply identifying a system, spurious input producing undesirable output are not normally considered as inputs and outputs when the system is examined in detail.

The terms input and output also may be used in the description of any type of system, whether or not it is a control system, and a control system may be part of a larger system, in which case it is called a subsystem or control subsystem, and its inputs and outputs may then internal variable of the larger system.

Example (1.2.1)

An electric switch is manufacture control system, controlling the flew of electricity. By definition the apparatus or person flipping the switch is not a part of this control system.

Flipping the switch on off may be considered as the input that is, the input can be in one of two states, on or off. The output is flow or no flow (two states) of electricity. The electric is one of the most rudimentary control systems.

1.3 Open Loop and Closed Loop Control Systems:

Control systems are classified into two general categories:

Open loop and closed loop system the distinction is determined by the control action, that quantity responsible for activating the system to produce the output.

The term control action is classical in the control system literature, but word action in this expression does not always directly imply change, motion, or activity for example, the control action in a system designed to have an object hit a target is usually the distance between the object and target distance, as such is not an action, but action (motion) is implied have, because the goal of such a control system is to reduce this distance to zero.

Definition (1.3.1)

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

Definition (1.3.2) a closed loop control system is one in which the control action is somehow dependent on the output.

Two outstanding features of open loop control system are:

1- Their ability to perform accurately is determined by their calibration. To calibrate means to establish or reestablish the input- output relation to obtain desired system accuracy.

2- They are not usually troubled with problems of instability a concept to be subsequently discussed in detail.

Definition (1.3.3)

Feedback is that property of closed loop system which permits the output (or some other controller) to be compared witch the input to the system (or an input to some other internally situated component or subsystem) so that the appropriate control action may be formed as some function of the output and input.

More generally, feedback is said to exist in a system when a closed sequence of cause –and- effect relations exists between system variable.

Characteristics of feedback

The presence of feedback typically imparts the following to a system.

1- Increased accuracy.

2- Tendency toward oscillation or in stability.

3- Reduced sensitive of the ratio of output to input to variation in system parameters and other Characteristics E- Reduced effects of external disturbances or noise.

4- Increased band width

The bandwidth of a system is a frequency response measure of how well the system responds to (or filters) variation or frequencies the input signal.

1.4 Analog and Digital Control Systems:

The signals in a control system are typically function of some independent variable, usually time, denoted.

Definition (1.4.1)

A signal dependent on a continuous of variable of t is called a continuous data signal or (less frequently) an analog signal.

Definition (1.4.2)

A signal defined at, or of interest at, only discrete (distinct) instants of the independent variable t (up on which it depends) is called a discrete time- a discrete data, a sampled data, or digital signal.

We remark that is a somewhat more specialized term, particularly in other contexts is as a synonym here because it is the convention in the control system literature.

Control system can be classified according to the types of signals they process: Continuous- time(analog), discrete- time (digital) or a combination of both (hybrid).

Definition (1.4.3)

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

Definition (1.4.4)

Discrete- time control system also called discrete data control systems, or sampled data control systems, have discrete – time signals or component at one or more points is the system, we note that discrete- time control system can have continuous- time signals , that is, they can be hybrid.

The distinguishing factor is that discrete- time or digital control system must include at least one discrete- data signals. Also, digital control system, particularly of sampled- data type, often have both open- loop and closed loop modes of operation.

1.5 Control System Models or Representation:

To solve control systems problem, we must put the specification or description of the system configuration and its components into a form amenable to analysis or design.

Three basic representations (model) of components and system are used extensively in the study of control systems:

1- Mathematical models in the form of differential equation, difference equation, and/or other mathematical relation for example, Laplace transforms.

2- Block diagrams.

3- Signal flow graphs.

1.5.1 Transfer Function:

In control theory , function called transfer function are commonly used to characterize the input – output relationships of components or systems that can be described by linear, time invariant ,differential equation, we begin by define transfer function and follow with a derivation of the transfer function of a differential equation system.

The transfer function of a linear, 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 (driving function) under the assumption that all initial conditions are zero.

Consider the linear time – invariant system defined by the following differential equation:

$$a_\circ y^n + a_1 y^{n-1} + \cdots + a_{n-1} y^\circ + a_n = b_\circ x^m + b_1 x^{m-1} + \cdots + b_{m-1} x^\circ + b_m$$

Where **y** is the output of the system and x is the input. The transfer function of this system is ratio of the Laplace transformed output to the Laplace transformation input when all initial condition are zero, or

Transfer function = $\frac{L[output]}{L[input]}\ zero\ inital\ condition$

1.5.2 Block Diagram:

A block diagram is a 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.

Figure 1.1

The operation of addition and subtraction have a special representation the block becomes a small circle called a summing point with the appropriate plus or minus sign associated with the arrows entering the circle the output is the algebraic sum of the inputs any number enter a summing point for examples

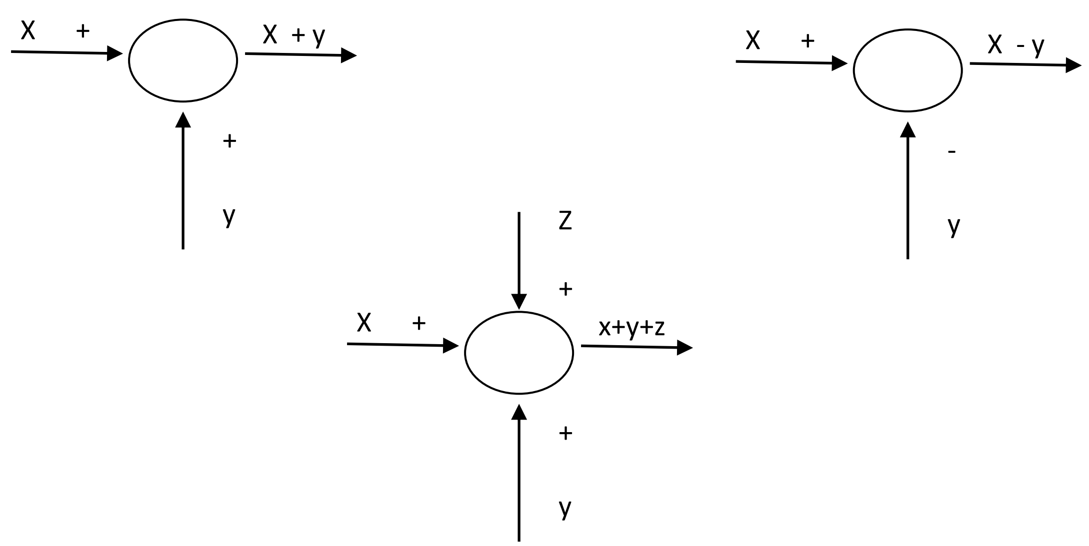


Figure 1.2

The basic configuration of a simple closed-loop (feedback) control system with a single input and a single output (abbreviated SISO) is illustrated in Figure 1.3 for system with continuous signals only

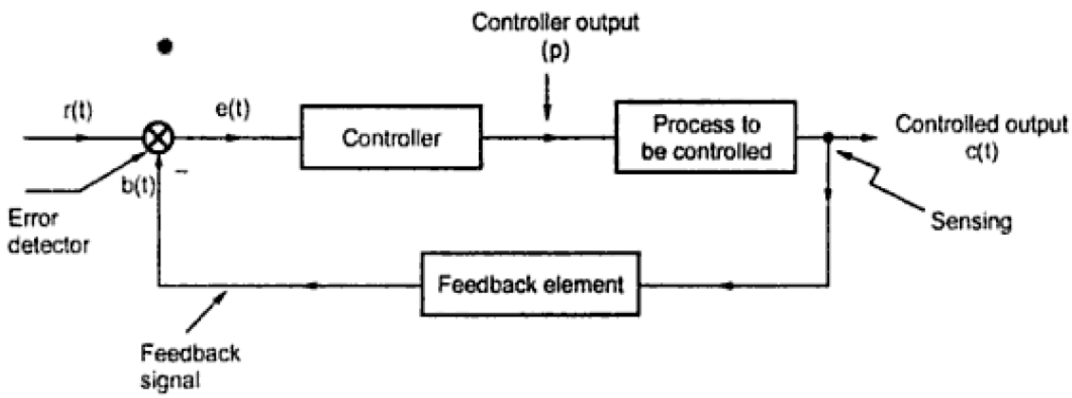


Figure 1.3 Basic Control systems

The plant (or **process** or **controlled system**) is the system, subsystem process or object controlled by the feedback control system.

The **controlled output C** is the output variable of the plant. Under the control of the feedback control system.

The **forward path** is the transmission path from the summing point to the controlled output C.

The **feed forward (control) elements** are the components of the forward path that generate the control signal u or m applied to the plant. Note feed forward elements typically include controller(s), compensator(s) or equalization and/or amplifiers.

The **control signal u (or manipulated variable m)** is output signal of the feed forward elements applied as input to the plant.

The **feedback bath** is the transmission path from the controlled output c back to the summing point.

The **feedback elements** establish the function relationship between the controlled output c and the primary feedback signal B. note feedback element typically include sensor of the controlled output, **c**ompensators and/or controller elements.

The Reference input R is one external signal applied to the feedback control system. Usually at the first summing point, in order to command a specified action of the plant.it usually represents ideal (or desired) plant output behavior.

The **primary feedback signal B** is function of the controller output C algebraically summed with the reference input R to obtain the actuating (error)signal E , that is R + B=E

Note: an open-loop system has no primary feedback signal.

The **actuating**(or **Error) Signal** is the reference input signal R puls or minus the primary feedback signal **B**. the control action is generated by the actuating (error)signal in a feedback control system(see definitions open – loop and closed loop control system) note: in an open-loop system, which has feedback the actuating signal is equal to **R.**

Negative feedback: means the summing point is asubtractor, that is E=R-B

Positive feedback: means the summing point is an adder, that is E=R+B

1.5.3 Signal flow graphs:

Block diagrams are adequate for the representation of the interrelationships of controlled and input variables. However, for a system with reasonably complex interrelationships, the block diagram reduction procedure is cumbersome and often quite difficult to complete. An alternative method for determining the relationship between system variables has been developed by Mason. The advantage of the line path method, called the signal-flow graph method, is the availability of a flow graph gain formula, which provides the relation between system variables without requiring any reduction procedure or manipulation of the flow graph. The transition from a block diagram representation to a directed line segment representation is easy to accomplish by reconsidering the systems of the previous section.

A **signal-flow graph** is a diagram consisting of nodes that are connected by several directed branches and is a graphical representation of a set of linear relations. Signal-flow graphs are particularly useful for feedback control systems because feedback theory is primarily concerned with the flow and processing of signals in systems. The basic element of a signal-flow graph is a unidirectional path segment called a **branch,** which relates the dependency of an input and an output variable in a manner equivalent to a block of a block diagram. Therefore, the branch relating the output $\theta(s)$ of a DC motor to the field voltage $V_f(s)$ is shown Figure 1.4.The input and output points or junctions are called **nodes**, is shown in Figure 1.5. The relation between each variable is written next to the directional arrow. All branches leaving a node will pass the nodal signal to the output node of each branch (unidirectional).The summation of all signals entering a node is equal to the node variable. A **path** is a branch or a continuous sequence of branches that can be traversed from one signal (node) to another signal (node). A **loop** is a closed path that originates and terminates on the same node, with no node being met twice along the path. Two loops are said to be **nontouching** if they do not have a common node. Two touching loops share one or more common nodes. Therefore, considering Figure 1.5 again, we obtain

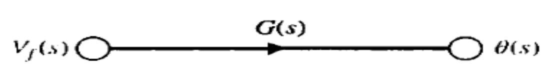


Figure 1.4 signal - flow graph of the DC motor

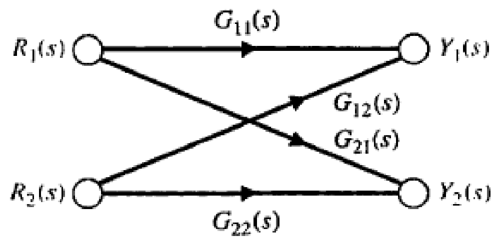


Figure 1.5 signal – flow graph of interconnected system

$$y_1(s) = G_{11}(s)R_1(s) + G_{12}(s)R_2(s) \tag{1.5.1}$$

and

$$y_2(s) = G_{21}(s)R_1(s) + G_{22}(s)R_2(s) \tag{1.5.2}$$

The flow graph is simply a pictorial method of writing a system of algebraic equations that indicates the interdependencies of the variables. As another example, consider the following set of simultaneous algebraic equations

$$a_{11}x_1 + a_{12}x_2 + r_1 = x_1 \tag{1.5.3}$$

$$a_{21}x_1 + a_{22}x_2 + r_2 = x_2 \tag{1.5.4}$$

The two input variables are r_1 and r_2, and the output variables are x_1 and x_2. A signal- flow graph representing equations (1.5.3) and (1.5.4) is shown in Figure1.6.

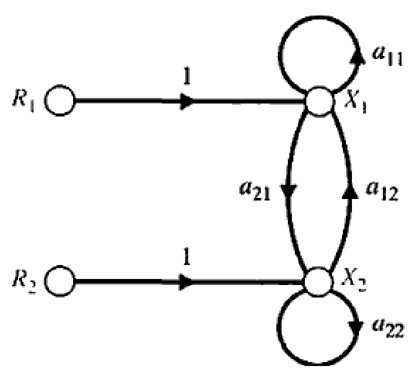


Figure 1.6 signal- Flow graphs of two algebraic equations

Equations (1.3) and (1.4) may be rewritten as

$$x_1(1 - a_{11}) + x_2(-a_{12}) = r_1 \tag{1.5.5}$$

$$x_1(-a_{21}) + x_2(1 - a_{22}) = r_2 \tag{1.5.6}$$

The simultaneous solution of equations (1.5.5) and (1.5.6) using Cramer's rule results in the solutions

$$x_1 = \frac{(1-a_{22})r_1 + a_{12}r_2}{(1-a_{11})(1-a_{22}) - a_{12}a_{21}} = \frac{1-a_{22}}{\Delta} r_1 + \frac{a_{12}}{\Delta} r_{2,} \tag{1.5.7}$$

$$x_2 = \frac{(1-a_{11})r_2 + a_{21}r_1}{(1-a_{11})(1-a_{22}) - a_{12}a_{21}} = \frac{1-a_{11}}{\Delta} r_1 + \frac{a_{21}}{\Delta} r_1 \tag{1.5.8}$$

The denominator of the solution is the determinant A of the set of equations and is rewritten as

$$\Delta = (1 - a_{11})(1 - a_{22}) - a_{12}a_{21} = 1 - a_{11} - a_{22} + a_{11}a_{22} - a_{12}a_{21} \tag{1.5.9}$$

In this case, the denominator is equal to 1 minus each self-loop a_{11}, a_{22}, and $a_{12}a_{21}$ plus the product of the two no touching loops a_{11} and a_{22}. The loops a_{22} and $a_{21}a_{12}$ are touching, as are a_{11} and $a_{21}a_{12}$.

The numerator for x_1 with the input r_1 is 1 times1 — a_{22} , which is the value of Δ excluding terms that touch the path 1 from $r_1\ to\ x_1$. Therefore the numerator from r_2 to x_1 is simply a_{12} because the path through a_{12} touches all the loops. The numerator for x_2 is symmetrical to that of x_2, In general, the linear dependence T_{ij} between the independent variable x_i (often called the input variable) and a dependent variable x_j is given by Mason's signal-flow gain formula

$$T_{ij} = \frac{\sum p_{ijk}\Delta_{ijk}}{\Delta} \tag{1.5.10}$$

p_{ijk} = gain of *kth* path from variable x_i to variable x_j

Δ = determinant of the graph,

Δ_{ijk}= cofactor of the path P_{ijk}, and the summation is taken over all possible *k* paths from x_i to x_j. The path gain or transmittance P_{ijk} is defined as the product of the gains of the branches of the path, traversed in the direction of the arrows with no node encountered more than once. The cofactor Δ_{ijk} is the determinant with the loops touching the *kth* path removed. The determinant Δ is

$$\Delta = 1 - \sum_{n=1}^{N} L_n + \sum_{\substack{n,m \\ nontouching}} L_n L_m + \sum_{\substack{n,m,p \\ nontouching}} L_n L_m L_p + \cdots, \qquad (1.5.11)$$

Where L_q equals the value of the *qth*. Loop transmittance. Therefore the rule for evaluating Δ in terms of loops $L_1, L_2, L_3, \dots \dots L_n$ is

Δ = 1 — (sum of all different loop gains) + (sum of the gain products of all combinations of two nontouching loops) — (sum of the gain products of all combinations of three nontouching loops) +………..

The gain formula is often used to relate the output variable *Y(s)* to the input Variable *R(s)* and is given in somewhat simplified form as

$$T = \frac{\sum_k p_k \Delta_k}{\Delta} \qquad (1.5.12)$$

Where

$$T(s) = y(s)/R(s).$$

1.6 Properties of Controller:

Consider a control system shown in the Figure 1.3 which includes a controller.

1.6.1Error The Error detectors compare the feedback signal $b(t)$with the reference input $r(t)$ to generate an error.

$$\therefore e(t) = b(t) - r(t)$$

This gives absolute indication of an error the range of the measured variable $b(t)$

Thus span $= b_{max} - b_{min}$

The Hence error can be expressed as

$$e_P = \frac{r-b}{b_{max} - b_{min}} \times 100$$

Where $e_P \equiv$ error as % of span

Example (1.6.1)

The range of measured variable for ascertains control system is 2 mv to 12 mv and a set point 7 mv. Find the error as percent of span when the measured variable is 6.5 mv.

Solution

$$b_{mix} = 12\,mv \quad ,b_{min} = 2mv \quad ,b = 6.5\,mv \quad r = 7\,mv$$

$e_P = \frac{r-b}{b_{max}-b_{min}} \times 100 = \frac{7-6.5}{12-2} \times 100 = 5\%$

1.6.2 Variable Range

In practical systems, the controlled variable has a range of values within which the control is required to be maintained. This range specified as the maximum and minimum values allowed for the controlled variable. It can be specified as some nominal values and plus minus tolerance allowed about this value such range is important for the design of controllers.

1.6.3 Controller Output Range

Similar to the controller variable a range is associated with a controlled output variable and minimum values. But often the controller output is expressed as a percentage where minimum controller output is 0% and maximum controller out is 100% but 0% controller output does not mean, zero output.

For example it is necessary of the system that a steam flow corresponding to $\left(\frac{1}{4}\right)^{th}$ opening of the values should be minimum.

Thus 0% controller output in such case corresponds to the $\left(\frac{1}{4}\right)^{th}$ opening of the value.

The controller output as a percent of full scale when the output changes within the specified range is expressed as

$$\boldsymbol{p = \frac{U - U_{min}}{U_{max} - U_{min}} \times 100}$$

Where

p ≡Controller output as a percent of full scale

U ≡Value of the output

U_{max} ≡Maximum value of controlling variable

U_{min} ≡Minimum value of controlling variable

1.6.4 Control Lag

The control system can have large as associate with it, the control lag is the time required by the process and controller loop to make the necessary changes to obtain the output at its set point the control lag t must be compared with the process lag while designing the controllers for example. In a process value is required to be open or closed for corresponding the output variable physically the of opening. Or closing of the value is very slow and is the part of the process lag. In such a case there is no point in designing a fast controller than the process lag.

1.6.5 Dead Zone

Many a times a dead zone is associated with a process control loop the time corresponding to dead zone is called dead time. The elapsed between the instant when error occurs and instant when first corrective action occurs is called dead time .Nothing happens the error occurs this part is also called dead hand the effect of such dead time must be considered while the design of the controllers.

1.7 Classification of Controllers:

The classification of the controllers is based on the response of the controllers and mode of response of the controller for example. In a simple temperature control of a room the heater is to be controlled it should be switched on or off by the controller when temperature crosses its set point. Such an operation and the mode of operation is called discontinuous mode of controller but in some process control systems simple on/off decision is not sufficient for example controlling the steam slow by opening or closing the value in such case a smooth opening or closing of value is necessary. The controller in such a case is said to be operating in a continuous mode thus the controllers are basically classified ad discontinuous controllers. The discontinuous mode controllers are further classified as ON, OFF controllers and multi position controllers. The continuous mode controllers are further classified as derivative controllers. Some continuous mode controllers can be combined to obtain composite controller mode.

For example of such composite controllers are PI, PD and PID controllers. The most of the controllers are placed in the forward path of control system. But in some cases input to the controller is controlled though a feedback path. The example of such a controller is rate feedback controller.

1.8 Continuous Controller Mode

In the discontinuous controller mode the output of the controller is discontinuous and not smoothly varying. But in the continuous controller output smoothly proportional of the error or proportional to some form of the error.

Depending upon which form of the error is used as the input to the controller to product the continuous controller output these controllers are classified as Proportional control mode, Integral control mode and derivative control mode

1.9 State Model of Linear System:

1-State:

The state of adynamic system is defined a minimal set of variables such that the knowledge of these variables at $t = t_0$ together with the knowledge of the input for $t \geq t_0$,completely determines the behaviour of the system for $t > t_0$.

2-State variables:

The variable involved in determining the state of adynamic system *x(t)* are called the state variable. $x_1(t), x_2(t), \dots \dots x_n(t)$ are nothing but the state variables these are normaly the energy storing elemnts contained in the system.

3-State vector:

The "n" state variable necessary to describe the complete behaviour of the system can be considered as "n"component of avector X(t) called the state vector at time "t". the state vector X(t) is vector sum of all the state variable.

4-State space:

The space whose co-ordinate exes are nothing but the "n" state variables with time as the implicit variable is called the state space.

5-State trajectory:

It is the locus of the tips of the state vectors,with as the implicit variable.Consider multiple input multiple output (MIMO) nth order system as shown in the figure

Number of inputs = m Number of outputs = p

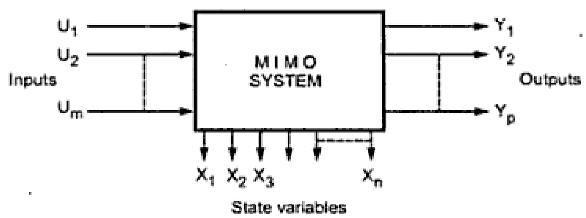


Figure 1.7

$$U(t) = \begin{bmatrix} U_1(t) \\ U_2(t) \\ : \\ : \\ U_m(t) \end{bmatrix}, \quad X(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \\ : \\ : \\ X_n(t) \end{bmatrix}, \quad Y(t) = \begin{bmatrix} Y_1(t) \\ y_2(t) \\ : \\ : \\ y_p(t) \end{bmatrix}$$

All are Colum vector having order $m \times 1, n \times 1$ *and* $p \times 1$ respectively.For such a system, the state variable representation can be arranged in the form of "n" first order differential equation

$$\frac{dX_1}{dt} = x_1^{\bullet} = f_1(X_1, X_2, \dots \dots, X_n, U_1, U_2, \dots \dots, U_m)$$

$$\frac{dX_2}{dt} = x_2^{\bullet} = f_2(X_1, X_2, \dots \dots, X_n, U_1, U_2, \dots \dots, U_m)$$

$$: \qquad\qquad\qquad\qquad :$$

$$\frac{dX_n}{dt} = x_n^{\bullet} = f_n(X_1, X_2, \dots \dots, X_n, U_1, U_2, \dots \dots, U_m)$$

Where $f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$ is the functional operator.

Integrating the above equation

$$X_i(t) = X_i(\mathrm{t_0}) + \int_{\mathrm{t_0}}^{\mathrm{t}} f_i(X_1, \mathrm{X_2}, \dots \dots \mathrm{X_n}, \mathrm{U_1}, \mathrm{U_2}, \dots \dots, \mathrm{U_n})\mathrm{dt}$$

Where i=1,2,……..,n

Thus "n" state variable and hence state vector at any time "t" can be determined uniquely.

Any "n" dimensional time invariant system has state equations in the functional form as

$$\dot{\mathrm{X}}(t) = f(X, U)$$

While output of such system are dependent on the state of system and instaneous input.

Functional output equation can be written as,

$$y(t) = g(X, U)$$

Where "g" is functional operator

For time variant system, the same equation can be written as,

$$\dot{\mathrm{X}}(t) = f(X, U) \text{……… state equation}$$

$$y(t) = g(X, U) \text{.......... Output equation}$$

The functional equation can be expressed interms of linear combination of system state and the input as,

$$\dot{X}_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1n}X_n + b_{11}U_1 + b_{12}U_2 + \cdots + b_{1m}U_m$$

$$\dot{X}_2 = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2n}X_n + b_{21}U_1 + b_{22}U_2 + \cdots + b_{2m}U_m$$

: :

$$\dot{X}_n = a_{n1}X_1 + a_{n2}X_2 + \cdots + a_{nn}X_n + b_{n1}U_1 + b_{n2}U_2 + \cdots + b_{nm}U_m$$

For the linear time invariant systems, the coefficients a_{ij} *and* b_{ij} are constant. Thus all the equation can be written in vector matrix form as,

$$\dot{X}(t) = A\,X(t) + B\,U(t)$$

Where

X(t) = state vector matrix of order $n \times 1$.

U(t) = input vector matrix order $n \times 1$.

A = system matrix or evolution matrix of order $n \times n$.

B = Input matrix or control matrix of order $n \times m$.

Similarly the output variable at time t can be expressed as the linear combination of the input variable and state variable at time t as,

$$y_1(t) = c_{11}X_1(t) + \cdots + c_{1n}X_n(t) + d_{11}U_1(t) + \cdots + d_{1m}(t)U_m$$

$$y_2(t) = c_{21}X_1(t) + \cdots + c_{2n}X_n(t) + d_{21}U_1(t) + \cdots + d_{2m}(t)U_m$$

: :

$$y_p(t) = c_{p1}X_1(t) + \cdots + c_{pn}X_n(t) + d_{p1}U_1(t) + \cdots + d_{pm}(t)U_m$$

For the linear time invariant system the coefficient c_{ij} and d_{ij} are constants.

Thus all the output equation can be written in vector matrix form as,

$$Y(t) = C\,X(t) + D\,U(t)$$

Where

Y(t) = output vector matrix of order $p \times 1$.

C = output matrix or observation matrix of order $p \times n$.

D = direct transmission matrix of order $p \times m$.

The two vector equation together is called the state model of the linear system.

$$\dot{X}(t) = A\,X(t) + B\,U(t) \text{......state equation}$$

$$Y(t) = C\,X(t) + D\,U(t) \text{......output equation}$$

This is state model of system.

For linear time – variant systems, the matrices A,B,C and D are also time dependent.

Thus,

$$\dot{X}(t) = A(t)X(t) + B(t)U(t)$$

$$Y(t) = C(t)X(t) + D(t)U(t)$$

1.9.1 State model of single input single output system:

Consider a single input single output system. i.e m = 1 and p = 1. But its order is "*n*" hence *n* state variable are required to define state of the system. In such a case, the state model is

$$\dot{X}(t) = A\,X(t) + B\,U(t)$$

$$Y(t) = C\,X(t) + D\,U(t)$$

Where A = $n \times n$ matrix , B = $n \times 1$ matrix

C = $1 \times n$ matrix , D constant

and U(t) = single scalar input variable.

1.10 Basic Features of Dynamic Mathematical Models

1.10.1 State variables and model parameters

The primary components of a dynamic mathematical model correspond to the molecular species involved in the system (which are represented in the corresponding interaction diagram). The abundance of each species is assigned to a state variable within the model. The collection of all of these state variables is called the state of the system. It provides a complete description of the system's condition at any given time. The model's dynamic behaviour is the time-course for the collection of state variables. Besides variables of state, models also include parameters, whose values are fixed. Model parameters characterize interactions among system components and with the environment. Examples of model parameters are: association constants, maximal expression rates, degradation rates, and buffered molecular concentrations. A change in the value of a model parameter corresponds to a change in an environmental conditions or in the system itself. Consequently, model parameter are typically held at constant values during simulation; these values can be varied to explore system behaviour under perturbations or in altered environments (e.g. under different experimental conditions). For any given model, the distinction between state variables and model parameters is clearcut. However, this distinction depends on the model's context and on the time-scale over which simulations run.

1.10.2 Steady-state behaviour and transient behaviour

Simulations of dynamic models represent time-varying system behaviour. Models of biological processes almost always arrive, in the long run, at steady behaviours. Most commonly, models exhibit a persistent operating state, called a steady state; some systems display sustained oscillations. The time-course that leads from the initial state to the long-time (or asymptotic) behavior is referred to as the transient. In some cases, we will focus on transient behaviour, as it reflects the immediate response of a system to perturbation. In other cases, our analysis will concern only the steady-state behavior, as it reflects the prevailing condition of the system over significant stretches of time.

1.10.3 Linearity and nonlinearity

A relationship is called linear if it is a direct proportionality. For example, the variables x and y are linearly related by the equation $x = k\,y$, where k is a fixed constant. Linearity allows for effortless extrapolation: a doubling of x leads to a doubling of y, regardless of their values. Linear relationships involving more than two variables are similarly transparent, e.g. $x = k_1 y + k_2 z$. A dynamic mathematical model is called linear if all interactions among its components are linear relationships. This is a highly restrictive condition, and consequently linear models display only a limited range of behaviours. Any relationship that is not linear is referred to (unsurprisingly) as nonlinear. Nonlinear relations need not follow any specific pattern, and so are difficult to address with any generality.

The nonlinearities that appear most often in biochemical and genetic interactions are saturations, in which one variable increases with another at a diminishing rate, so that the dependent variable tends to a limiting, or asymptotic value. Two kinds of saturating relationships that we will encounter repeatedly are shown in

Figure 1.8. Panel A shows a hyperbolic saturation, in which the rate of increase of y declines continuously as the value of x increases. Panel B shows a sigmoidal saturation, in which y initially grows very slowly with x, then passes through a phase of rapid growth before saturating as the rate of growth drops.

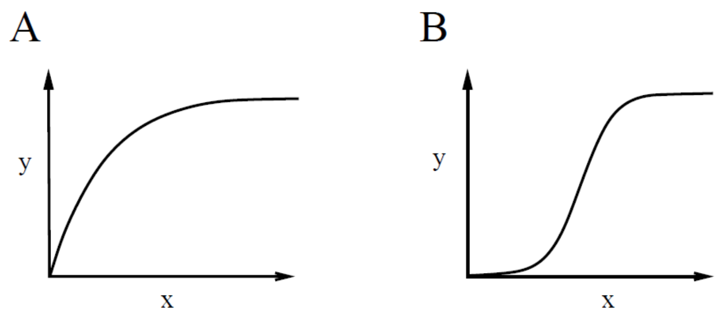


Figure 1.8 Common nonlinear relationships in cell biological processes. A. Hyperbolic saturation. As x increases, y also increases, but at an ever-diminishing rate. The value of y thus approaches a limiting, or asymptotic, value. B. Sigmoidal nonlinearity. The values of y show a slow rate of increase for small values of x, followed by a rapid 'switch-like' rise toward the limiting value.

1.10.4 Global and local behaviour

Nonlinear dynamic models exhibit a wide range of behaviours. In most cases, a detailed analysis of the overall, global, behaviour of such models would be overwhelming. Instead, attention can be focused on specific aspects of system behaviour. In particular, by limiting our attention to the behaviour near particular operating points, we can take advantage of the fact that, over small domains, nonlinearities can always be approximated by linear relationships (e.g. a tangent line approximation to a curve). This local approximation allows one to apply linear analysis tools in this limited purview. Intuition might suggest that this approach is too handicapped to be of much use. However, the global behaviour of systems is often tightly constrained by their behavior around a handful of nominal operating points; local analysis at these points can then provide comprehensive insight into global behaviour. Local approximations are of particular use in biological modelling because self-regulating (e.g. homeostatic) systems spend much of their time operating around specific nominal conditions.

1.10.5 Deterministic models and stochastic models

The notion of determinism reproducibility of behaviouris a foundation for much of scientific investigation. A mathematical model is called deterministic if its behaviour is exactly reproducible. Although the behaviour of a deterministic model is dependent on a specified set of conditions, no other forces have any influence, so that repeated simulations under the same conditions are always in perfect agreement. (To make an experimental analogy, they are perfect replicates.)

In contrast, stochastic models allow for randomness in their behaviour. The behaviour of a stochastic model is influenced both by specified conditions and by unpredictable forces. Repeated stochastic simulations thus yield distinct samples of system behaviour.

Deterministic models are far more tractable than stochastic models, for both simulation and model analysis. Our focus will be on deterministic models.

1.11 Mean-Value Theorems:

If $f(x)$ is continuous on an interval$a \leq x \leq b$, then there exists a number $\bar{x}$, such that

$$\int_a^b f(\bar{x})dx = f(\bar{x})(b-a), \quad where \;\; a < \bar{x} < b. \tag{1.11.1}$$

Recalling Fundamental Theorem of Integral Calculus

$\int_a^b f(x)dx = \int_a^b F'(x)dx = F(b) - F(a)$ and $\frac{df}{dx} = F'(x) = f(x)$, one can rewrite (1.11.1)

$$\frac{[F(b)-F(a)]}{b-a} = F'(\bar{x}) \quad for\ some \quad a < \bar{x} < b. \tag{1.11.2}$$

Expression (1.11.2) is the mean-value theorem of differential calculus, while (1.11.1) is the mean-value theorem of integral calculus.

Mean-value theorem (1.11.2) has an interesting geometric interpretation. The left side is the slope of the hypotenuse of the right triangle ABC in Figure 1.9, with base of length b - a and height F(b) – F(a). Then (1.11.2) indicates that there is at least one point, say *x*, at which the curve F(x) has the same slope as the hypotenuse AB of the triangle joining the end points.

In (1.11.2), *x* is strictly between a and b. The mean-value theorem can be used to show that

$$if\ F(b) - F(a) = F(a)(b-a), then\ either \qquad (1.11.3)$$

a = b or else F´´(r) = 0 for some r strictly between a and b

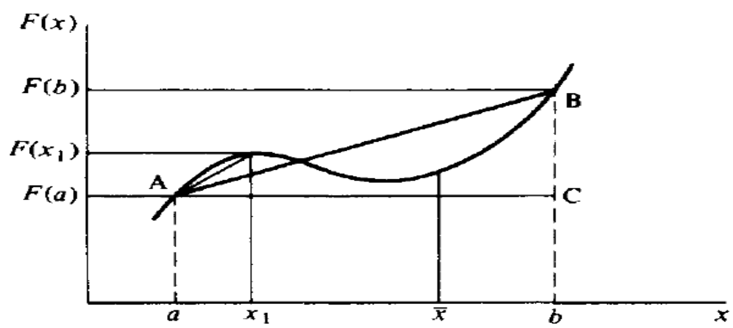


Figure 1.9

This is clear from the graph. Algebraically, if a = b, the result is immediate. So suppose a < b. By the mean-value theorem, there is a number q, a < q < b, such that

$$F(b) - F(a) = (b-a)F'(q)$$

Combining this with the hypothesis of (3) gives

$$[F'(a) - F'(q)](b-a) = 0$$

Again, by the mean-value theorem, there is a number r, a < r < q, such that

$$F'(a) - F'(q) = (a-q)F''(r)$$

and therefore

$$(a-q)(b-a)F''(r)=0.$$

Since a < q < b, F''(r) = 0, as was to be shown. The case of a > b is analogous.

Any curve can be approximated arbitrarily well in the neighborhood of a point by a polynomial of sufficiently high order, according to

Taylor's Theorem. If the function *f(x)* and its first *n - 1* derivatives are continuous in the interval $a < x < b$ and its nth derivative for each x, $a < x < b$ exists, then for $a < x_0 < b$, the Taylor series expansion of *f(x)* about the point x_0 is

$$f(x)=f(x_0)+\frac{\sum_{i=1}^{n}(x-x_0)^i f^{(i)}(x_i)}{i!}+R_n \tag{1.11.4}$$

$$\text{Where } f^i \equiv \frac{d^i f}{dx^i} \; and \; R_n=\frac{(x-x_0)^n f^n(\bar{x})}{n!} \quad for\ some\ \bar{x}, a<\bar{x}<b.$$

This Taylor series expansion of *f(x)* extends (2). We call R_{n+1} the remainder and (n - l)th the order of the approximation obtained by deleting R_{n+1} The special case (2) gives the zero order approximation and $R_1 = (b — a)f'(x)$. The first order approximation is

$f(x) = f(x_0) + (x - x_0)f'(x_0)$, a straight line through $(x_0, f(x_0))$ with slope $f'(x_0)$. The second order approximation gives a quadratic approximation to the curve at the point x_0 deleting the remainder in (4) gives an nth degree polynomial approximation to *f(x)*.

The mean-value theorem for a function of two variables, *F(x, y)*, is developed as follows. For points *(x, y)* and (x_0, y_0) in the domain of F, let $h = x — x_0$ and $k = y — y_0$. Define the function

$$f(t)=F(x_0+th, y_0+tk) \tag{1.11.5}$$

of the single variable *t* on the interval $0 \le t \le 1$. Then

$$f(1)-f(0)=F(x,y)-F(x_0,y_0)$$

But from (2),

$$f(1)-f(0)=f'(\bar{t}) \quad for\ some\ \bar{t}, 0<\bar{t}<1. \tag{1.11.6}$$

Differentiating (5) with respect to *t* gives

$$f'(t) = hF_x + kF_y \tag{1.11.7}$$

where F_x and F_y are evaluated at $(x_0 + th, y_0 + tk)$. Combining (5)-(7) gives the mean-value theorem for a function of two variables:

$$F(x,y) - F(x_0, y_0) = (x - x_0)F_x(\bar{x}, \bar{y}) + (y - y_0)F_y(\bar{x}, \bar{y}) \tag{1.11.8}$$

For some $\bar{x}$,$\bar{y}$ where $x_0 < \bar{x} < x \; and \; y_0 < \bar{y} < y$.

The Taylor series expansion or generalized mean-value theorem (3) can likewise be extended to a function of two variables. Let $h = x - x_0$ and $k = y - y_0$. Then

$$F(x,y) = F + (hF_x + kF_y) + R_2 \tag{1.11.9}$$

Where $R_2 = h^2F_{xx}/2 + hkF_{xy} + k^2F_{yy}/2$. On the right side of (9), F, F_x, and F_y are evaluated at x_0, y_0 and the second partial derivatives in R_2 are evaluated at some point between x, y and x_0, y_0. Expanding further, we get

$$F(x,y) = F + (hF_x + kF_y) + \left(\frac{h^2F_{xx}}{2} + hkF_{xy} + \frac{k^2F_{yy}}{2}\right) + R_3 \tag{1.11.10}$$

where F and its partial derivatives on the right side are all evaluated at $(x_0 > y_0)$ and R_3 is the remainder.

More generally, the first order Taylor series expansion of a function $F(x_1, \ldots, x_n)$ of n variables about the point $x^\circ = (x^\circ, \ldots, x^\circ)$ is

$$F(x_1, \ldots\ldots, x_n) = F + \sum_{i=1}^{n} h_i F_i + R_2 \tag{1.11.11}$$

Where

$$R_2 = \left(\frac{1}{2}\right)\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j F_{ij}, \quad h_i = x_i - x_i^0 \;\; i = 1, \ldots\ldots\ldots, n.$$

On the right side of (11), F and F_i, are evaluated at $x°$ and the F_{ij} are evaluated at some point between x and x^0. The second order expansion is

$$F(x_1, \dots \dots, x_n) = F + \sum_{i=1}^{n} h_i F_i + \left(\frac{1}{2}\right) \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j F_{ij} + R_3 \qquad (1.11.12)$$

where F and its first and second partial derivatives F_i, F_{ij} on the right are all evaluated at x^0, and R_3 is the remainder term. The series can be expanded to as many terms as desired.

1.12 Concave and Convex Functions

A function *f (x)* is said to be *concave* on the interval $a \le x \le b$ if for all $0 \le t \le 1$ and for any $a \le x_1 \le x_2 < b$

$$tf(x_1) + (1-t)f(x_2) \le f(tx_1 + (1-t)x_2) \qquad (1.12.1)$$

A weighted average of the values of the function at any two points is no greater than the value of the function at the same weighted average of the arguments. If the inequality in (1.12.1) is strong, then the function *f (x)* is said to be strictly concave. The chord joining any two points on the graph of a concave function is not above the graph. If the function is strictly concave, the chord is strictly below the graph. Linear functions are concave, but not strictly concave (see Figure 1.10).

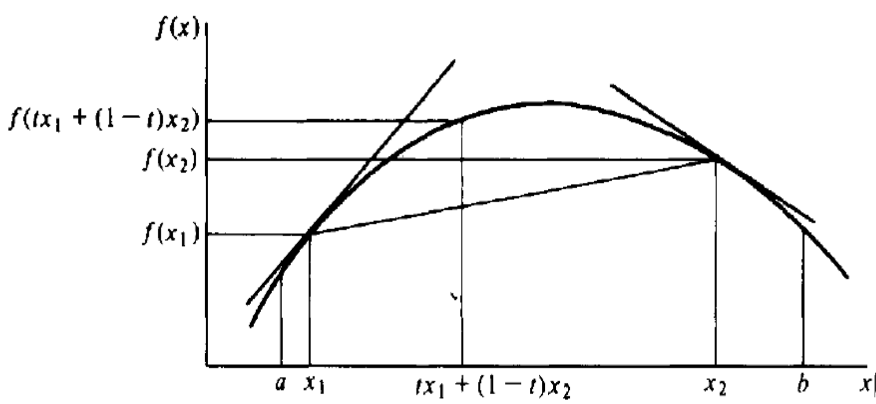


Figure 1.10

A function *f (x)* is said to be convex if the inequality in (1.12.1) is reversed. It is strictly convex if the reversed inequality is strong. For example, $f\ (\ x\ )\ =\ -x^2$ is a strictly concave function, whereas $f\ (\ x\ )\ =\ x^2$ is strictly convex. These functions illustrate the general principle that if *f (x)* is a concave function, then -*f(x)* is a convex function.

Concave functions have several important properties. First, a concave function *f (x)* is continuous on any open interval a < x < b on which it is defined. Second, if *f* is concave and differentiable,

$$(x_2 - x_1)f^{'}(x_1) \geq f(x_2) - f(x_1) \geq (x_2 - x_1)f^{'}(x_2) \tag{1.12.2}$$

The slope of the line joining two points on its graph is less than the slope of the tangent line at its left endpoint and greater than the slope of its tangent line on its right endpoint (see Figure 1.10). For a convex function, the inequalities in (1.12.2) are reversed. Third, the second derivative of a twice differentiable concave function is nonpositive. To see this, use (1.11.4) to write

$$f(x_2) - f(x_1) - (x_2 - x_1)f'(x_1) = (x_2 - x_1)^2 f''(\bar{x})/2 \tag{1.12.3}$$

for some $\bar{x}$ where $x_1 < \bar{x} < x_2$. But from (1.12.2) or Figure 1.10, it is apparent that the left side of (1.12.2) is nonpositive. Therefore, the right side of (1.12.3) must also be nonpositive. Since $(x_2 - x_1)^2$ is positive, *f''(x)* must be nonpositive. Further,

Since (3) holds for any$a \leq x_1 < x_2 \leq b$, *f''(x)* < 0 for all *a* < *x* < *b*. Similarly, the second derivative of a convex function is nonnegative.

The condition *f''(x)* < 0 for all *x* in the domain is both necessary and sufficient for a twice differentiable function *f* to be concave. If *f''(x) < 0*, then *f* is strictly concave. The converse of this, however, is not true. The function

$f\ (x)\ =\ -x^4$ appears to be strictly concave when plotted but, *f''(O) = 0*. Despite this possibility, we assume that $f''(x)\ <\ 0$ for strictly concave functions. The reader should keep in mind the possibility of exceptional points.

The definition of a *concave function of two variables f(x, y)* is a direct extension of the definition (1):

$$tf(x_1,y_1) + (1-t)f(x_2,y_2) \leq f(tx_1 + (1-t)x_2, ty_1 + (1-t)y_2\ ) \qquad (1.12.4)$$

for all 0 < *t* < 1 and any pair of points$(x_1,y_1),(x_2,y_2)$ in the domain of ***f*.** The extension of (2) is

$$f(x_2,y_2) - f(x_1,y_1) \leq (x_2-x_1)f_x(x_1,y_1) + (y_2-y_1)f_y(x_1,y_1) \qquad (1.12.5)$$

which holds for a differentiable concave function of two variables at any pair of points in its domain. Inequalities (1.12.4) and (1.12.5) are reversed for a convex function. to find the analog of the sign condition on the second derivative, write the counterpart to (1.12.3), employing (1.11..9).

Let $h = x_2 - x_1$and $k = y_2 - y_1$. Then,

$$f(x_2,y_2) - f(x_1,y_1) - hf_x(x_1,y_1) - kf_y(x_1,y_1)$$
$$= h^2 f_{xx}/2 + hkf_{xy} + k^2 f_{yy}/2 \qquad (1.12.6)$$

where the second partial derivatives are evaluated at an appropriate point $\bar{x}, \bar{y}$ between (x_1,y_1) and (x_2,y_2). The left side of (6) is nonpositive by (5), and therefore, the *quadratic form* on the right side must be nonpositive as well. (A *quadratic form* is a function of the form $f(x_1,\ldots,x_n) = \sum_{j=1}^{n}\sum_{n=1}^{n} a_{ij}x_ix_j$ where $a_{ij} = a_{ji},\quad i,j = 1,2,\ldots\ldots\ldots,n$. if $f_{xx} \neq 0$,add and subtract $k^2 f_{xy}^2/f_{xx}$ and collect terms to write the right side of (1.12.6) equivalently

$$f_{xx}[\left(h + \frac{kf_{xy}}{f_{xx}}\right)^2 + \frac{(f_{xx}f_{yy}-f_{xy}^2)k^2}{f_{xx}^2}] \leq 0 \qquad (1.12.7)$$

Since (1.12.7) must hold for any choice of *h* and *k*, including *k* = 0, it follows that

$$f_{xx} \leq 0 \qquad (1.12.8)$$

In addition, (1.12.7) must hold in case $h = -kf_{xy} / f_{xx}$, so in view of (1.12.8), we must have

$$f_{xx}f_{yy} - f_{xy}^2 \geq 0 \qquad (1.12.9)$$

Thus, if *f(x, y)* is concave and if $f_{xx} \neq 0$, then (1.12.8) and (1.12.9) hold. Note that (1.12.8) and (1.12.9) together imply that

$$f_{yy} \leq 0 \qquad (1.12.10)$$

In case $f_{xx} = 0$, one can conduct the argument under the supposition that $f_{yy} \neq 0$, to conclude that (1.12.9) and (1.12.10) hold. These conditions are sufficient for concavity. If the inequalities are strict, *f* is strictly concave. If *f(x, y)* is a convex function, then$f_{xx} > 0$, $f_{yy} > 0$, and (1.12.9) holds as well.

A function $f(x_1, \dots \dots \dots \dots \dots, x_n)$ is concave if

$$tf(x^*) + (1-t)f(x^0) \leq f(tx^* + (1-t)x^0) \qquad (1.12.11)$$

for all $0 \leq t \leq 1$ and any pair of points $x^* = [x_1^*, \dots, x_n^*]$, $x^0 = [x_1^0, \dots, x_n^0]$ in the domain of *f*. The extension of (1.12.2) may be stated that if $f(x_1, \dots, x_n)$ is concave and differentiable, then

$$f(x^*) - f(x^0) \leq \sum_{i=1}^{n} (x_i^* - x_i^0) f_i(x_0) \qquad (1.12.12)$$

where x^* and x^0 are any two points in the domain of *f*. And, letting

$h_i = x_i^* - x_i^0, i = 1, \dots, m,$

$$f(x^*) - f(x^0) - \sum_{i=1}^{n} h_i f_i(x^0) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j f_{ij} \qquad (1.12.13)$$

by the Taylor series expansion (1.11.11), where the second partial derivatives are evaluated at an appropriate point *x* between x^* and x^0. Since the left side of (1.12.13) is nonpositive by (1.12.12), the quadratic form on the right of (1.12.13) must be nonpositive.

To state an equivalent condition for the quadratic form in h_i on the right of (1.12.13) to be nonpositive, we need some definitions. The coefficient matrix of second partial derivatives of a function *f*,

$$H = \begin{bmatrix} f_{11} & f_{12} \cdots\cdots\cdots f_{1n} \\ f_{21} & f_{22} \cdots\cdots\cdots f_{2n} \\ \vdots & \vdots \\ f_{n1} & f_{n2} \cdots\cdots\cdots f_{nn} \end{bmatrix},$$

is called a Hessian matrix of *f*. The quadratic form on the right on (1.12.13) can be written

$$\boldsymbol{hHh^T} \leq 0 \tag{1.12.14}$$

Where$\boldsymbol{h} = [h_1 \ , \ldots, h_n]$, and $\boldsymbol{h^T}$ is the transpose of $\boldsymbol{h}$. The quadratic form $\boldsymbol{hHh^T}$is said to be negative semidefinite if (1.12.14) holds for all h. (It is negative definite if (1.12.14) holds with strict inequality for all $h \neq 0$.) Equivalently, we say that the Hessian matrix $\boldsymbol{H}$ is negative semi definite if $\boldsymbol{hHh^T}$ is negative semidefinite. The matrix is negative semidefinite if its principal minors alternate in sign, beginning with negative:

$$f_{11} < 0, \quad \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} < 0, \quad \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} < 0,$$

$$(-1)^n |H| \geq 0. \tag{1.12.15}$$

The last principal minor, namely the determinant of $\boldsymbol{H}$ itself, may be zero. (If $\boldsymbol{H}$ is negative definite, then the principal minors alternate in sign and none may be zero.) It is clear from (1.12.12) and (1.12.13) that $\boldsymbol{H}$ is negative semi definite for all *x* if f is concave

If $f(x_1, , \ldots, x_n)$ is twice continuously differentiable and convex, then the sign in (1.12.12) is reversed, and thus the Hessian must be positive semidefinite. The matrix $\boldsymbol{H}$ is positive semidefinite if all its principal minors are positive, except possibly $|H|$ which may be zero. It is positive semidefinite if and only If *f* is convex.

The notion of concavity has been generalized in several ways. A function $f(x_1, , \ldots, x_n)$ is said to be quasiconcave is

$$f(tx^* + (1-t)x^0) \geq min[f(x^*), f(x^0)] \tag{1.12.16}$$

for any x^*, x° in the domain of f and for all $0 \leq t \leq 1$. Equivalently, $f(x_1, \ldots, x_n)$ is quasiconcave if and only if the set

$$A_a = \{x^* : f(x) \geq a\}$$

Is convex for every number a. A function g is quasiconvex if $-g$ is quasiconcave. Every concave function is quasiconcave, but a quasiconcave function need not be concave, nor even continuous.

1.13 Maxima and Minima:

The weierstrass theorem assures us that a continuous function assumes a maximum and a minimum on a closed bounded domain. If the hypotheses are not satisfied, then there may be no maximum and/or minimum. For instance, $f(x) = 4x - x^2$ has no maximum on $0 \leq x < 2$ since the interval is not closed, the function attains values arbitrarily close to 4, but the value 4 is not achieved on the interval. The function $f(x) = 4 + 1/x$ is not continuous on $-1 \leq x \leq 1$ and has no maximum on that interval, it becomes arbitrarily large as x approaches zero from the right. The maximum may occur on the interior of the domain or at a boundary point. It may be attained at just one or at several points in the domain.

If $f(x^*) > f(x)$ for all x near x^* that is, for all x such that $x^* - \varepsilon < x < x^* + \varepsilon$ for some $\varepsilon > 0$ then x^* is said to provide a *strict local maximum.* If $f(x^*) > f(x)$ for all x in the domain of f, then x^* provides a *strict global maximum.* Local and global minima are defined analogously. Suppose $f(x)$ is twice continuously differentiable and attains its maximum at x^* on a < x < b. From the mean-value theorem

$$f(x) - f(x^*) = f'(\bar{x})(x - x^*) \tag{1.13.1}$$

for some $\bar{x}$ between x and x^*. Since x^* maximizes f, the left side of (1) must be nonpositive and therefore the right side as well. Thus $f' \geq 0$ when $x < x^*$, and $f' \leq 0$ when $x > x^*$. Since f^* is continuous, we conclude that

$$f'(x^*) = 0 \tag{1.13.2}$$

Furthermore, from Taylor's theorem

$$f(x) - f(x^*) - (x - x^*)f'(x^*) = (x - x^*)^2 f''(\bar{x})/2 \tag{1.13.3}$$

for some $\bar{x}$ between x and x^*.Since (1.13.2) holds and x^* is maximizing, the left side of (3) is nonpositive. This implies

$$f''(x^*) \leq 0 \tag{1.13.4}$$

Thus, conditions (1.13.2) and (1.13.4) are necessary for a point x^* to maximize a twice continuously differentiable function $f(x)$ on the interior of its domain. At a local maximum, the function is stationary (1.13.2) and locally concave (1.13.4). Further, if x^* satisfies (1.13.2) and also

$$f''(x) < 0 \tag{1.13.5}$$

for all x near x^*, i.e., $x^* - \varepsilon < x < x^* + \varepsilon$, then it follows from (1.13.3) that $f(x^*) > f(x)$. Therefore, (1.13.2) and (1.13.5) are sufficient conditions for a point x^* to provide a local maximum.

Similar arguments show that necessary conditions for a local minimum are (1.13.2) and

$$f''(x^*) \geq 0 \tag{1.13.6}$$

Sufficient conditions for a local minimum are (1.13.2) and

$$f''(x) > 0 \tag{1.13.7}$$

for all x near x^*.

To find the maximum of a function of one variable, one compares the values of the function at each of the local maxima and at the boundary points of the domain, if any, and selects the largest. If the function is strictly concave over its entire domain (globally strictly concave), the maximizing point will be unique.

In Figure 1.11, the boundary point $x = g$ maximizes the function $f(x)$ over $a \le x \le g$. Points b and d are local maxima and satisfy (1.13.2) and (1.13.4). Points c and e are local minima, satisfying (1.13.2) and (1.13.6).

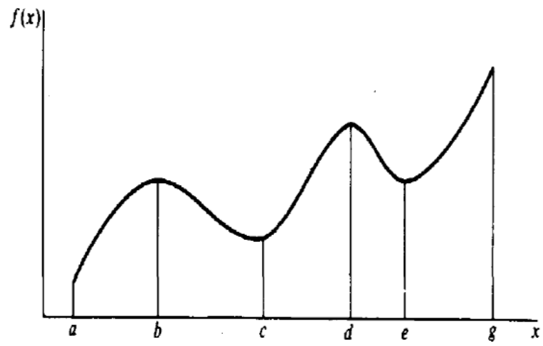


Figure 1.11

For a twice continuously differentiable function $f(x, y)$ of two variables, one can repeat the above arguments. Let x^*, y^* provide an interior maximum. Mean-value theorem (1.11..8) gives

$$f(x,y) - f(x^*,y^*) = (x - x^*)f_x(\bar{x},\bar{y}) + (y - y^*)f_y(\bar{x},\bar{y}) \tag{1.13.8}$$

for some $\bar{x}, \bar{y}$ between x^*, y^* and x, y. Since x^*, y^* is maximizing, the left side is nonpositive for all x, y. Taking $y = y^*$, we find that $x - x^*$ and f_x must have opposite signs for any x, so that

$f_x = 0$. Similarly, $f_y = 0$. Thus, it is necessary for f to be stationary at x^*, y^*:

$$f_x(x^*,y) = 0, \qquad f_y(x^*,y^*) = 0. \tag{1.13.9}$$

From the Taylor expansion (1.11.9),

$$f(x,y) - f(x^*,y^*) - hf_y(x^*,y^*) - kf_y(x^*,y^*)$$
$$= h^2 f_{xx}/2 + hkf_{xy} + k^2 f_{yy}/2 \tag{1.13.10}$$

where $h = x - x^*$, $k = y - y^*$, and the second partial derivatives on the right are evaluated at some point between x, y and x^*, y^*. But, since x^*, y^* is maximizing and since (9) holds, the left side of (1.13.10) must be nonpositive, so the right side must be as well. As shown in (3.6)-(3.10), at x^*, y^*

$$f_{xx} \leq 0, \quad \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx}f_{yy} - f_{xy}f_{yx} \geq 0 \qquad (1.13.11)$$

Thus, (1.13.9) and (1.13.11), local stationarity and local concavity, are necessary for a local maximum at x^*, y^*. Similarly, one shows that local stationarity (1.13.9) and local convexity,

$$f_{xx} \geq 0, \quad f_{xx}f_{yy} - f_{xy}^2 \geq 0 \qquad (1.13.12)$$

are necessary for a local minimum at x^*, y^*. Sufficient conditions for a local optimum are (1.13.9) and strong inequalities in (1.13.11) (maximum) or (1.13.12) (minimum) for all *x*, *y* near x^*, y^*.

To find the minimum of a function *f(x, y)*, one compares the values of the function at each of the local maxima and along the boundary of the domain and selects the largest value. If *f(x, y)* is strictly concave throughout its domain, then a local maximum will be the global maximum.

The way is clear to show that if $x^* = [x_1^*, \ldots, x_2^*]$ maximizes the twice continuously differentiable function $f(x_1, \ldots \ldots, x_n)$ on the interior of its domain, then

$$f_i(x^*) = 0, \qquad i = 1, \ldots \ldots .., n \qquad (1.13.13)$$

Furthermore, let $x = [x_1, \ldots \ldots \ldots \ldots .., x_2]$ and $h_i = x_i - x_i^*$ Then, by Taylor's theorem (1.11.11) we have

$$f(x, y) - f(x^*, y^*) - \sum_{i=1}^{n} h_i f_i(x^*) = (\tfrac{1}{2}) \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j f_{ij} \qquad (1.13.14)$$

where the second partial derivatives on the right are evaluated at some point between *x* and x^*. Since x^* is maximizing and since (1.13.13) holds, the quadratic form on the right must be nonpositive, that is, negative semidefinite. Thus, the principal minors of the Hessian matrix of *f* must alternate in sign, starting with negative (review (1.12.15)). And if the quadratic form in (1.13.14) is negative definite at x^* and (1.13.13) holds, then x^* provides a local maximum.

Analogously, necessary conditions for a minimum of *f* are stationarity (1.13.13) and local convexity, that is, positive semidefmiteness of the Hessian. Positive definiteness of the Hessian and stationarity are sufficient conditions for a local minimum.

Chapter 2

Calculus of Variations

2.1 Introduction

Calculus of variations or variational calculus deals with finding the optimum (maximum or minimum) value of a functional. Variational calculus that originated around 1696 became an independent mathematical discipline after the fundamental discoveries of L. Euler (1709-1783), whom we can claim with good reason as the founder of calculus of variations.

We start with some basic definitions and a simple variational problem of extremizing a functional. We then incorporate the plant as a conditional optimization problem and discuss various types of problems based on the boundary conditions. We briefly mention both the Lagrangian and Hamiltonian formalisms for optimization.

2.2 Basic Concepts

2.2.1 Function and Functional

We discuss some fundamental concepts associated with functionals along side with those of functions.

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

Example (2.2.1)

Consider

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

Forn $t = 1, x = 3, t = 2, x = 9$ and so on. Other functions are

$x(t) = 2t, x(t_1, t_2) = t_1^2 + t_2^2.$

Next we consider the definition of a functional based on that of a function.

(b) Functional: A variable quantity *J* is a functional dependent on a function *f (x)*, written as *J = J (f (x))*, if to each function *f (x)*, there corresponds a value *J*, i.e., we have a correspondence: to the function *f (x)* there corresponds a number *J*. Functional depends on several functions.

Example (2.2.2)

Let $x(t) = 2t^2 + 1$ then

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

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

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

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

2.2.2 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 [38, 39, 40].

Definition (2.2.1)

The increment of the function *f*, denoted by Δ*f*, is defined as

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

It is easy to see from the definition that Δf depends on both the independent variable t and the increment of the independent variable Δt, and hence strictly speaking, we need to write the increment of a function as $\Delta f(t, \Delta t)$.

Example (2.2.3)

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

find the increment of the function f (t) .

Solution: The increment Δf becomes

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

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

$$= (t_1 + \Delta t_1)^2 + (t_2 + \Delta t_2)^2 + 2(t_1 + \Delta t_1)(t_2 + \Delta t_2) - (t_1^2 + t_2^2 + 2t_1t_2)$$

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

(b) Increment of a Functional: Now we are ready to define the increment of a functional.

Definition (2.2.2) the increment of the functional J, denoted by ΔJ, is defined as

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

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 $\delta x(t)$, strictly speaking, we need to write the increment as $\Delta J(x(t), \delta x(t))$.

Example (2.2.4)

Find the increment of the functional

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

Solution: The increment of J is given by

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

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

$$= \int_{t_0}^{t_f} [4x(t)\delta x(t) + 2(\delta x(t))^2 + 1]dt$$

2.2.3 Differential and Variation

Here, we consider the differential of a function and the variation of a functional.

(a) Differential of a Function: Let us define at a point t^* the increment of the function *J* as

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

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

$$\Delta f = f(t^*) + \left(\frac{\partial f}{dt}\right)_* \Delta t + \frac{1}{2!}\left(\frac{\partial^2 f}{dt^2}\right)_* (\Delta t)^2 + \cdots - f(t^*).$$

Neglecting the higher order terms in Δt

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

Here, *df* is called the differential of *f* at the point t^*. $\dot{f}(t^*)$ is the derivative or slope of *f* at t^*. In other words, the differential *df* is the first order approximation to increment Δt. Figure 2.1 shows the relation between increment, differential and derivative.

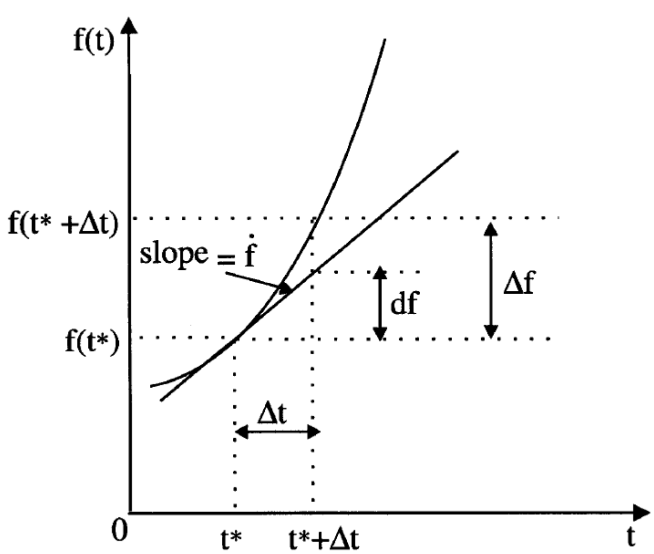


Figure 2.1 increment Δ*f*, differential *df*, and Derivative ḟ of a function *f(t)*

Example (2.2.5)

Let $f(t) = t^2 + 2t$. Find the increment and the derivative of the function *f*(t).

Solution

By definition, the increment Δf is

$$\Delta f \triangleq f(t+\Delta t) - f(t)$$
$$= (t+\Delta t)^2 + 2(t+\Delta t) - (t^2+2t)$$
$$= 2t\Delta t + 2\Delta t + \cdots + higer\ order\ terms$$
$$= 2(t+1)\Delta t$$
$$= \dot{f}(t)\Delta t$$

Here $\dot{f}(t) = 2(t+1)$.

(b) Variation of a Functional: Consider the increment of a functional

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

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

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

Where

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

are called the first variation (or simply the variation) and the second variation of the functional *J*, respectively. The variation δJ of a functional *J* is the linear (or first order approximate) part (in $\delta x(t)$) of the increment δJ. Figure 2.2 shows the relation between increment and the first variation of a functional.

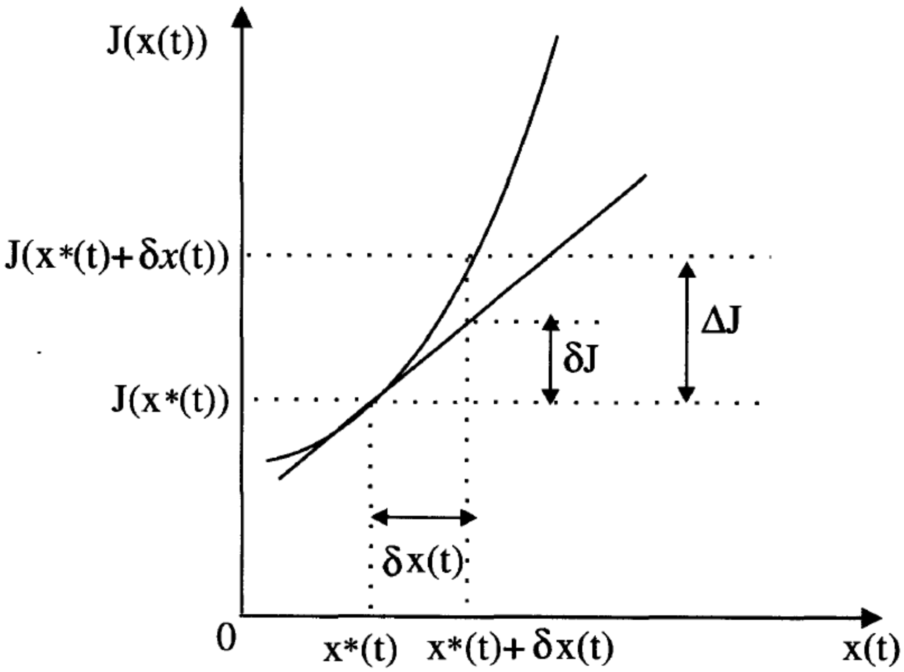


Figure 2.2 Increment ΔJ and the first variation δJ of the function J

Example (2.2.6)

Given the functional

$$J\big(x(t)\big) = \int_{t_0}^{t_f} [2x^2(t) + 3x(t) + 4]dt$$

evaluate the variation of the functional.

Solution: First, we form the increment and then extract the variation as the first order approximation. Thus

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

$$= \int_{t_0}^{t_f} [(2\big(x(t) + \delta x(t)\big)^2 + 3\big(x(t) + \delta x(t)\big) + 4)dt - (2x^2(t) + 3x(t) + 4)]dt$$

$$= \int_{t_0}^{t_f} [4x(t)\delta x(t) + 2\big(\delta x(t)\big)^2 + 3\delta x(t)]dt$$

Considering only the first order terms, we get the (first) variation as

$$\delta J\big(x(t), \delta x(t)\big) = \int_{t_0}^{t_f} [4x(t) + 3]\delta x(t)dt$$

2.3 Optimum of a Function and a Functional

We give some definitions for optimum or extremum (maximum or minimum) of a function and a functional [39, 38, 41]. The variation plays the same role in determining optimal value of a functional as the differential does in finding extremal or optimal value of a function.

Definition (2.3.1)

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 ϵ such that for all points t 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

$$\Delta f = f(t) - f(t^*) \geq 0$$

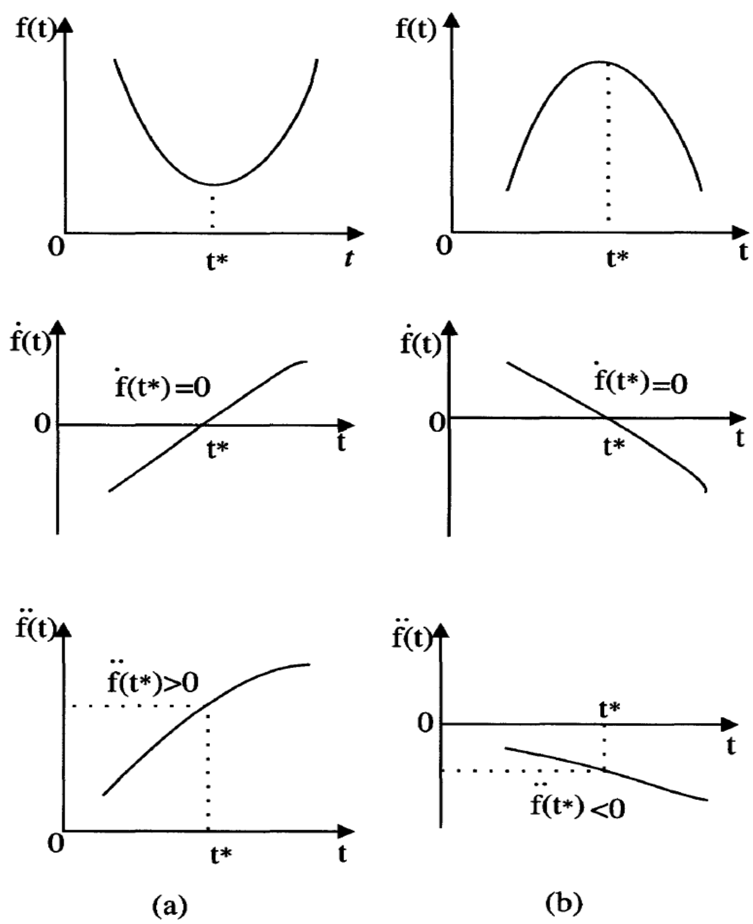


Figure 2.3 (a) Minimum (b) Maximum of a function *f(t)*

then, $f(t^*)$ is a relative local minimum. On the other hand, if

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

then, $f(t^*)$ is a relative local maximum. If the previous relations are valid for arbitrarily large ϵ, then, $f(t^*)$ is said to have a global absolute optimum. Figure 2.3 illustrates the (a) minimum and (b) maximum of a function.

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 the second differential is positive, i.e., $d^2f > 0$.
2. for maximum is that the second differential is negative, i.e., $d^2f < 0$.

If $d^2f = 0$, it corresponds to a stationary (or inflection) point.

Definition (2.3.2)

Optimum of a Functional: A functional J is said to have a relative optimum at x^* if there is a positive ϵ such that for all functions x in a domain Ω which satisfy $|x - x^*| < \epsilon$, the increment of J has the same sign.

In other words, if

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

then $J(x^*)$ is a relative minimum. On the other hand, if

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

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

Analogous to finding extremum or optimal values for functions, in variational problems concerning functionals, the result is that the variation must be zero on, an optimal curve. Let us now state the result in the form of a theorem, known as fundamental theorem of the calculus of variations, the proof of which can be found in any book on calculus of variations [39, 38, 41].

Theorem (2.3.1)

For $x^*(t)$ to be a candidate for an optimum, the (first) variation of J must be zero on $x^*(t)$, i.e., $\delta J\big(x^*(t), \delta x(t)\big) = 0$ for all admissible values of $\delta x(t)$.

This is a necessary condition. As a sufficient condition for minimum, the second variation $\delta^2 J > 0$, and for maximum $\delta^2 J < 0$.

2.4 The Basic Variational Problem

2.4.1 Fixed-End Time and Fixed-End State System

We address a fixed-end time and fixed-end state problem, where both the initial time and state and the final time and state are fixed or given a priori. Let *x (t)* be a scalar function with continuous first derivatives and the vector case can be similarly dealt with. The problem is to find the optimal function $x^*(t)$ for which the functional

$$J\big(x(t)\big) = \int_{t_0}^{t_f} V(x(t), \dot{x}(t), t)dt \qquad (2.4.1)$$

has a relative optimum. It is assumed that the integrand V has continuous first and second partial derivatives w.r.t. all its arguments; t_0 and t_f are fixed (or given a priori) and the end points are fixed, i.e.,

$$x(t = t_0) = t_0 \, , x\big(t = t_f\big) \qquad (2.4.2)$$

We already know from Theorem 2.3.1 that the necessary condition for an optimum is that the variation of a functional vanishes. Hence, in our attempt to find the optimum of *x(t)*, we first define the increment for *J*, obtain its variation and finally apply the fundamental theorem of the calculus of variations (Theorem 2.1).

Thus, the various steps involved in finding the optimal solution to the fixed-end time and fixed-end state system are first listed and then discussed in detail.

- **Step 1**: Assumption of an optimum
- **Step 2:** Variations and increment
- **Step 3:** First variation

• **Step 4:** Fundamental Theorem

• **Step 5:** Fundamental Lemma

• **Step 6:** Euler-Lagrange Equation

• **Step 1: Assumption of an optimum:** Let us assume that $x^*(t)$ is the optimum attained for the function $x(t)$. Take some admissible function $x_a(t) = x^*(t) + \delta x(t)$ close to $x^*(t)$, where $\delta x(t)$ is the variation of $x^*(t)$ as shown in Figure 2.4. The function $x_a(t)$ should also satisfy the boundary conditions (2.4.2) and hence it is necessary that

$$\delta x(t_0) = \delta\, x(t_f) = 0 \tag{2.4.3}$$

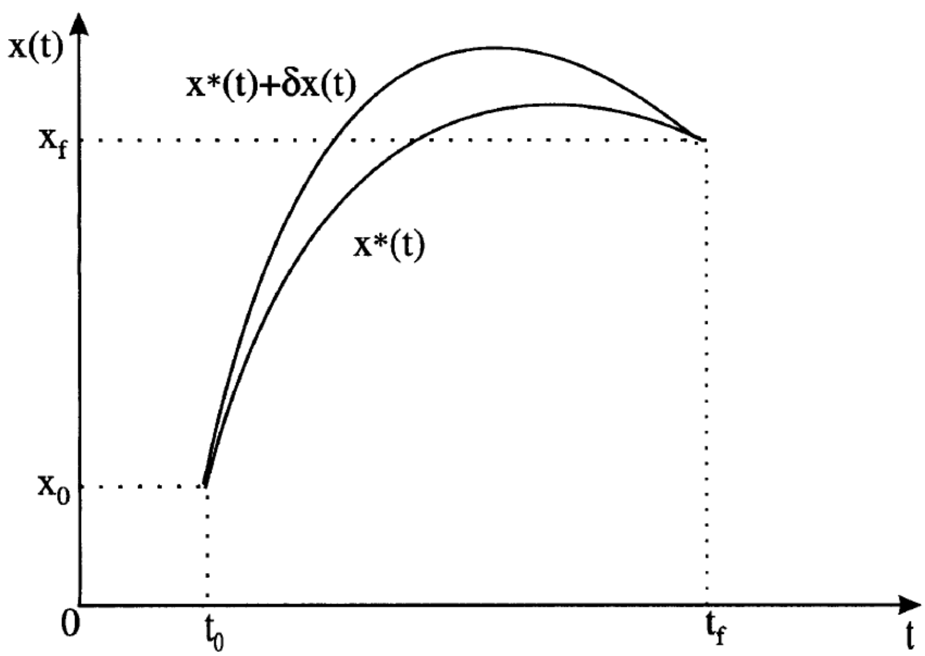


Figure 2.4 Fixed End Time and Fixed End State System

• **Step 2: Variations and Increment:** Let us first define the increment as

$$\Delta J\big(x^*(t), \delta x(t)\big) \triangleq J(x^*(t) + \delta x(t), \dot{x}^*(t) + \delta\dot{x}^*(t), t) - J(x^*(t), \dot{x}^*(t), t)$$

$$= \int_{t_0}^{t_f} V(x^*(t) + \delta x(t), \dot{x}^*(t) + \delta\dot{x}^*(t), t)dt$$

$$- \int_{t_0}^{t_f} V(x^*(t) +, \dot{x}^*(t), t)dt \qquad (2.4.4)$$

which by combining the integrals can be written as

$\Delta J\big(x^*(t), \delta x(t)\big) =$

$$\int_{t_0}^{t_f} [\, V(x^*(t) + \delta x(t), \dot{x}^*(t) + \delta\dot{x}^*(t), t) - V(x^*(t), \dot{x}^*(t), t)]dt \qquad (2.4.5)$$

Where

$$\dot{x}(t) = \frac{dx(t)}{dt} \, and \; \delta\dot{x}^*(t) = \frac{d}{dt}\{\delta\dot{x}^*(t)\} \qquad (2.4.6)$$

Expanding V in the increment (2.4.5) in a Taylor series about the point $x^*(t)$ and $\dot{x}^*(t)$, the increment ΔJ becomes (note the cancelation of $V(x^*(t), \dot{x}^*(t), t)$

$\Delta J = \Delta J\big(x^*(t), \delta x(t)\big)$

$$\int_{t_0}^{t_f} [\, \frac{\partial V(x^*(t), \dot{x}^*(t), t)}{\partial x} \delta x(t) + \frac{\partial V(x^*(t), \dot{x}^*(t), t)}{\partial \dot{x}} \delta\dot{x}^*(t)$$

$$+ \frac{1}{2!}\{ \frac{\partial^2 v(..)}{\partial x^2} + \big(\delta x(t)\big)^2 + \frac{\partial^2 v(..)}{\partial \dot{x}^2} + \big(\delta\dot{x}(t)\big)^2$$

$$2 \frac{\partial^2 v(..)}{\partial x^2 \partial \dot{x}^2} \delta x(t)\delta\dot{x}(t)\} + \cdots]dt \qquad (2.4.7)$$

Here, the partial derivatives are w.r.t. *x(t)* and $\dot{x}(t)$ at the optimal condition (*) and * is omitted for simplicity .

• **Step 3: First Variation:** Now, we obtain the variation by retaining the terms that are linear in $\delta x(t)$ and $\delta\dot{x}(t)$ as

$$\int_{t_0}^{t_f} [\, \frac{\partial V(x^*(t), \dot{x}^*(t), t)}{\partial x} \delta x(t) + \frac{\partial V(x^*(t), \dot{x}^*(t), t)}{\partial \dot{x}} \delta\dot{x}^*(t)]dt \qquad (2.4.8)$$

To express the relation for the first variation (2.4.8) entirely in terms containing $\delta x(t)$ (since $\delta\dot{x}(t)$ is dependent on $\delta x(t)$), we integrate by parts the term involving $\delta\dot{x}(t)$as (omitting the arguments in V for simplicity)

$$\int_{t_0}^{t_f} \left(\frac{\partial v}{\partial \dot{x}}\right)_* \delta\dot{x}(t)dt \quad = \int_{t_0}^{t_f} \left(\frac{\partial v}{\partial \dot{x}}\right)_* \frac{d}{dt}(\delta x(t))dt$$

$$\left[\left(\frac{\partial v}{\partial \dot{x}}\right)_* \delta\dot{x}(t)\right]_{t_0}^{t_f} - \int_{t_0}^{t_f} \delta x(t) \frac{d}{dt}\left(\frac{\partial v}{\partial \dot{x}}\right)_* dt \tag{2.4.9}$$

In the above, we used the well-known integration formula

$\int udv = uv - \int vdu$ where $u = \frac{\partial v}{\partial \dot{x}} and\ v = \delta x(t)$. Using (2.4.9), the relation (2.4.8) for first variation becomes

$\delta J\big(x^*(t), \delta x(t)\big)$

$$= \int_{t_0}^{t_f} \left(\frac{\partial v}{\partial x}\right)_* \delta x(t)dt + \left[\left(\frac{\partial v}{\partial \dot{x}}\right)_* \delta x(t)\right]_{t_0}^{t_f} - \int_{t_0}^{t_f} \frac{d}{dt}\left(\frac{\partial v}{\partial \dot{x}}\right)_* \delta x(t)dt$$

$$= \int_{t_0}^{t_f}\left[\left(\frac{\partial v}{\partial x}\right)_* - \frac{d}{dt}\left(\frac{\partial v}{\partial \dot{x}}\right)_*\right]\delta x(t)dt + \left[\left(\frac{\partial v}{\partial \dot{x}}\right)_* \delta x(t)\right]_{t_0}^{t_f} \tag{2.4.10}$$

Using the relation (2.4.3) for boundary variations in (2.4.10), we get

$$\delta J\big(x^*(t), \delta x(t)\big) = \int_{t_0}^{t_f}\left[\left(\frac{\partial v}{\partial x}\right)_* - \frac{d}{dt}\left(\frac{\partial v}{\partial \dot{x}}\right)_*\right]\delta x(t)dt \tag{2.4.11}$$

• **Step 4: Fundamental Theorem**: We now apply the fundamental theorem of the calculus of variations (Theorem 2.1), i.e., the variation of *J* must vanish for an optimum. That is, for the optimum $x^*(t)$ to exist, $\delta J\big(x^*(t), \delta x(t)\big) = 0$. Thus the relation (2.4.11) becomes

$$\int_{t_0}^{t_f}\left[\left(\frac{\partial v}{\partial x}\right)_* - \frac{d}{dt}\left(\frac{\partial v}{\partial \dot{x}}\right)_*\right]\delta x(t)dt = 0 \tag{2.4.12}$$

Note that the function $\delta x(t)$ must be zero at t_0 and t_f, but for this, it is completely arbitrary.

• **Step 5: Fundamental Lemma:** To simplify the condition obtained in the equation (2.4.12), let us take advantage of the following lemma called the fundamental lemma of the calculus of variations [38, 39, 41].

Lemma (2.4.1)

If for every function g(t) which is continuous,

$$\int_{t_0}^{t_f} g(t)\delta x(t)dt = 0 \qquad (2.4.13)$$

Where the function $\delta x(t)$ is continuous in the interval $[t_0, t_f]$ then the function g(t) must be zero everywhere throughout the interval $[t_0, t_f]$.(see Figure 2.5)

Proof

We prove this by contradiction. Let us assume that g(t) is nonzero (positive or negative) during a short interval $[t_a, t_b]$. Next, let us select $\delta x(t)$, which is arbitrary, to be positive (or negative) throughout the interval where g(t) has a nonzero value. By this selection of $\delta x(t)$, the value of the integral in (2.4.13) will be nonzero. This contradicts our assumption that g(t) is non-zero during the interval.

Thus g(t) must be identically zero everywhere during the entire interval $[t_0, t_f]$ in (2.4.13). Hence the lemma.

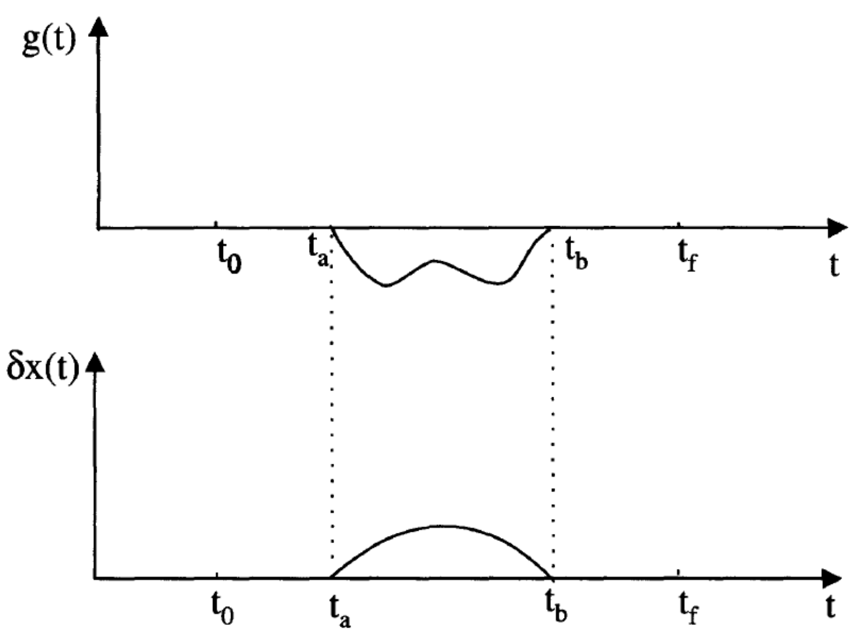


Figure 2.5 A Nonzero g(t) and an Arbitrary $\delta x(t)$

• **Step 6: Euler-Lagrange Equation:** Applying the previous lemma to (2.4.12), a necessary condition for $x^*(t)$ to be an optimal of the functional J given by (2.4.1) is

$$\left(\frac{\partial v((x^*(t),\dot{x}^*(t),t)}{\partial x}\right)_* - \frac{d}{dt}\left(\frac{\partial v(x^*(t),\dot{x}^*(t),t)}{\partial \dot{x}}\right)_* = 0 \tag{2.4.14}$$

or in simplified notation omitting the arguments in V,

$$\left(\frac{\partial v}{\partial x}\right)_* - \frac{d}{dt}\left(\frac{\partial v}{\partial \dot{x}}\right)_* = 0$$

for all $t \in [t_0, t_f]$. This equation is called Euler equation, first published in 1741 [42].

Theorem (2.4.1)

Let $I : X \rightarrow \mathbb{R}$ be a functional that is differentiable at $x^* \in X$. If I has a local extremum at x^*, then $I'(x^*) = 0$.

Proof

To be explicit, suppose that I has a minimum at x^*: there exists r > 0 such that $I(x^* + h) \geq I(x^*)$ for all h such that $\|h\| < 2\,r$. Suppose that $[I'(x^*)](h_0) \neq 0 \; for\; some\; h_0 \in X$.

$$h_n = -\frac{1}{n}\frac{[I'(x^*)](h_0)}{|[I'(x^*)](h_0)|}h_0$$

We note that $\|h_n\| \rightarrow 0\; as\; n \rightarrow \infty$, and so with N chosen large enough, we have $\|h_n\| < r$ for all n > N. It follows that for n >N,

$$0 \leq \frac{I(x^* - h_n) - I(x^*)}{\|h_n\|} = \frac{|[I'(x^*)](h_0)|}{\|h_0\|} + \mathcal{E}(h_n).$$

Passing the limit as n→∞, we obtain $-|[I'(x^*)](h_0)| \geq 0$, a contradiction

2.5 The simplest variational problem. Euler-Lagrange equation

The simplest variational problem can be formulated as follows:

Let $F(x, x^*, t)$ be a function with continuous first and second partial derivatives with respect to (x, x^*, t).

Then find $x \in C^1[t_i, t_f]$ such that $x(t_i) = x_i$ and $x(t_f) = x_f$, and which is an extremum for the functional

$$I(x) = \int_{t_i}^{t_f} F(x(t), x'(t), t)dt$$

In other words, the simplest variational problem consists of finding an extremum of a functional of the form (2.13), where the class of admissible curves comprises all smooth curves joining two fixed points; see Figure 2.6. We will apply the necessary condition for an extremum (established in Theorem 2.4.1) to the solve the simplest variational problem described above. This will enable us to solve the brachistochrone problem.

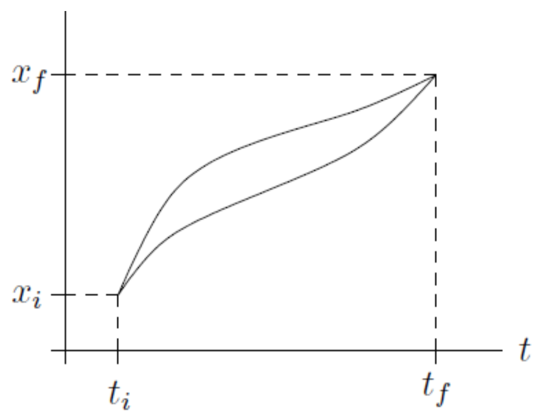


Figure 2.6: Possible paths joining the two fixed points $x(t_i) = x_i$ and $x(t_f) = x_f$.

Theorem (2.5.1)

Let I be a functional of the form

$$I(x) = \int_{t_i}^{t_f} F(x(t), x'(t), t)dt$$

where $F(x, x^*, t)$ is a function with continuous first and second partial derivatives with respect to (x, x^*, t) and $x \in C^1[t_i, t_f]$ such that $x(t_i) = x_i$ and $x(t_f) = x_f$. If I has an extremum at x^*, then x^* satisfies the Euler-Lagrange equation:

$\frac{\partial F}{\partial x}(x^*(t), x'^*(t), t) - \frac{d}{dt}(\frac{\partial F}{\partial x'}(x^*(t), x'^*(t), t) = 0 , t \in (t_i, t_f)$ (2.5.1)

(This equation is abbreviated by $F_x - \frac{d}{dt}F_{x'} = 0$)

Proof

The proof is long and so we divide it into several steps.

Step1. First of all we note that the set of curves in $C^1[t_i, t_f]$satisfying $x(t_i) = x_i$ and $x(t_f) = x_f$ do not form a linear space! So Theorem (2.4.1) is not applicable directly. Hence we introduce a new linear space X, and consider a new functional $\tilde{I} : X \to \mathbb{R}$ which is defined in terms of the old functional *I*.

Introduce the linear space

$X = \{h \in C^1[t_i, t_f] \mid h(a) = h(b) = 0\}$, with the $C^1[t_i, t_f]$-norm. Then for all $h \in X, x^* + h$ satisfies $(x^* + h)(t_i) = x_i \; and \; (x^* + h)(t_f) = x_f$.

Defining $\tilde{I}(h) = I(x^* + h)$, we note that $\tilde{I} : X \to \mathbb{R}$ has an extremum at 0. It follows from Theorem (2.4.1) that $\tilde{I}(0) = 0$. Note that by the 0 in the right hand side of the equality, we mean the zero functional, namely the continuous linear map from X to $\mathbb{R}$, which is defined by $h \to 0$ for all $h \in X$.

Step2. We now calculate $\tilde{I}(0)$. We have

$$\tilde{I}(h) - \tilde{I}(0) = \int_{t_i}^{t_f} F\big((x^* + h)(t), (x^* + h)'(t), t\big)dt - \int_{t_i}^{t_f} F(x^*, {x'}^*(t), t)dt$$

$$= \int_{t_i}^{t_i} [F(x^*(t) + h(t), x'^*(t) + h'(t), t) - F(x^*(t), x'^*(t), t)]dt$$

Recall that from Taylor's theorem, if F possesses partial derivatives of order 2 in some neighborhood N of (x_0, x'^*_0, t), then for all $(x, x^*, t) \in N$, there exists $a\, \Theta \in [0, 1]$ such that

$$F(x, x', t) = F(x_0, x'^*_0, t)$$

$$+ ((x - x_0)\frac{\partial}{\partial \mathrm{x}} + (x' - x'_0)\frac{\partial}{\partial \mathrm{x}'} + ((\mathrm{t} - \mathrm{t}_0)\frac{\partial}{\partial \mathrm{t}})\mathrm{F}\Big|_{(x_0, x'_0, t_0)} +$$

$$\frac{1}{2!}((x-x_0)\frac{\partial}{\partial \mathrm{x}}+(x'-x'_0)\frac{\partial}{\partial \mathrm{x}'}+(\mathrm{t}-\mathrm{t}_0)\frac{\partial}{\partial \mathrm{t}})^2 F\Big|_{(x_0,x'_0,t_0)+\theta(x,x',t)-(x_0,x'_0,t_0)}$$

Hence for $h\ \in\ X$ such that$\|\ h\|$ is small enough

$$\tilde{\mathrm{I}}(h)-\tilde{\mathrm{I}}(0)=\int_{t_i}^{t_f}[\frac{\partial F}{\partial x}(x^*(t),x'^*(t),t)h(t)+\frac{\partial F}{\partial x'}(x^*(t),x'^*(t),t)h'(t)]dt$$

$$+\frac{1}{2!}\int_{t_i}^{t_f}(h(t)\frac{\partial}{\partial x}+h'(t)\frac{\partial}{\partial x'})^2\Big| F_{(x^*(t)+\theta(t)h(t),x'^*(t)+\theta(t)h'(t),t)}dt$$

It can be checked that there exists a M>0 such that

$$\frac{1}{2!}\int_{t_i}^{t_f}(h(t)\frac{\partial}{\partial x}+h'(t)\frac{\partial}{\partial x'})^2\Big| F_{(x^*(t)+\theta(t)h(t),x'^*(t)+\theta(t)h'(t),t)}dt\le M\|h\|$$

And so $\tilde{\mathrm{I}}(0)$ is map

$$h\to\int_{t_i}^{t_f}[\frac{\partial F}{\partial x}(x^*(t),x'^*(t),t)h(t)+\frac{\partial F}{\partial x'}(x^*(t),x'^*(t),t)h'(t)]dt \qquad (2.5.2)$$

Step3. Next we show that if the map in (2.10) is the zero map, then this implies that (2.9) holds. Define

$$A(t)=\int_{t_i}^{t_f}\frac{\partial F}{\partial X}(x^*(\tau),x'^*(\tau),\tau)d\tau$$

Integrating by parts, we find that

$$\int_{t_i}^{t_f}[\frac{\partial F}{\partial x}(x^*(t),x'^*(t),t)h(t)dt=-\int_{t_i}^{t_f}A(t)h'(t)dt$$

and so from (2.10), it follows that $\tilde{\mathrm{I}}'(0)=0$ implies that

$$\int_{t_i}^{t_f}[-A(t)+\frac{\partial F}{\partial x'}(x^*(t),x'^*(t),t)]h'(t)dt=0\ for\ all\ h\in X$$

Step4. Finally we will complete the proof by proving the following.

Lemma (2.5.1)

If $K \in C[t_i, t_f]$ and $\int_{t_i}^{t_f} k(t)h'(t)dt = 0$ for all

$h \in C^1[t_i, t_f]with \quad h(t_i) = h(t_f) = 0$

then there exists a constant k such that $K(t) = k$ for all $t \in [t_i, t_f]$.

Proof

Let k be the constant defined by the condition

$$\int_{t_i}^{t_f} [k(t) - k]dt = 0$$

and Let $h(t) = \int_{t_i}^{t_f} [k(\tau) - k]d\tau$

furthermore

$$\int_{t_i}^{t_f} [k(t) - k]^2 dt = \int_{t_i}^{t_f} [k(t) - k]h'(t)dt$$

$$= \int_{t_i}^{t_f} k(t)h'(t)dt - k(h(t_f) - h(t_i) = 0$$

Thus $k(t) - k = 0 \; for \; all \; t \in [t_i, t_f]$

Applying lemma 2.2 we obtain

$-A(t) + \frac{\partial F}{\partial x'}(x^*(t), x'^*(t), t) = k \; forall \; t \in [t_i, t_f]$.

Differentiating with respect to t, we obtain (2.5.2). This completes the proof of **Theorem (2.5.2)**

Since the Euler-Lagrange equation is in general a second order differential equation, it solution will in general depend on two arbitrary constants, which are determined from the boundary conditions $x(t_i) = x_i$ and $x(t_f) = x_f$.

The problem usually considered in the theory of differential equations is that of finding a solution which is defined in the neighborhood of some point and satisfies given initial conditions (Cauchy's problem). However, in solving the Euler-

Lagrange equation, we are looking for a solution which is defined over all of some fixed region and satisfies given boundary conditions. Therefore, the question of whether or not a certain variational problem has a solution does not just reduce to the usual existence theorems for differential equations.

Note that the Euler-Lagrange equation is only a necessary condition for the existence of an extremum. This is analogous to the case of $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$, for which
$f'(0) = 0$, although f clearly does not have a minimum or maximum at 0. See Figure 2.7. In fact, the existence of an extremum is sometimes clear from the context of the problem. From example, in the brachistochrone problem, it is clear from the physical meaning. Similarly in the problem concerning finding the curve with the shortest distance between two given points, this is clear from the geometric meaning. If in such scenarios, there exists only one critical curve2 satisfying the boundary conditions of the problem, then this critical curve must a fortiori be the curve for which the extremum is achieved.

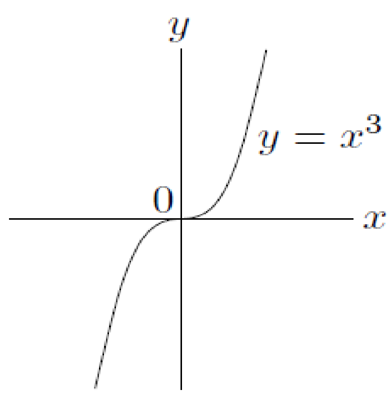


Figure 2.7: The derivative vanishes at 0, although it is not a point at which the function has a maximum or a minimum.

The Euler-Lagrange equation is in general a second order differential equation, but in some special cases, it can be reduced to a first order differential equation or where its solution can be obtained entirely by evaluating integrals.

Chapter 3

The Optimal Control

3.1 Introduction

Optimal control theory is about controlling the given system in some 'best' way. The optimal control strategy will depend on what is defined as the best way. This is usually specified in terms of a performance index functional. As a simple example, consider the problem of a rocket launching a satellite into an orbit about the earth. An associated optimal control problem is to choose the controls (the thrust attitude angle and the rate of emission of the exhaust gases) so that the rocket takes the satellite into its prescribed orbit with minimum expenditure of fuel or in minimum time. We first look at a number of specific examples that motivate the general form for optimal control problems, and having seen these, we give the statement of the optimal control.

Example of optimal control:

Example (3.1.1) (Economic growth) We first consider a mathematical model of a simplified economy in which the rate of output Y is assumed to depend on the rates of input of capital K (for example in the form of machinery) and labour force L, that is,

$$Y = P(K, L)$$

where P is called the production function. This function is assumed to have the following 'scaling' property

$$P(\alpha K, \alpha L) = \alpha P(K, L)$$

with $= \frac{1}{L}$, and defining the output rate per worker as $y = \frac{Y}{L}$ and the capital rate per worker as $k = \frac{K}{L}$, we have

$$y = \frac{Y}{L} = \frac{1}{L} P(K\ ,L) = P\left(\frac{K}{L}\ ,\frac{K}{L}\right) = P(K, 1) = \Pi(\mathrm{k}), say.$$

A typical form of Π is illustrated in Figure 2.1; we note that Π(k) > 0, but Π"(k) < 0. Output is either consumed or invested, so that

$$Y\ =\ C\ +\ I,$$

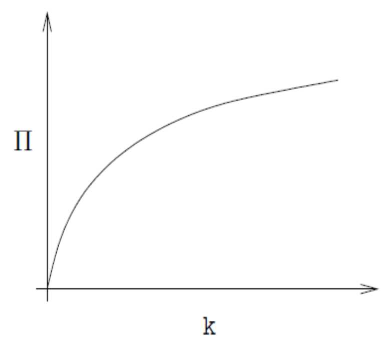


Figure3.1 Production function

Where *C* and *I* are the rates of consumption and investment, respectively.

The investment is used to increase the capital stock and replace machinery, that is

$$I(t) = \frac{dK}{dt}(t) + \mu K(t)$$

Where μ is called the rate of depreciation. Defining $\mu = \frac{C}{L}$ as the consumption rate per worker, we obtain

$$y(t) = \ \Pi\big(k(t)\big) = \ c(t) + \frac{1}{L(t)}\frac{dK}{dt}(t) + \mu k(t)$$

Since

$$\frac{d}{dt}\left(\frac{K}{L}\right) = \frac{1}{L}\frac{dK}{dt} - \frac{k}{L}\frac{dl}{dt}$$

it follows that

$$\Pi(k) = \ c + \frac{dk}{dt} + \frac{L'}{L}k + \mu k$$

Assuming that labour grows exponentially, that is $L(t) = L_0 e^{\lambda(t)}$, we have

$$\frac{dk}{dt}(t) = \Pi(k(t)) - (\lambda + \mu)k(t) - c(t),$$

This is the governing equation of this economic growth model. The consumption rate per worker, namely c, is the control input for this problem.

The central planner's problem is to choose c on a time interval [0, T] in some best way. But what are the desired economic objectives that define this best way? One method of quantifying the best way is to introduce a 'utility' function U; which is a measure of the value attached to the consumption. The function U normally satisfies $U''(\mathrm{c}) \leq 0$, which means that a fixed increment in consumption will be valued increasingly highly with decreasing consumption level.

This is illustrated in Figure 3.2. We also need to optimize consumption for [0, T], but with some discounting for future time. So the central planner wishes to maximize the 'welfare' integral

$$W(c) = \int_0^T e^{-\delta t} U(C(t))dt$$

Where δ is known as the discount rate, which is a measure of preference for earlier rather than later consumption. If δ = 0, then there is no time discounting and consumption is valued equally at all times; as δ increases, so does the discounting of consumption and utility at future times.

The mathematical problem has now been reduced to finding the optimal consumption path $\{c(t), t \in [0, T]\}$, which maximizes W subject to the constraint

$\frac{dk}{dt}(t) = \Pi(k(t)) - (\lambda + \mu)k(t) - c(t),$

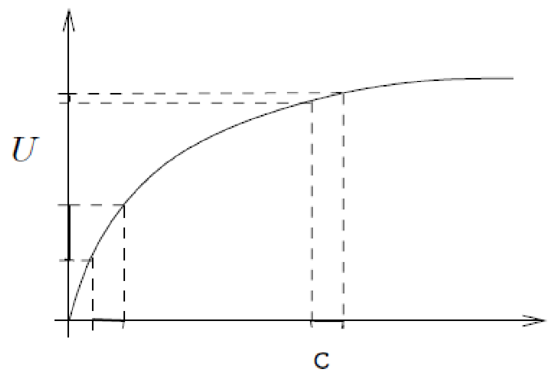


Figure 3.2: Utility function U.

and with $k(0) = k_0$.

Example (3.1.2) (Exploited populations) many resources are to some extent renewable (for example, fish populations, grazing land, forests) and a vital problem is their optimal management. With no harvesting, the resource population x is assumed to obey a growth law of the form

$$\frac{dx}{dt}(t) = \rho(x(t)) \tag{3.1.1}$$

A typical example for ρ is the Verhulst model

$$\rho(x) = \rho_0 X(1 - \frac{X}{X_s})$$

where X_s is the saturation level of population, and ρ_0 is a positive constant. With harvesting, (3.1) is modified to

$$\frac{dx}{dt}(t) = \rho(x(t)) - h(t)$$

Where h is the harvesting rate. Now h will depend on the fishing effort e (for example, size of nets, number of trawlers, and number of fishing days) as well as the population level, so that we assume

$$h(t) = e(t)x(t).$$

Optimal management will seek to maximize the economic rent defined by

$$r(t) = ph(t) - ce(t),$$

assuming the cost to be proportional to the effort, and where p is the unit price.

The problem is to maximize the discounted economic rent, called the present value V , over some period [0, T], that is,

$$V(e) = \int_0^T e^{-\delta t}(pe(t)x(t) - ce(t))dt\,,$$

Subject to

$$\frac{dx}{dt}(t) = \rho(x(t)) - e(t)x(t),$$

and the initial condition $x(0) = x_0$.

3.2 Functional

The examples from the previous section involve finding extremum values of integrals subject to a differential equation constraint. These integrals are particular examples of a ‘functional’.

A functional is a correspondence which assigns a definite real number to each function belonging to some class. Thus, one might say that a functional is a kind of function, where the independent variable is itself a function.

Examples (3.2.1)

The following are examples of functionals:

1. Consider the set of all rectifiable plane curves1. A definite number associated with each such curve, is for instance, its length. Thus the length of a curve is a functional defined on the set of rectifiable curves.

2. Let x be an arbitrary continuously differentiable function defined on $[t_i, t_f\,]$. Then the formula

$$I(x) = \int_{t_i}^{t_f} \left(\frac{dx}{dt}(t)\right)^2 dt$$

defines a functional on the set of all such functions x.

3. As a more general example, let $F(x, x', t)$ be a continuous function of three variables. Then the expression

$$I(x) = \int_{t_i}^{t_f} F\left(x(t), \frac{dx}{dt}(t), t\right) dt$$

Where x ranges over the set of all continuously differentiable functions defined on the interval $[t_i, t_f]$, defines a functional.

By choosing different functions F, we obtain different functionals. For example, if

$$F(x, x', t) = \sqrt{1 + (x')^2}$$

Then *I(x)* is the length of the curve $\{x(t), t \in [t_i, t_f]\}$, as in the first example, while if

$$F(x, x', t) = (x')^2$$

Then *I(x)* reduce to the case considered in the second example.

4. Let $f(x, u)$ and $F(x, u, t)$ be continuously differentiable functions of their arguments. Given a continuous function u on$[t_i, t_f]$, let x denote the unique solution of

$$\frac{dx}{dt}(t) = f(x(t), u(t)), x(t_i) = x_i, t \in [t_i, t_f].$$

Then *I* given by

$$I_{x_i}(u) = \int_{t_i}^{t_f} F(x(t), u(t), t)\ dt$$

defines a functional on the set of all continuous functions u on $[t_i, t_f]$.

The examples discussed in above can be put in the following form. As mentioned in the introduction, we assume that the state of the system satisfies the coupled first order differential equations

$$\frac{dx_1}{dt}(t) = f_1\big(x_1(t), \dots \dots . . x_n(t)\,, u_1(t), \dots \dots . . , u_m(t)\big), \;\; x_1(t_0) = x_{i,1}$$

$$\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots$$

$$\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots$$

$$\frac{dx_1}{dt}(t) = f_n\big(x_1(t), \dots \dots . . x_n(t)\,, u_1(t), \dots \dots . . , u_m(t)\big), \;\; x_n(t_0) = x_{i,n}$$

on $[t_i, t_f\ ]$, and where the m variables $u_1, \dots \dots, u_m$ form the control input vector u. We can conveniently write the system of equations above in the form

$$\frac{dx}{dt}(t) = f\big(x(t), u(t)\big), \qquad x(t_i) = x_i, \qquad i \in \big[t_i, t_f\ \big].$$

We assume that $u \in (C[t_i, t_f\ ])^m$, that is, each component of u is a continuous function on $[t_i, t_f\ ]$.

It is also assumed that $f_1, \dots, f_n$ possess partial derivatives with respect to x_k, $1 \leq k \leq n$ and $u_l, 1 \leq l \leq m$ and these are continuous. (So f is continuously differentiable in both variables.) The initial value of *x* is specified $(x_i\ at\ time\ t_i)$, which means that specifying *u(t)* for $t \in [t_i, t_{\ f}]$ determines *x* .

The basic optimal control problem is to choose the control $u \in (C[t_i, t_f\ ])^m$ such that:

1. The state *x* is transferred from x_i to a state at terminal time t_f where some (or all or none) of the state variable components are specified; for example, without loss of generality $x(t_f\ )k$ is specified for $k \in \{1, \dots, r\}$.

2. The functional

$$I_{x_i}(u) = \int_{t_i}^{t_f} F(x(t), u(t), t)\ dt$$

is minimized.

A function u^* that minimizes the functional I is called an optimal control, the corresponding state x^* is called the optimal state, and the pair (x^*, u^*) is called an optimal trajectory.

Definition (3.2.1)

Let $I \subseteq \mathbb{R}$ 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 $t \in I$.

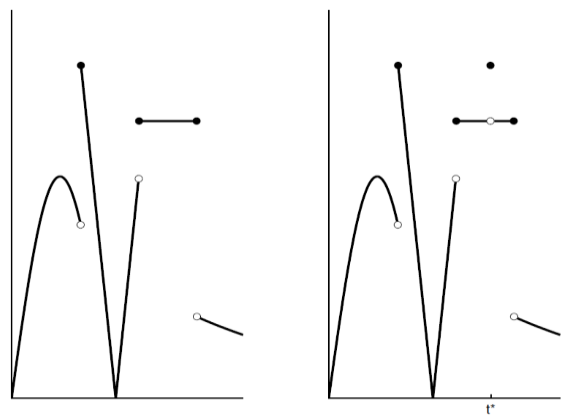


Figure 3.3:the graph to the left is an example of a piecewise continuous function.

The graph to the right is not ,because the value of the function at t^* is not the left or right limit.

Although somewhat nonstandard terminology, requiring piecewise continuous functions to equal their left or right limits eliminates a great many headaches farther down the road. In words, a piecewise continuous function can have finitely many "jump discontinuities" from one continuous segment to another. It cannot have a value that is an isolated single point (Figure 3.4).

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

$$x'(t) = g(t, x(t), u(t)) \tag{3.2.1}$$

it is meant a continuous function $x : I \to \mathbb{R}$ which is differentiable, with x' satisfying the above expression, wherever u is continuous. Equivalently, if $I = [a , b]$, then x satisfies

$$x(t) = x(a) + \int_a^t g\big(s, x(s), u(s)\big) ds.$$

An initial condition for $x(a)$ will normally be specified.

Definition (3.2.2)

Let $x : I \to \mathbb{R}$ 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.

Note, if u is piecewise continuous, and x satisfies (3.2.1), then x is piecewise differentiable. Also, the actual value of u at its discontinuities is irrelevant in determining x. All controls considered will be piecewise continuous, and we will not be concerned with values at discontinuities.

Definition (3.2.3)

Let $k : I \to \mathbb{R}$ We say k is continuously differentiable if k' exists and is continuous on I.

Definition (3.2.4)

A function $k(t)$ is said to be concave on [a , b] if

$$\alpha k(t_1) + (1-\alpha)k(t_2) \le k(\alpha t_1 + (1-\alpha)t_2)$$

$for\ all\ 0 \le \alpha \le 1\ and\ for\ any\ a \le t_1, t_2 \le b$

A function k is said to be *convex* on $[a; b]$ if it satisfies the reverse inequality, or equivalently, if $-k$ is concave. The second derivative of a twice differentiable concave function is non-positive; relating this to terminology used in calculus, concave here is “concave down" and convex is “concave up." If k is concave and differentiable, then we have a tangent line property

$$k(t_2) - k(t_1) \ge (t_2 - t_1)k'(t_2)$$

for all $a \leq t_1, t_2 \leq b$. In words, the slope of the secant line joining two points is less than the slope of the tangent line at the left point, and greater than the slope of the tangent line at the right point. See Figure 3.4.

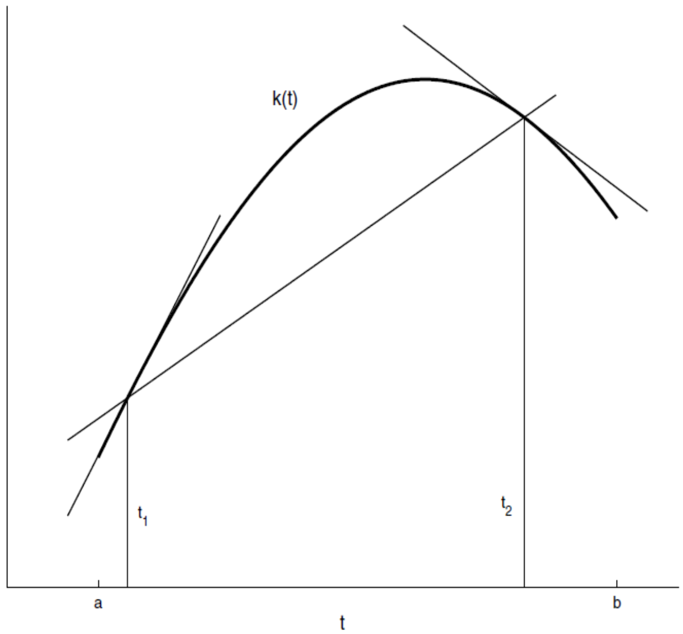


Figure 3.4 The graph of a concave function *k*(*t*). The secant line and tangent lines for two points t_1 *and* t_2 are shown.

Analogously, a function $k(x, y)$ in two variables is said to be concave if

$$\alpha k(x_1, y_1) + (1-\alpha)k(x_2, y_2) \leq k(\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha)y_2)$$

for all $0 \leq \alpha \leq 1$ and all $(x_1, y_1), (x_2, y_2)$ in the domain of *k*. If *k* is such a function and has partial derivatives everywhere, then the analogue to the tangent line property is

$$k(x_1, y_1) - k(x_2, y_2) \geq (x_1 - x_2)k_x(x_1, y_1) + (y_1 - y_2)k_y(x_1, y_1)$$

for all pairs of points $(x_1, y_1), (x_2, y_2)$ in the domain of *k*.

Definition (3.2.5)

A function k is called Lipschitz if there exists a constant c (particular to k) such that

$|k(t_2) - k(t_1)| \leq c|t_1 - t_2|$for all points t_1, t_2 in the domain of *k*. The constant *c* is called the Lipschitz constant of *k*.

3.3 The Basic Problem and Necessary Conditions:

In our basic optimal 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:

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

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

$u(t) \rightarrow x = x(u)$ (of course, *x* is really a function of the independent variable *t*, we write *x*(*u*) simply to remind us of the dependence on *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

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

$$\text{Subject to } x'(t) = g(t, x(t), u(t)) \qquad (3.3.1)$$

$$x(t_0) = x_0 \; and \; x(t_1) free.$$

Such a maximizing control is called an optimal control. By $x(t_1)$ free, it ismeant that the value of $x(t_1)$ is unrestricted. For our purposes, f and g will always be continuously differentiable functions in all three arguments. Thus, as the control(s) will always be piecewise continuous, the associated states will always be piecewise differentiable.

The principle technique for such an optimal control problem is to solve a set of “necessary conditions" that an optimal control and corresponding state must satisfy. It is important to understand the logical difference between necessary conditions and suffcient conditions of solution sets.

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

Sufficient Conditions: If $u^*(t)$, $x^*(t)$ satisfy the following conditions ..., then $u^*(t)$, $x^*(t)$ are optimal. For now, let us derive the necessary conditions. Express our objective functional in terms of the control:

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

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

The necessary conditions that we derive were developed by Pontryagin and his co-workers in Moscow in the 1950's . Pontryagin introduced the idea of “adjoint" functions to append the differential equation to the objective functional. Adjoint functions have a similar purpose as Lagrange multipliers in multivariate calculus, which append constraints to the function of several variables to be maximized or minimized.

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 another piecewise continuous control. Let x^ϵ be the state corresponding to the control u^ϵ, namely, x^ϵ satisfies

$$\frac{d}{dt}x^\epsilon(t) = g(t, x^\epsilon(t), u^\epsilon(t)) \qquad (3.3.2)$$

wherever u^ϵ is continuous.

Since all trajectories start at the same position, we take $x^\epsilon(t_0) = x_0$ (Figure 3.5).

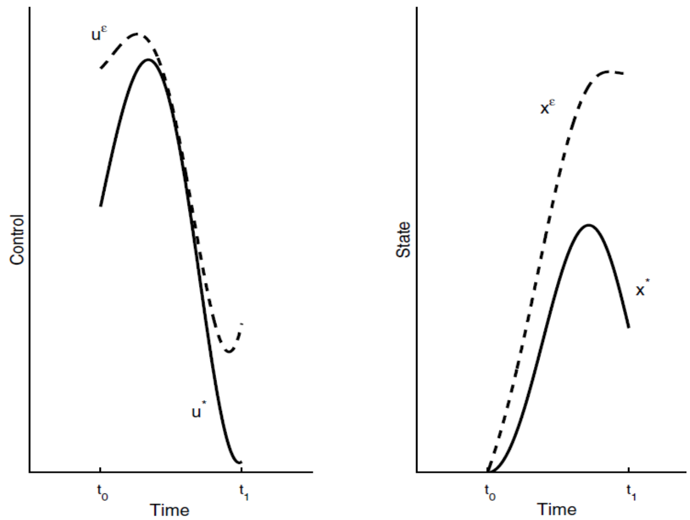


Fig 3.5 The optimal control u* and state x* (in solid) plotted together with u^* and x^* (dashed).

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

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

In fact, something similar is true for x^ϵ. Because of the assumptions made on g, it follows that

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

for each fixed t. Further, the derivative

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

exists for each t. The actual value of quantity will prove unimportant. We need only to know that it exists.

The objective functional at u^ϵ is

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

We are now ready to introduce the adjoint function or variable ¸. Let $\lambda(t)$ be a piecewise differentiable function on $[t_0, t_1]$ to be determined. By the Fundamental Theorem of Calculus,

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

which implies

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

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

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

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

$$+ \lambda(t_0)x^\epsilon(t_0) - \lambda(t_1)x^\epsilon(t_1)$$

where we used the product rule and the fact that $g(t,x^\epsilon,u^\epsilon) = \frac{d}{dt}x^\epsilon$ at all but finitely many points. Since the maximum of J with respect to the control u occurs at u^*, the derivative of $J(u^*)$ with respect to ϵ (in the direction h) is zero, i.e.,

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

This gives a limit of an integral expression. A version of the Lebesgue Dominated Convergence Theorem above allows us to move the limit (and thus the derivative) inside the integral. This is due to the compact interval of integration and the piecewise differentiability of the integrand. Therefore,

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

$$= \int_{t_0}^{t_1} \frac{\partial}{\partial \epsilon} f\big(t, x^\epsilon(t), u^\epsilon(t)\big) + \big(\lambda'(t) x^\epsilon(t)\big)$$

$$+ \lambda(t) g\big(t, x^\epsilon(t), u^\epsilon(t)\big) dt\Big|_{\epsilon=0} - \frac{d}{d\epsilon} \lambda(t_1) x^\epsilon(t_1)\Big|_{\epsilon=0}$$

Applying the chain rule to *f* and *g*, it follows

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

$$- \lambda(t_1) \frac{\partial x^\epsilon}{\partial \epsilon}(t_1)\Big|_{\epsilon=0} \qquad (3.3.3)$$

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

Rearranging the terms in (3.4.3) gives

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

$$\lambda(t_1) \frac{\partial x^\epsilon}{\partial \epsilon}(t_1)\Big|_{\epsilon=0} \qquad (3.3.4)$$

We want to choose the adjoint function to simplify (3.3.4) by making the coefficients of

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

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

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

and the boundary condition

$\lambda(t_1) = 0$ (*transversality condition*).

Now (3.5) reduces to

$$0 = \int_{t_0}^{t_1} \Big(f_u\big(t, x^*(t), u^*(t)\big) + \lambda(t) g_u\big(t\,, x^*(t), u^*(t)\big) \Big) h(t) dt.$$

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

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

In this case

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

which implies the optimality condition

$f_u\big(t, x^*(t), u^*(t)\big) + \lambda(t) g_u\big(t, x^*(t), u^*(t)\big) = 0 \; for \; all \; t_0 \le t \le t_1.$

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,

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

= integrand + adjoint * RHS of DE*:*

We are maximizing *H* with respect to *u* at u^*, and the above conditions can be written in terms of the Hamiltonian:

$\frac{\partial H}{\partial u} = 0 \; at \; u^* \Rrightarrow \; f_u + \lambda g_u$ (*Optimality condition*),

$\lambda' = -\frac{\partial H}{\partial x} \Rrightarrow \lambda' = -(f_x + \lambda g_x)$ (*adjoint equation*),

$\lambda(t_1) = 0$ (*transversality condition*).

We are given the dynamics of the state equation:

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

3.4 Pontryagin's Maximum Principle:

The Pontryagin maximum principle is stated somewhat differently from our usage. Our version is correct only under more stringent conditions than have been fully stated. We shall set forth the Pontryagin maximum principle and then note the differences between it and the version given in earlier solutions.

The problem

Find a piecewise continuous control vector $u(t) = [u_1(t), \ldots\ldots, u_m(t)]$ and an associated continuous and piecewise differentiable state vector $x(t) = [x_1(t), \ldots\ldots, x_n(t)]$, defined on the fixed time interval $[t_0, t_1]$, that will

$$\max \int_{t_0}^{t_1} f\big(t, x(t), u(t)\big) dt \tag{3.4.1}$$

subject to the differential equations

$$x'_i(t) = g_i\big(t, x(t), u(t)\big)\ , i = 1, \ldots\ldots, n \tag{3.4.2}$$

initial conditions

$$x_i(t_0) = x_{i0}\ ,\ i = 1, \ldots\ldots, n \tag{3.4.3}$$

terminal conditions

$$x_i\ (t_1\ ) = x_{i1}\ ,\quad i = 1, \ldots\ldots, p$$

$$x_i(t_1) \geq x_{i1}\ ,\qquad i = p + 1, \ldots\ldots, q \quad (x_{i1}\ ,\ i = 1, \ldots\ldots, q, \quad fixed)$$

$$x_i\ (t_1\ )\ free\ ,\quad i = q + 1, \ldots\ldots, n \tag{3.4.4}$$

and control variable restriction

$$u(t) \in U, \quad U\ \ a\ given\ set\ in\ R^m \tag{3.4.5}$$

We assume that $f, g_i, \partial f/\partial x_j\ , and\ \ \partial g_i/\partial x_j$ are continuous functions of all their arguments, for all

$i\ =\ 1\ ,\ldots, n\ and\ j\ =\ 1\ ,\ldots, n$.

Theorem (3.4.1)

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 (3.5.1),with associated state x^*, and u^* ¸ a piecewise differentiable function with ¸$\lambda(t) \geq 0$¸ 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,

$$\begin{aligned} H(t, x^*(t), u^*(t), \lambda(t)) - H(t, x^*(t), u(t), \lambda(t)) \\ &= [f(t, x^*(t), u^*(t)) + \lambda(t) g(t, x^*(t), u^*(t))] \\ &\quad -[f(t, x^*(t), u(t)) + \lambda(t) g(t, x^*(t), u(t))] \\ &= [\mathrm{f}(\mathrm{t}, \mathrm{x}^*(\mathrm{t}), \mathrm{u}^*(\mathrm{t})) - \mathrm{f}(\mathrm{t}, \mathrm{x}^*(\mathrm{t}), \mathrm{u}(\mathrm{t}))] \\ &\quad +\lambda(t)[g(t, x^*(t), u^*(t)) - g(t, x^*(t), u(t))] \end{aligned}$$

The transition from line 3 to line 4 is attained from applying the tangent line property to *f* and *g*, and because ¸λ(t)≥ 0 .

An identical argument generates the same necessary conditions when the problem is minimization rather than maximization. In a minimization problem, we are minimizing the Hamiltonian point wise, and the inequality in Pontryagin's Maximum Principle in reversed. Indeed, for a minimization problem with *f*, *g* being convex in *u*, we can derive

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

by the same argument as in Theorem 3.4.1.

We have converted the problem of finding a control that maximizes (or minimizes) the objective functional subject to the differential equation and initial

condition, to maximizing the Hamiltonian point wise with respect to the control. Thus to find the necessary conditions, we do not need to calculate the integral in the objective functional, but only use the Hamiltonian. Later,
we will see the usefulness of the property that the Hamiltonian is maximized pointwise by an optimal control.

We can also check concavity conditions to distinguish between controls that maximize and those that minimize the objective functional [14]. If

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

then the problem is maximization, while

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

goes with minimization.

We can view our optimal control problem as having two unknowns, u^* and x^*, at the start. We have introduced an adjoint variable λ, which is similar to a Lagrange multiplier. It attaches the differential equation information onto the maximization of the objective functional. The following is an outline of how this theory can be applied to solve the simplest problems.

1. Form the Hamiltonian for the problem.
2. Write the adjoint differential equation, transversality boundary condition, and the optimality condition. Now there are three unknowns, u^*, x^*, and λ.
3. Try to eliminate u¤ by using the optimality equation $H_u = 0$, i.e., solve for u^* in terms of x^* and λ.
4. Solve the two differential equations for x^* and λ ¸ with two boundary conditions, substituting u^* in the differential equations with the expression for the optimal control from the previous step.
5. After finding the optimal state and adjoint, solve for the optimal control.

Example (3.4.1)

$$\min_{u} \int_{0}^{1} u^2(t)dt$$

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

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

The optimality condition is

$$0 = \frac{\partial H}{\partial u} = 2u + \lambda \quad at\, u^* \implies u^* = -\frac{1}{2}\lambda$$

We see the problem is indeed minimization as

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

The adjoint equation is given by

$$\lambda' = -\frac{\partial H}{\partial x} = -\lambda \implies \lambda(t) = ce^{-t}$$

for some constant c. But, the transversality condition is

$$\lambda(1) = 0 \implies ce^{-1} = 0 \implies c = 0$$

Thus, $\lambda=0$, so that $u^* = \frac{-\lambda}{2} = 0$.

So, x^* satisfies $x' = x$ and $x(0) = 1$.

Hence, the optimal solutions are

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

and the state function is plotted in Figure 3.6

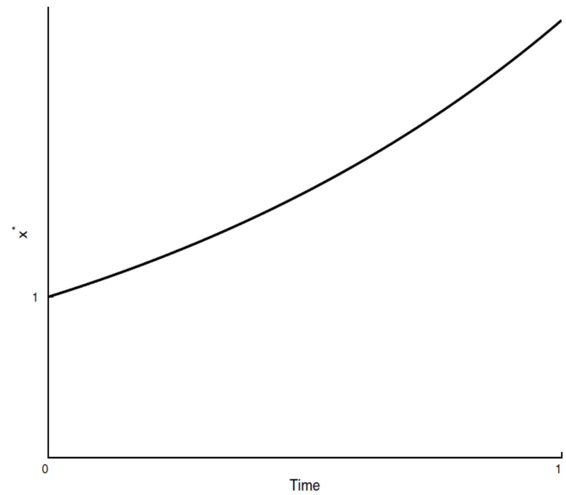


Figure 3.6 Optimal state for Example 3.4.1 plotted as a function of time.

Example (3.4.2)

$$\min_u \frac{1}{2}\int_{t_0}^{t_1} \left(3x^2 + u^2(t)\right)dt$$

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

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

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

The optimality condition gives

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

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 λ,

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

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 λ are

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

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

Thus, using the optimality equation, the optimal solutions are

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

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

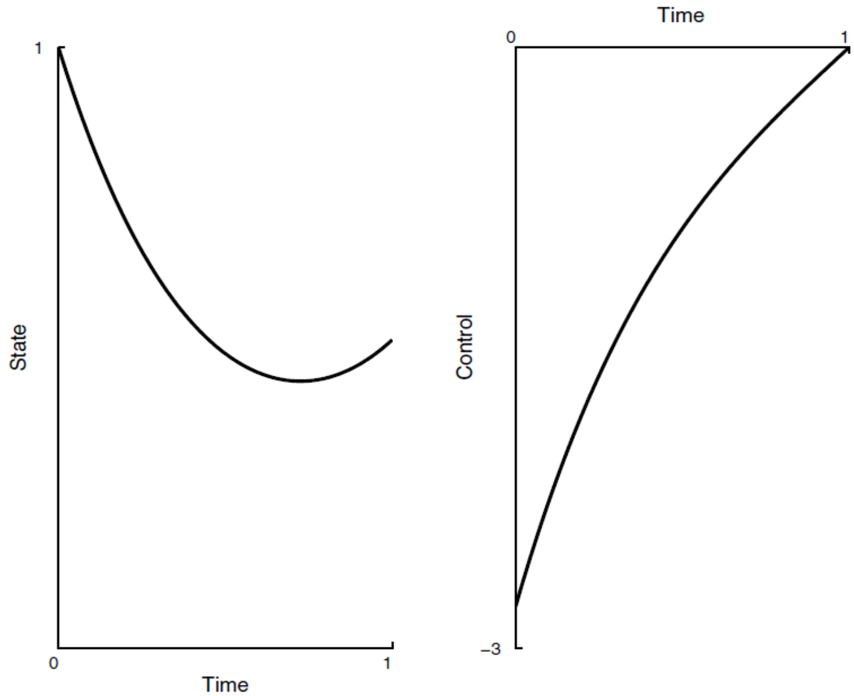


Figure 3.7 Optimal control and state

3.5 Existence and Other Solution Properties:

We developed necessary conditions to solve basic optimal control problems. However, some difficulties can arise with this method. It is possible that the necessary conditions could yield multiple solution sets, only some of which are optimal controls. Further, recall that in the development of the necessary conditions, we began by assuming an optimal control exists. It is also possible that the necessary conditions could be solvable when the original optimal control problem has no solution. We expect the objective functional evaluated at the optimal state and control to give a finite answer. If this objective functional value turns out to be ∞ $or - \infty$, we would say the problem has no solution. An example of this is given below.

Example (3.5.1)

$$\min_{u} \int_0^1 \big(x(t) + u(t)\big)dt$$

Subject to $x'(t) = 1 - u^2(t), x(0) = 1.$

The Hamiltonian and the optimality condition are:

$$H = x + u + \lambda(1 - u^2)$$

$$\frac{\partial H}{\partial u} = 1 - 2\lambda u = 0 \; at \; u^* \Longrightarrow u^* = \frac{1}{2\lambda}$$

From the adjoint equation and its boundary condition,

$$\lambda'(t) = -\frac{\partial H}{\partial x} = -1 \; and \; \lambda(1) = 0.$$

We can directly calculate

$$\lambda(t) = 1 - t.$$

Note that the concavity with respect to the control *u* is correct for a maximization problem,

$$H_{uu} = -2\lambda \leq 0$$

as $\lambda(t) \geq 0$. Next, we calculate the optimal state using the differential equation and its boundary condition

$$x' = 1 - u^2 = 1 - \frac{1}{4(1-t)^2} \quad and \quad x(0) = 1,$$

and find that

$$x^*(t) = t - \frac{1}{4(1-t)} + \frac{5}{4} \quad ,$$

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

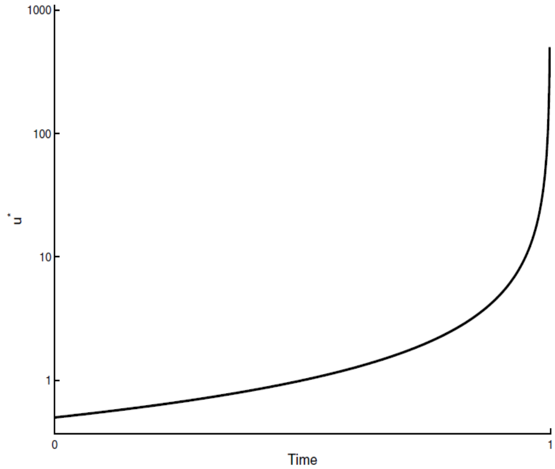


Figure 3.8 the graph of *uplotted in logarithmic scale. The value of *u**tends to infinity as *t* approaches 1.**

3.5.1 Existence and Uniqueness Results

Theorem (3.5.1)

Consider

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

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

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 $t_0 \leq t \leq t_1$:

$$f_u + \lambda g_u = 0,$$

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

$$\lambda(t_1) = o,$$

$$\lambda(t) \geq 0.$$

Then for all controls u, we have

$$J(u^*) \geq J(u).$$

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

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

This gives

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

$$\geq \int_{t_0}^{t_1} (x^*(t) - x(t)) f_x(t, x^*, u^*) + (u^*(t) - u(t)) f_u(t, x^*, u^*) dt \qquad (3.5.1)$$

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 by the hypothesis, the last term in (3.6.1) becomes

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

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

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

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

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

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 $J(u^*) - J(u) \geq 0.$

3.5.2 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 (3.5.2)

Let u^* be an optimal control, and x^* the resulting state, for the problem

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

$$\text{Subject to } x'(t) = g\big(t, x(t), u(t)\big),\ x(t_0) = x_0 \tag{3.5.2}$$

Let $t^\wedge$ be a fixed point in time such that $t_0 < t^\wedge < t_1$. Then, the restricted functions

$u^{\wedge *} = u^{\wedge *}\big|_{[t^\wedge, t_1]}$, form an optimal pair for the restricted problem

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

$$\text{Subject to } x'(t) = g\big(t, x(t), u(t)\big),\ x(t) = x^*(t^\wedge) \tag{3.5.3}$$

Further, if u^* is the unique optimal control for (3.5.2), then $u^{\wedge *}$ is the unique optimal control for (3.5.3).

Proof

This proof is done by contradiction. Suppose, to the contrary, that $u^{\wedge *}$is not optimal, i.e., there exists a control $u_1^{\wedge}$ on the interval $[t^{\wedge}, t_1]$ such that $J^{\wedge}(u_1^{\wedge}) > J^{\wedge}(u^{\wedge *})$ Construct a new control u_1 on the whole interval $[t_0, t_1]$ as follows

$$u_1(t) = \begin{cases} u^*(t)\ For\ t_0 \le t \le t_1, \\ u_1^{\wedge}\ For\ t^{\wedge} \le t \le t_1. \end{cases}$$

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

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

$$= J^{\wedge}(u_1^{\wedge}) - J^{\wedge}(u^{\wedge *}) > 0.$$

However, this contradicts our initial assumption that u^* was optimal for (3.5.2). Thus, no such control $u_1^{\wedge}$ exists, and $u^{\wedge *}$ is optimal for (3.5.3).

The proof of the result concerning uniqueness follows in almost exactly the same manner and is left.

Example (3.5.2)

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

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

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

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

and the optimality condition leads to

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

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.,

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

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

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:

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

while the optimality is also unchanged,

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

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

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

Of course, we see the same optimal pair as above, as called for by the Principle of optimality.

Example (3.5.3)

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

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

Again, the Hamiltonian is the same, so that the adjoint and optimality conditions are unchanged. However, the transversality condition is now different,

$$\lambda'(t) = -\frac{\partial H}{\partial u} = -1 - \lambda\,, \lambda(1) = 0 \Longrightarrow \lambda(t) = e^{1-t} - 1$$

So that

$$u^*(t) = -\lambda(t) = 1 - e^{1-t}.$$

Using this in the state equation,

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

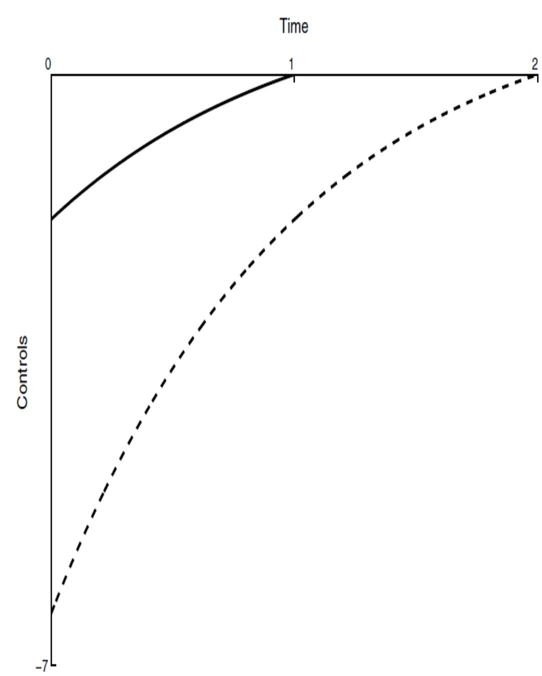


Figure 3.9 Optimal controls for Examples 3.6.1 (dashed) and 3.6.2. (solid) plotted together.

3.6 State Conditions at the Final Time

Up to this point, we have viewed the value of the state at the terminal time to be immaterial, i.e., the objective functional (our goal) did not explicitly depend on $x(t_1)$. However, there are situations where we might wish to take it into consideration.

3.6.1 Payoff Terms

Many times, in addition to maximizing (or minimizing) terms over the entire time interval, we will wish to also maximize a function value at one particular point in time, specifically, the end of the time interval. For example, suppose you want to minimize the tumor cells at the final time in a cancer model, or the number of infected individuals at the final time in an epidemic model.

The necessary conditions must be appropriately altered. In general, consider the following set-up,

$$\max_u \left[\emptyset(x(t_1) + \int_{t_0}^{t_1} f\big(t, x(t), u(t)\big) dt \right]$$

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

Where$\emptyset(x(t_1)$ is a goal with respect to the final position or population level, $x(t_1)$. We call $\emptyset(x(t_1)$a payoff term. It is sometimes referred to as the salvage term.

Consider the resulting change in the derivation of the necessary conditions. Our objective functional becomes

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

In the calculation of

$$0 = \lim_{\epsilon \to 0} \frac{J(u^\epsilon) - J(u^*)}{\epsilon}$$

the only change occurs in the conditions at the final time

$$0 = \int_{t_0}^{t_1} \left[(f_x + \lambda g_x + \lambda') \frac{dx^\epsilon}{d\epsilon}\bigg|_{\epsilon=0} + (f_u + \lambda g_u)h \right] dt$$

$$- \Big(\lambda(t_1) - \emptyset\big(x(t_1)\big)\Big) \frac{\partial x^\epsilon}{\epsilon}(t_1)\bigg|_{\epsilon\to 0} \qquad (3.6.1)$$

So, if we choose the adjoint variable λ to satisfy the previous adjoint equation and also

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

$$\lambda(t_1) = \emptyset'\big(x^*(t_1)\big)$$

then (3.6.1) reduces to

$$0 = \int_{t_0}^{t_1} (f_u + \lambda g_u) h \, dt$$

and the optimality condition

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

follows as before. So, the only change in the necessary conditions is in the transversality condition

$$\lambda(t_1) = \emptyset'\big(x^*(t_1)\big).$$

To clarify how to calculate this adjoint final time condition, consider the following examples.

Example (3.6.1)

$$\max_u \int_0^{T} f\big(t, x(t), u(t)\big) dt + 5x(T)^3$$

Subject to $x'(t) = g\big(t, x(t), u(t)\big), x(0) = x_0$.

Here we have

$$\emptyset(s) = 5s^3 \implies \emptyset'(s) = 15s^2$$

so that the transversality condition is

$$\lambda(T) = 15x^*(T)^2.$$

Example (3.6.2)

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

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

Note, this problem is identical to Example (3.5.1), except for the addition of the payoff term. So now, our goal includes minimizing the term $x(1)^2$, in addition to the square integral of the control. We can view this as minimizing a population, with exponential growth, at the end of a time frame. We should expect u to be negative, in order to decrease x, but $|u|$ cannot be too large because of the integral. The Hamiltonian in this example is

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

The optimality condition gives

$$0 = \frac{\partial H}{\partial u} = u + \lambda \Longrightarrow u^*(t) = -\lambda(t).$$

Also, the adjoint equation is

$$\lambda'(t) = -\frac{\partial H}{\partial x} = -\lambda \Longrightarrow \lambda(t) = Ce^{-t},$$

for some constant C. Hence,

$$u^*(t) = -\lambda(t) = -Ce^{-t},$$

So,

$$x'(t) = x - Ce^{-t}, x(0) = 1,$$

Which gives

$$x^*(t) = \frac{C}{2}e^{-t} + Ke^{t}$$

Where K is a constant. Recall, the transversality condition here is

$$\lambda(1) = \emptyset'\big(x(1)\big) = \big(x^2(1)\big)' = 2x(1).$$

We have the system of linear equations

$$1 = x(0) = \frac{C}{2} + K$$

$$Ce^{-1} = \lambda(1) = 2x(1) = Ce^{-1} + 2Ke^{1}$$

which can be solved to give C = 2, K = 0. Thus

$$x^{*}(t) = e^{-t} \ , \qquad u^{*}(t) = -2e^{-t}$$

and u^* is negative as expected.

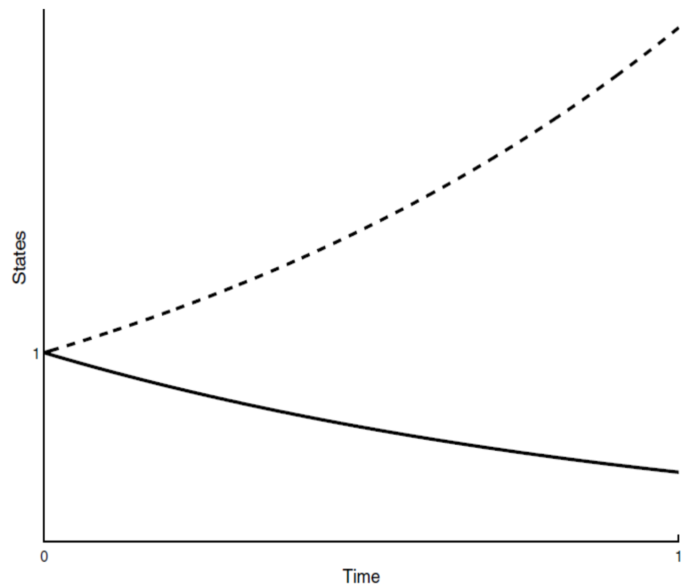


Figure 3.10 The optimal state for Example 3.2 (solid) and Example 3.1(dashed).

Example (3.6.3)

Let $x(t)$ represent the number of tumor cells at time t (with exponential growth factor α), and $u(t)$ the drug concentration. We wish to simultaneously minimize the number of tumor cells at the end of the treatment period and the accumulated harmful effects of the drug on the body. So, the problem is

$$\min_{u} x(T) + \int_0^T u(t)^2 dt$$

Subject to $x'(t) = \alpha x(t) - u(t), x(0) = x_0 > 0.$

This model is very simple and unrealistic; we use it for illustrative purposes only. A more sophisticated and interesting model is used in Lab

Note that $\emptyset(s) = s$ here, so that $\emptyset'(s) = 1$. First, we construct the Hamiltonian and then calculate the necessary conditions:

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

$$\frac{\partial H}{\partial u} = 2u - \lambda = 0 \; at \; u^* \Longrightarrow u^* = \frac{\lambda}{2},$$

$$\lambda' = -\frac{\partial H}{\partial x} = -\alpha\lambda \Longrightarrow \lambda = Ce^{-\alpha t},$$

$$\lambda(T) = 1.$$

This gives the adjoint variable,

$$\lambda(t) = e^{\alpha(T-t)}.$$

Hence, we obtain the optimal control

$$u^*(t) = \frac{e^{\alpha(T-t)}}{2}.$$

and we can then solve for the optimal state

$$x' = \alpha x - u = \alpha x - \frac{e^{\alpha(T-t)}}{2}, x(0) = x_0.$$

This ODE can be solved using an integration factor to find

$$x^*(t) = x_0 e^{\alpha t} + e^{\alpha T}\frac{e^{-\alpha t} - e^{\alpha t}}{4\alpha}.$$

3.6.2 States with Fixed Endpoints

There are various possibilities of fixing the position of the state at the beginning or at the end of the time interval or both. The objective functional could depend on the final or initial position. Consider the problem

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

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

$$x(t_0) free, x(t_1) = x_1 \ fixed.$$

This is different than the problems we have been examining, as the state is fixed at the end of the time interval, not at the beginning. However, the same argument we used in section 3.1, with the adjoint chosen appropriately, shows that the necessary conditions for an optimal pair u^*, x^* will be as before, with only the transversality condition changed. Specifically,

$$\lambda(t_0) = \emptyset'\big(x(t_0)\big).$$

Consider the problem below, where the state is fixed at both the beginning and end of the time interval,

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

$$\text{Subject to } x'(t) = g\big(t, x(t), u(t)\big) \qquad (3.6.2)$$

$$x(t_0) = x_0, x(t_1) = x_1 \ both \ fixed.$$

The maximization here is over all admissible controls. That is, the set of controls which adhere to all stated restrictions (explicit and implicit). In the case of (3.6.2), this would mean all controls which steer the state from the fixed initial condition to the fixed final condition. A slight modification of the necessary conditions is needed to solve such a problem.

Example (3.6.4)

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

$$\text{Subject to } x'(t) = u(t), x(0) = 0, x(4) = 1.$$

We begin by forming the Hamiltonian

$$H = u^2 + x + \lambda u .$$

We have no transversality condition, as x has both boundary conditions, but we make use of the adjoint condition,

$$\lambda'(t) = -\frac{\partial H}{\partial x} = -1 \Longrightarrow \lambda(t) = k - t$$

for some constant *k*. Then, the optimality condition gives

$$0 = \frac{\partial H}{\partial x} = 2u + \lambda \;\Longrightarrow u^* = -\frac{\lambda}{2} = \frac{t-k}{2}$$

Solving the state equation with this control gives

$$x^*(t) = \frac{t^2}{4} - \frac{kt}{2} + c$$

for some constant c. Using the boundary conditions, *x(0) = 0* implies c = 0, and $x(4) = 1$ gives $k = \frac{3}{2}$. So,

$$u^*(t) = \frac{2t-3}{4} \quad and \quad x^*(t) = \frac{t^2 - 3t}{4}$$

Example (3.6.5)

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

Subject to $x'^{(t)} = x(t) + u(t),\ x(0) = 1, x(1) = 0.$

This is another variation on Examples 3.5.1 and 3.7.2 The objective functional once again does not depend on *x*, but we must choose a control that moves *x* from 1 to 0. Again, we expect a negative *u*. The Hamiltonian is

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

As before, the optimality condition gives

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

Also

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

for some constant C. Thus,

$$u^*(t) = -\lambda(t) = -Ce^{-t}$$

So that

$$x'(t) = x - Ce^{-t}$$

$$x^*(t) = \frac{C}{2}e^{-t} + ke^{t}$$

Enforcing the boundary conditions on *x*, we find

$$1 = x(0) = \frac{C}{2} + k$$

$$0 = x(1) = \frac{C}{2}e^{-1} + ke$$

Which gives $C = \frac{2e^2}{e^2-1}$ and $= \frac{1}{1-e^2}$, so that

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

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

Note, in Example 3.6.2 we wanted to minimize the value of $x(1)^2$ and the cumulative effect of the control. So, we wanted to push $x(1)$ close to 0. Here, we choose the control with the smallest cumulative effect that forces the state to 0. If we plug the optimal control from this example into the objective functional, we find $J(u^*) = 2(1 - e^{-2})^{-1}$, whereas the value of $J(u^*)$ in Example 3.6.2 was 1. Not fixing the final state allows more freedom in the choice of controls, and the objective functional can be reduced further. The two optimal states are shown in Figure 3.11

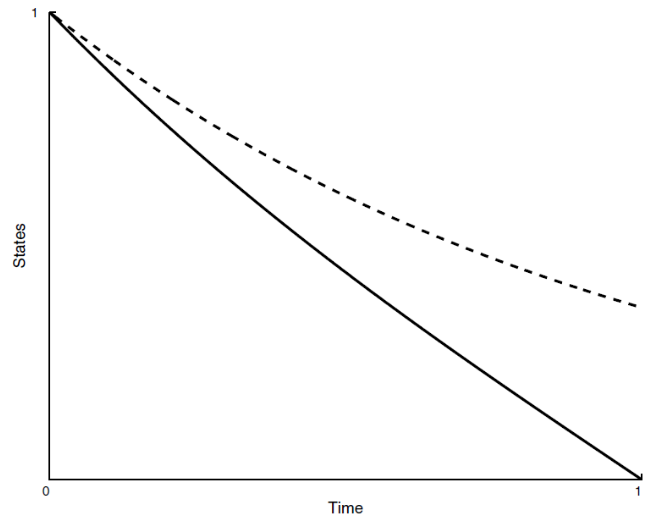


Figure 3.11 the optimal state from Example 3.6.4 (solid) is forced to 0. The optimal state from Example 3.6.2 (dashed) is not.

3.7 Bounded Controls

Many problems require bounds on the control to achieve a realistic solution. Suppose, for instance, that our control is the amount of a chemical used in a system. Then, clearly we require this amount to be nonnegative, i.e., $u \geq 0$. Often, the control must also be bounded above. Perhaps there are physical limitations on the amount of chemicals or environmental regulations which prohibit a certain level of use. We could also have a problem where the control is the percentage of some strength or use. Then $0 \leq u \leq 1$ would be our bounds.

Necessary Conditions:

In order to solve problems with bounds on the control, we must develop alternate necessary conditions. Consider the problem

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

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

$$a \leq u(t) \leq b$$

where a, b are fixed, real constants and $a < b$. Let $J(u)$ be the value of the objective functional at control u, where $x = x(u)$ is the associated state, namely,

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

Let u^*, x^* be an optimal pair. Let $h(t)$ be a piecewise continuous function where there exists a positive constant ϵ_0, such that for all $\epsilon \in (0, \epsilon_0)$, $u^\epsilon(t) = u^*(t) + \epsilon h(t)$ is admissible, i.e.,

$$a \le u^\epsilon(t) \le b \; for \; all \; t$$

Due to bounds on the controls, the derivative of the objective functional may not be zero at the optimal control, since u^* may be at the bounds (endpoints of its range) at some points in time; we may only know the sign of this derivative. To calculate this sign, we also restrict the sign of the ϵ parameter. Let $x^\epsilon(t)$ be the corresponding state variable for each $\epsilon \in (0, \epsilon_0]$. Precisely as was done in section above, introduce a piecewise differentiable adjoint variable ¸$\lambda(t)$ and apply the fundamental theorem of calculus to write $J(u^\epsilon)$ as

$$J(u^\epsilon) = \int_{t_0}^{t_1} [\, f(t, x^\epsilon, u^\epsilon) + \lambda(t) g(t, x^\epsilon, u^\epsilon) + x^\epsilon(t)\lambda'(t)]dt$$
$$-\lambda(t_0)x_0 + \lambda(t_1)x^\epsilon(t_1) + \emptyset(x(t_1)) \qquad (3.7.1)$$

As the maximum of $J(u)$ with respect to u at u^*

$$0 \ge \frac{d}{d\epsilon} J(u^\epsilon)\Big|_{\epsilon=0} = \lim_{\epsilon \to 0} \frac{J(u^\epsilon) - J(u^*)}{\epsilon} \qquad (3.7.2)$$

Note, the constant ϵ was chosen to be positive, so the limit can only be taken from one side. The numerator is clearly non-positive, as u^* is maximal. This gives the inequality shown, instead of equality as in section above. However, this is all we will need. As we did before, choose the adjoint variable so that

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

$$\lambda(t_1) = \emptyset'\big(x^*(t_1)\big)$$

Then (3.8.1),(3.8.2) reduce to

$$0 \geq \int_{t_0}^{t_1} (f_u + \lambda f_u) h dt \qquad (3.7.3)$$

and this inequality holds for all *h* as described above.

Let *s* be a point of continuity of u^* with $a \leq u^*(s) \leq b$.suppose $f_u + \lambda g_u > 0$ at *s*. As u^* is continuous at *s*, so is $f_u + \lambda g_u$. Thus, there is a small interval *I*, containing *s*, on which $f_u + \lambda g_u$ is strictly positive and $u^* < b$. Let

$$M = max\{u^*(t) : t \in I\} < b$$

Define a particular *h* by

$$h(t) = \begin{cases} b - m & if\ t \in I, \\ 0 & if\ t \in I. \end{cases}$$

Note, $h > 0$ on *I*. Further, it is easily seen that $a \leq u^* + \epsilon h \leq b$ for all$\epsilon \in [0,1]$ But,

$$\int_{t_0}^{t_1} (f_u + \lambda f_u) h dt = \int_I (f_u + \lambda f_u) h dt > 0$$

which contradicts (3.8.3). So

$f_u + \lambda g_u \geq 0\ at\ s$. Further, this holds for all points of continuity *s*. In summary,

$$\begin{aligned} u^*(t) = a \quad & implies\ f_u + \lambda g_u \leq 0\ at\ t, \\ a < u^*(t) < b \quad & implies\ f_u + \lambda g_u = 0\ at\ t, \\ u^*(t) = b \quad & implies\ f_u + \lambda g_u \geq 0\ at\ t. \end{aligned} \qquad (3.7.4)$$

The conditions (3.8.4) are equivalent to

$$\begin{aligned} f_u + \lambda g_u < 0\ at\ t \quad & implies\ \ u^*(t) = a \\ f_u + \lambda g_u = 0\ at\ t \quad & implies\ \ a \leq u^*(t) \leq b \\ f_u + \lambda g_u > 0\ at\ t \quad & implies\ \ u^*(t) = b \end{aligned} \qquad (3.7.5)$$

This holds for all points of continuity t of u^*. As they are irrelevant to the objective functional and the state equation, we neglect the remaining points. These new necessary conditions can be compiled as before.

Forming the Hamiltonian

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

the necessary conditions for x^* and λ ¸ are unchanged, namely

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

$$\lambda'(t) = -\frac{\partial H}{\partial x} \;\; , \qquad \lambda(t_1) = \emptyset'\big(x(t_1)\big).$$

It follows from the derivation above

$$\begin{cases} u^* = u \;\; if \;\; \frac{\partial H}{\partial u} < 0 \\ a \leq u^* \leq b = u \;\; if \;\; \frac{\partial H}{\partial u} = 0 \\ u^* = b \;\; if \;\; \frac{\partial H}{\partial u} > 0 \end{cases} \tag{3.7.6}$$

Example (3.7.1)

$$\max_u \int_0^2 [2x(t) - 3u(t) - u(t)^2] \, dt$$

Subject to $x'(t) = x(t) + u(t),\; x(0) = 5$

$$0 \leq u(t) \leq 2$$

Form the Hamiltonian

$$H = 2x - 3u - u^2 + x\lambda + u\lambda$$

Then the adjoint calculation yields :

$$\lambda'(t) = -\frac{\partial H}{\partial x} = -2 - \lambda \Longrightarrow \lambda = -2 + c_1 e^{-t}$$

$$\lambda(2) = 0 \Longrightarrow c_1 = 2e^2 \;\Longrightarrow\; \lambda(t) = 2e^{2-t} - 2$$

Now that we have found the adjoint value, we turn our attention to u^*, which requires considering the sign of $\frac{\partial H}{\partial U}$:

$$\frac{\partial H}{\partial u} = -2 - 3u - \lambda$$

$$0 > \frac{\partial H}{\partial u}\ at\ t \Rightarrow u(t) = 0 \Rightarrow 0 > -3 + \lambda = -3 + (2e^{2-t} - 2)$$

$$\Longrightarrow t > 2 - \ln({}^{5}/_{2})$$

$$0 < \frac{\partial H}{\partial u}\ at\ t \Rightarrow u(t) = 2 \Rightarrow 0 < -3 - 2(2) + \lambda = -7(2e^{2-t} - 2)$$

$$t < 2 - \ln\left({}^{9}/_{2}\right)$$

$$0 = \frac{\partial H}{\partial u}\ at\ t \Rightarrow u(t) = \frac{1}{2}(\lambda - 3) \Rightarrow 0 \leq \frac{1}{2}(\lambda - 3) \leq 2$$

$$2 - \ln\left({}^{9}/_{2}\right) \leq t \leq 2 - \ln\left({}^{5}/_{2}\right).$$

Hence, the optimal control is

$$u^*(t) = \begin{cases} 2\ \ when\ \ 0 \leq t < 2 - \ln\left({}^{9}/_{2}\right) \\ e^{2-t} - {}^{5}/_{2}\ when\ \ 2 - \ln\left({}^{9}/_{2}\right) \leq t \leq 2 - \ln\left({}^{5}/_{2}\right) \\ 0\ \ when\ \ 2 - \ln\left({}^{5}/_{2}\right) \leq t \leq 2\,. \end{cases}$$

To find the optimal state, insert the values for *u*¤ into the differential equation for *x*, and solve the three cases. We find the optimal state to be

$$x^*(t) = \begin{cases} k_1 e^t - 2\ when\ \ 0 \leq t < 2 - \ln({}^{9}/_{2}) \\ k_2 e^t - \frac{1}{2}e^{2-t} + {}^{5}/_{2}\ when\ \ 2 - \ln({}^{9}/_{2}) \leq t \leq 2 - \ln({}^{5}/_{2}) \\ k_3\, e^t \qquad when\ \ 2 - \ln({}^{5}/_{2}) \leq t \leq 2\,. \end{cases}$$

where k_1, k_2, and k_3 are constants. Using $x(0) = 5$, it follows $k_1 = 7$. Recall, the state must be continuous. So, requiring x^* to agree at

$$t = 2 - \ln\left({}^{9}/_{2}\right) \text{ and } t = 2 - \ln\left({}^{5}/_{2}\right)$$

, we find values for k_1 and k_3, so that

$$x^*(t) = \begin{cases} 7e^t - 2 \;\; when \;\; 0 \le t < 2 - \ln\left({}^{9}/_{2}\right) \\ \left(7 - \frac{81}{8}\, e^{-2}\right) e^t - \frac{1}{2} e^{2-t} + {}^{5}/_{2} \;\; when \;\; 2 - \ln\left({}^{9}/_{2}\right) \le t \le 2 - \ln\left({}^{5}/_{2}\right) \\ (7 - 7e^{-2})\, e^t \qquad when \;\; 2 - \ln\left({}^{5}/_{2}\right) \le t \le 2\,. \end{cases}$$

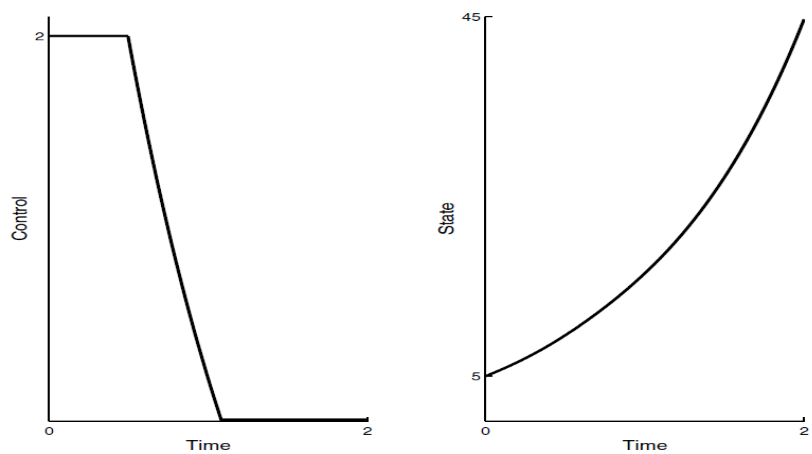


Figure 3.12 optimal control and state

Example (3.7.2)

This example deals with a one-sided control constraint.

$$\max_u \; x(4) - \int_0^4 u(t)^2 dt$$

$$\text{Subject to } \; x'(t) = x(t) + u(t)\,, \;\; x(0) = 0$$

$$x(t) \le 5.$$

The Hamiltonian in this problem is

$$H = -u^2 + \lambda x + \lambda\, u$$

Recall, the payoff term $\emptyset = x(4)$ is not included in the Hamiltonian, but instead incorporated into the transversality condition. Specifically, since $\emptyset(x) = x \; and \; \emptyset' = 1$, we have

$$\lambda'(t) = -\frac{\partial H}{\partial x} = -\lambda$$

$$\lambda(4) = \emptyset'\big(x^*(4)\big) = 1.$$

This gives

$$\lambda(t) = e^{4-t}.$$

If you refer back to equation (3.8.5), you will see $\frac{\partial H}{\partial u} < 0$ implies u^* is at the lower bound. However, we have no lower bound in this problem. The control u can range over all values less than or equal to 5. So, $\frac{\partial H}{\partial u} < 0$ cannot occur. To find a representation of u^*, we need only consider the other two cases:

$$\frac{\partial H}{\partial u} = \lambda - 2u$$

$$\frac{\partial H}{\partial u} > 0 \Longrightarrow u^*(t) = 5 \Longrightarrow \lambda - 10 > 0 \Longrightarrow e^{4-t} > 10 \Longrightarrow t < 4 - \ln(10)$$

$$\frac{\partial H}{\partial u} = 0 \Longrightarrow u^*(t) \leq 5 \Longrightarrow e^{4-t} = \lambda = 2u \leq 10 \Longrightarrow t \geq 4 - \ln(10)$$

Hence, the above two cases give

$$u^* = \begin{cases} 5 & when \; 0 \leq t < 4 - \ln(10) \\ \frac{1}{2} e^{4-t} & when \; 4 - \ln(10) \leq t \leq 4 \end{cases}$$

To finish the example, we simply solve the state equation to find x^*

$$x'(t) = x + 5 \;, x(0) = 0 \Longrightarrow x(t) = 5e^t - 5 \; on \; [0,4 - \ln(10)]$$

$$x'(t) = x + \frac{1}{2} e^{4-t} \Longrightarrow x(t) = -\frac{1}{4} e^{4-t} + ke^t \; on \; [4 - \ln(10) \;,4]$$

for some constant k. We require that x^* be continuous, so these two expressions must agree at $t = 4 - \ln(10)$. This gives $k = 5 - 25e^{-4}$. Hence

$$x^* = \begin{cases} 5e^t - 5 & when\ 0 \le t \le 4 - \ln(10) \\ -\frac{1}{4}e^{4-t} + (5 - 25e^{-4})e^t & when\ 4 - \ln(10) \le t \le 4. \end{cases}$$

3.8 Optimal Control of Several Variables

We have only examined problems with one control and one dependent state variable. Often, though, we will wish to consider more variables. For example, consider a system modeling antibiotics used to fight a viral infection. In addition to the number of viral particles in the blood, we might also want to follow the number of antibodies or white blood cells. These quantities would be represented as additional state variables. Further, suppose the patient was taking two different antibiotics that caused the body to generate antibodies at different rates or times. These would need to be separate control variables; see [20]. Further, we could examine an SIR epidemic model with vaccination levels as a control [15, 16, 17, 18], or a tuberculosis epidemic model involving decisions in allocating efforts [19].

Necessary Conditions

The methods developed for one control and state are easily extended to optimal control of multiple state and control variables. Consider a problem with n state variables, m control variables, and a payoff function $\emptyset'$

$$\max_{u_1,\dots\dots,u_m} \int_{t_0}^{t_1} f(t, x_1(t), \dots\dots, x_n(t), u_1(t), \dots\dots, u_m(t))\, dt$$
$$+ \emptyset\big(x_1(t_1), \dots\dots,, x_n(t_1)\big)$$

Subject to $x_i'(t) = g_i\big(t, x_1(t), \dots\dots,, x_n(t), u_1(t), \dots\dots, u_m(t)\big)$

$$x_i(t_0) = x_{i0} \quad for\ i = 1,2, \dots\dots, n$$

Where the functions f, g_i are continuously differentiable in all variables. We make no requirements on m, n. In fact, $m < n$, $m = n$, or $m > n$ are all acceptable. Use vector notation to change the problem to a more familiar form.

Let $x^{\rightarrow}(t) = [x_1(t), \dots \dots, x_n(t)]$, $u^{\rightarrow}(t) = [u_1(t), \dots \dots, u_m(t)]$, $x_0{}^{\rightarrow}(t) = [x_{10}, \dots \dots, x_{n0}]$ and $g^{\rightarrow}(t, x^{\rightarrow}, u^{\rightarrow}) = [\, g_1(t, x^{\rightarrow}, u^{\rightarrow}), \dots \dots, g_n(t, x^{\rightarrow}, u^{\rightarrow})]$. Then, we can write the problem as

$$\max_{u^{\rightarrow}} \int_{t_0}^{t_1} f\big(t, x^{\rightarrow}(t), u^{\rightarrow}(t)\big) dt + \emptyset(x^{\rightarrow}(t_1))$$

$$\text{Subject to } x^{\rightarrow\prime}(t) = g^{\rightarrow}\big(t, x^{\rightarrow}(t), u^{\rightarrow}(t)\big) \;, x^{\rightarrow}(t_0) = x_0{}^{\rightarrow} \,.$$

Let $u^{\rightarrow *}$ be a vector of optimal control functions and $x^{\rightarrow *}$ be the vector of corresponding optimal state variables. With n states, we will need n adjoints, one for each state. Introduce a piecewise differentiable vector-valued function $\lambda^{\rightarrow}(t) \;=\; [\lambda_1(t), \dots \dots \lambda_n(t)]$, where each ¸$\lambda_i$ is the adjoint variable corresponding to x_i. Define the Hamiltonian

$$H(t, x^{\rightarrow}, u^{\rightarrow}, \lambda^{\rightarrow}) = f(t, x^{\rightarrow}, u^{\rightarrow}) + \lambda^{\rightarrow}(t).\, g(t, x^{\rightarrow}, u^{\rightarrow})$$

where . is the dot product of vectors. By essentially the same argument presented in section above, we find the variables satisfy identical optimality, adjoint, and transversality conditions in each vector component. Namely, $u^{\rightarrow *}$ maximizes $H(t, x^{\rightarrow *}, u^{\rightarrow}, \lambda^{\rightarrow})$ with respect to $u^{\rightarrow}$ at each t, and $u^{\rightarrow *}$, $x^{\rightarrow *}$, and $\lambda^{\rightarrow}$ ¸ satisfy

$$x_i{}'(t) = \frac{\partial H}{\partial \lambda_i} = g_i(t, x^{\rightarrow}, u^{\rightarrow}),\;\; x_i(t_0) = x_{i0} \; for \; i = 1, \dots \dots, n$$

$$\lambda_j{}'(t) = -\frac{\partial H}{\partial x_j} \;\;,\;\; \lambda_j(t_1) = \emptyset_{x_j}\, (x^{\rightarrow}(t_1) \; for \; j = 1, \dots \dots, n$$

$$0 = \frac{\partial H}{\partial u_k} \; at \; u_k{}^{*} \; for \; k = 1, \dots \dots, m$$

Where

$$H(t, x^{\rightarrow}, u^{\rightarrow}, \lambda^{\rightarrow}) = f(t, x^{\rightarrow}, u^{\rightarrow}) + \sum_{i=1}^{n} \lambda_i{}^{\rightarrow}(t) g_i(t, x^{\rightarrow}, u^{\rightarrow}).$$

By $\emptyset_{x_j}$, it is meant the partial derivative in the x_j component. Note, if $\emptyset = 0$, then $\lambda_j(t_1) = 0$ for all j, as usual.

Modifications of the problems yield adjustments on the conditions similar to those in previous chapters. For example, if a particular state variable x_i satisfies $x_i(t_0) = x_{i0}$, $x_i(t_1) = x_{i1}$ both fixed, then the corresponding adjoint λ_i has no boundary conditions. Similarly, if bounds are placed on a control variable $a_k \leq u_k \leq b_k$ then the optimality condition is changed from.

$$\frac{\partial H}{\partial u_k} = 0 \ to \qquad \begin{cases} u_k = a_k & if \ \frac{\partial H}{\partial u_k} < 0, \\ a_k \leq u_k \leq b_k & if \ \frac{\partial H}{\partial u_k} = 0, \\ u_k \leq b_k & if \ \frac{\partial H}{\partial u_k} > 0. \end{cases}$$

We illustrate these ideas with a few examples.

Example (3.8.1)

$$\min_u \int_0^1 (x_2(t) + u(t)^2) dt$$

$$\text{Subject to } x_1{}'(t) = x_2(t), x_1(0) = 0, x_1(1) = 1$$

$$x_2{}'(t) = u(t), x_2(0) = 0$$

Introduce two adjoint variables, one for each state variable, and form the Hamiltonian,

$$H = x_2 + u^2 + \lambda_1 x_2 + \lambda_2 u$$

Form the adjoint and transversality conditions

$$\lambda'_1(t) = -\frac{\partial H}{\partial x_1} = 0$$

$$\lambda'_2(t) = -\frac{\partial H}{\partial x_2} = -\lambda'_1 - 1, \lambda'_2(1) = 0$$

The first adjoint λ'_1 is simply a constant, say *C*. Then, λ'_2 can be solved as follows,

$$\lambda'_1(t) = C$$

$$\lambda'_2(t) = -(C+1)(t-1).$$

Using the optimality condition,

$$0 = \frac{\partial H}{\partial u} = 2u + \lambda_2 \Longrightarrow u^* = -\frac{\lambda_2}{2} = \frac{(C+1)}{2}(t-1).$$

Finally, we make use of the state equations and boundary conditions to find

$$x'_2 = u \Longrightarrow x_2(t) = \frac{(C+1)}{2}\left(\frac{t^2}{2} - t\right), as\ x_2(0) = 0$$

$$x'_1 = x_2 \Longrightarrow x_1(t) = \frac{(C+1)}{2}\left(\frac{t^3}{6} - \frac{t^2}{2}\right), as\ x_2(0) = 0\ .$$

Noting that $x_1(1) = 1$, it follows $C = -7$. Thus, the optimal solution set is

$$u^*(t) = 3 - 3\mathrm{t},$$

$$x_1{}^*(t) = \frac{3}{2}t^2 + \frac{1}{2}t^3$$

$$x_2{}^*(t) = 3t + \frac{3}{2}t^2.$$

The optimal states $x_1{}^*$ and $x_2{}^*$are shown in Figure 3.13

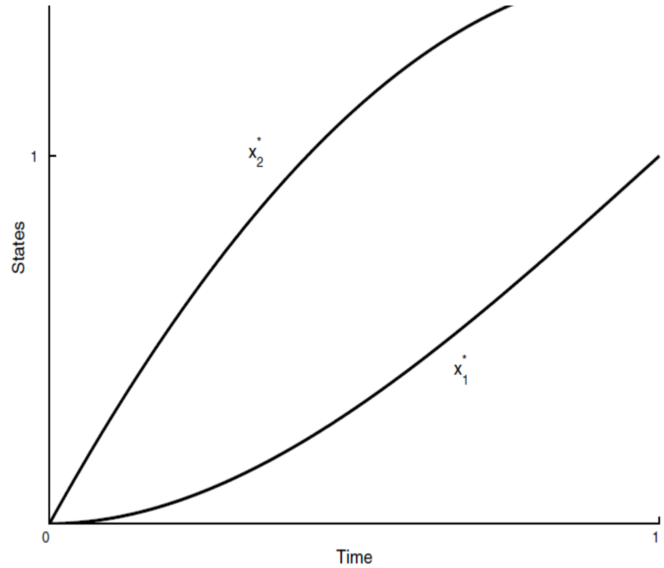


Figure 3.13 optimal state for example 3.9.1

Example (3.8.2)

$$\max_{u_1,u_2} \int_0^1 \left(x(t) - \frac{1}{8}u_1(t)^2 - \frac{1}{2}u_2(t)^2\right) dt$$

$$\text{Subject to } x'(t) = u_1(t) + u_2(t)\,,\;\; x(0) = 0$$

$$1 \leq u_1(t) \leq 2$$

The Hamiltonian is

$$H = x - \frac{1}{8}{u_1}^2 - \frac{1}{2}{u_2}^2 + \lambda u_1 + \lambda u_2.$$

The adjoint and transversality conditions yield

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

The second control has no bounds, so we can easily solve for it,

$$0 = \frac{\partial H}{\partial u_2} = -u_2 + \lambda \Longrightarrow {u_2}^* = \lambda = 1 - t.$$

To find ${u_2}^*$, we note

$$\frac{\partial H}{\partial u_1} = \lambda - \frac{u_1}{4}$$

$$\frac{\partial H}{\partial u_1} < 0 \Rightarrow u_1(t) = 1 \Rightarrow 1 - t < \frac{1}{4} \Rightarrow t > \frac{3}{4}$$

$$\frac{\partial H}{\partial u_1} = 0 \Rightarrow u_1{}^* = 4\lambda = 4 - 4t \Rightarrow 1 \leq 4 - 4t \leq 2 \Rightarrow \frac{1}{2} \leq t \leq \frac{3}{4}$$

$$\frac{\partial H}{\partial u_1} > 0 \Rightarrow u_1(t) = 2 \Rightarrow 1 - t > \frac{1}{2} \Rightarrow t < \frac{1}{2}$$

By plugging the three cases back into the state equation, and requiring continuity, we can find x^* Then, the optimal solution set (Figure 3.14) is

$$u_1{}^*(t) = \begin{cases} 2 & 0 \leq t < \frac{1}{2} \ , \\ 4 - 4t & \frac{1}{2} \leq t \leq \frac{3}{4} \ , \\ 1 & \frac{3}{4} < t \leq 1 \ , \end{cases} \qquad , u_1{}^*(t) = 1 - t$$

$$x^*(t) = \begin{cases} 3t - \frac{1}{2}t^2 & 0 \leq t \leq \frac{1}{2} \ , \\ 5t - \frac{5}{2}t^2 - \frac{1}{2} & \frac{1}{2} \leq t \leq \frac{3}{4} \ , \\ 2t - \frac{1}{2}t^2 + \frac{5}{8} & \frac{3}{4} < t \leq 1 \ , \end{cases}$$

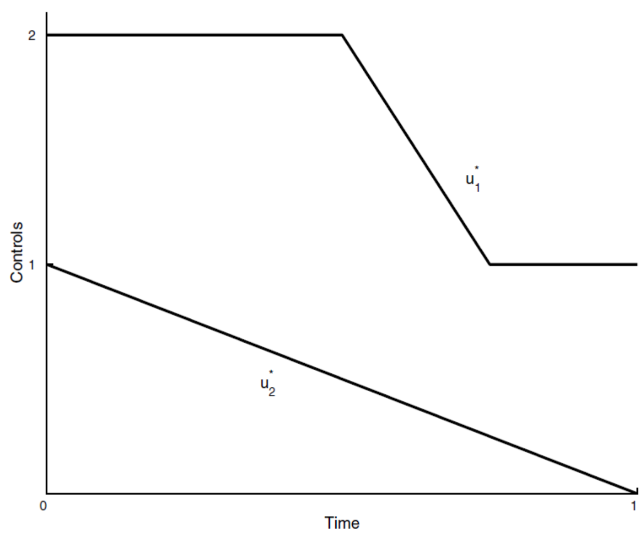


Figure 3.14 Optimal control for example 3.8.2

3.8.1 Linear Quadratic Regulator Problems

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. For example, one might use such systems to model chemostats [21]. Our state system is given by

$$x'(t) = A(t)x(t) + B(t)u(t) \tag{3.8.1}$$

Where *x* is an *n*-dimensional column vector and *u* is a *m*-dimensional column vector. The matrices *A*(*t*),*B*(*t*) have sizes *n* x *n* and *n* x *m* respectively.

Note that entries of matrices of *A*,*B* can be functions of time. The objective functional is

$$J(u) = \frac{1}{2}\left[x^T(T)M_x(T) + \int_0^T x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)dt\right] \tag{3.8.2}$$

Where the symmetric matrices *M*, *Q*(*t*), and *R*(*t*) are sizes *n* x *n*, and *m* x *m* respectively, with *M*, *Q*(*t*) being positive semidefinite and *R*(*t*) being positive defnite for all $0 \leq t \leq T$.

The positive defnite property guarantees *R*(*t*) is invertible.

The superscript T refers to transpose of the matrix. We can interpret the objective functional as minimizing a weighted sum of the components of the state and the control. The matrices would be chosen to decide which components to emphasize. In practice, the state might be the difference between a quantity (like the levels of microorganisms in a chemo stat) and its desired profile, and the objective functional can drive certain components of the quantity close to the profile.

Like the control and state, we write λ to mean an *n*-dimensional column vector of adjoints. The Hamiltonian becomes

$$H = \frac{1}{2}x^T Qx + \frac{1}{2}u^T Ru + \lambda^T (Ax + Bu).$$

Some care must be taken in differentiating matrix expressions, particularly if not familiar with the process. We suppress the details here, but encourage the reader to check the calculations term-by-term. The optimality equation is

$$Ru + B^T\lambda = 0 \Longrightarrow u^* = -R^{-1}B^T\lambda$$

and the adjoint equation is

$$\lambda' = -Qx - A^T\lambda\,, \quad \lambda(T) = Mx(T).$$

The assumptions of symmetry for *M*, *Q*, and *R* are buried in the above calculations. We choose to solve this problem in a different way due to the structure of the transversality condition and the adjoint differential equation; this method is called the sweep method [24, 22, 23]. Instead of using *λ*, we find a matrix function *S*(*t*) such that λ(*t*) = *S*(*t*)*x*(*t*). By the product rule for matrices,

$$\lambda'^{(t)} = S'(t)x(t) + s(t)x'^{(t)}.$$

Using the expressions for λ' and *x*'given by the state and adjoint equations, we have

$$-Qx - A^T\lambda = S'x + SAx + SBu.$$

Making use of the characterization of the control and the identity *λ* = *Sx*,

$$\begin{aligned} -S'(t) &= Qx + A^T\lambda + SAx - SBR^{-1}B^{-1}\lambda \\ &= Qx + A^T Sx + SAx - SBR^{-1}B^{-1}Sx \\ &= [Q + A^T S + SA - SBR^{-1}B^T S]x. \end{aligned}$$

From the transversality condition, we obtain the matrix *Riccati* equation that *S*(*t*) must satisfy. Namely,

$$-S' = A^T S + SA - SBR^{-1}B^T S + Q\,, \qquad S(T) = M.$$

Reconsidering the characterization, we see the control is a linear function of the state only, a type of feedback control

$$u = -R^{-1}B^T Sx.$$

The matrix $R^{-1}B^T S$ is called the *gain*. After solving the Riccati matrix equation for *S*, the control is given by an equation in *x*, and *x* is given by an ODE in *u*, so that the problem can be solved using simple ODE methods. Therefore, we have totally eliminated the adjoint ¸ from the problem. See the book by Morris about feedback control [26] and a recent application of
the Riccati approach [20].

Example (3.8.3)

We consider a simple one dimensional example.

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

Subject to $x'(t) = u(t)\,, x(0) = x_0$.

In this case, all the matrices are scalars (size 1 £ 1) and *S*(*T*) = *M* = 0, *A* = 0, *B* = *Q* = *R* = 1

The Riccati equation is

$$-S' = 1 - S^2 \ , \qquad S(T) = 0$$

Solving as a separable equation, and using partial fractions,

$$\frac{1}{2}Ln\left|\frac{S-1}{S+1}\right| = \int \frac{S'}{S^2-1}dt = \int dt = t + c.$$

Which along with *S*(*T*) = 0 gives

$$S(t) = \frac{1 - e^{2(t-T)}}{1 + e^{2(t-T)}}$$

The optimal control satisfies $u = -Sx$, so that the optimal state satisfies $x' = -Sx$.

Using partial fractions (or an integral table) we can find an antiderivative of S, and solve the separable equation to see

$$x(t) = C(e^{t-T} + e^{T-t}).$$

Taking into account $x(0) = x_0$

$$x^*(t) = x_0 \frac{e^t + e^{2T-t}}{1 + e^{2T}} \ \ and\ u^*(t) = \ x_0 \frac{e^t - e^{2T-t}}{1 + e^{2T}}.$$

3.8.2 Higher Order Differential Equations

Optimal control of systems can be employed to solve maximization (or minimization) problems involving higher order differential equations. Consider the following problem,

$$\max_{u_1, \ldots\ldots, u_m} \int_{t_0}^{t_1} f(\ t\ , x(t), x'(t), \ldots \ldots x^n(t), u_1(t), \ldots \ldots, u_m(t))\, dt$$

$$\text{Subject to } x^{n+1}(t) = g(t\ , x(t), x'(t), \ldots \ldots x^n(t), u_1(t), \ldots \ldots, u_m(t))$$

$$x(t_0) = \alpha_1\ , \qquad x'(t_0) = \alpha_2\ , \ldots \ldots \ldots \ldots, x^n(t_0) = \alpha_n + 1\ for\ n > 1$$

Pontryagin's Maximum Principle, as we have developed it, does not directly deal with this type of problem. However, it is easily converted to a systems problem by introducing n+1 state variables defined by

$x_1(t) = x(t), x_2(t) = x'(t), \ldots \ldots, x_{n+1}(t) = x^n(t)$. Then, the above problem becomes

$$\max_{u_1, \ldots\ldots, u_m} \int_{t_0}^{t_1} f(\ t\ , x(t), x'(t), \ldots \ldots, x_{n+1}(t), u_1(t), \ldots \ldots, u_m(t))dt$$

$$\text{Subject to } \ x'_1(t) = x_2(t), x_1(t_0) = \alpha_1$$

$$x'_2(t) = x_3(t), x_2(t_0) = \alpha_1$$

$$\vdots$$

$$\vdots$$

$$x'_n(t) = x_{n+1}(t), x_n(t_0) = \alpha_n$$

$$x'_{n+1}(t) = g\big(t, x_1(t), x_2(t), \dots\dots, x_{n+1}(t), u_1(t), \dots\dots, u_m(t)\big)$$

$$x_{n+1}(t_0) = \alpha_{n+1}$$

Example (3.8.4)

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

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

Let $x_1 = x \ and \ x_2 = x'$, to covert the problem to

$$\min_u \frac{1}{2}\int_0^{\pi} (u(t)^2 - x_1(t)^2)dt$$

Subject to $x'_1(t) = x_2(t), x_1(0) = 1,$

$$x'_2(t) = u(t), x_2(0) = 1.$$

Introduce two adjoint variables $\lambda_1 \ and \ \lambda_2$ and set up the Hamiltonian:

$$H = \frac{1}{2}u^2 - \frac{1}{2}{x_1}^2 + \lambda_1 x_2 + \lambda_2 u$$

$$0 = \frac{\partial H}{\partial u} = u + \lambda_2 \Longrightarrow u^* = -\lambda_2$$

$$\lambda'_1(t) = -\frac{\partial H}{\partial x_1} = x_1\ , \lambda_1(\pi) = 0,$$

$$\lambda'_2(t) = -\frac{\partial H}{\partial x_2} = -\lambda_1\ , \lambda_2(\pi) = 0.$$

Note ${\lambda_2}^{(4)} = -{\lambda_1}''' = -{x_1}'' = -x_2' = -u = \lambda_2$. Thus

$\lambda_2(t) = Ae^t + Be^{-t} + Ccost + Dsint$ for some constant A,B,C,D. Making use of the adjoint and state equation , we see

$$x_1(t) = -Ae^t - Be^{-t} + Ccost + Dsint,$$

$$x_2(t) = -Ae^t + Be^{-t} - Csint + Dcost,$$

$$\lambda_1(t) = -Ae^t + Be^{-t} - Csint - Dcost,$$

$$\lambda_2(t) = Ae^t + Be^{-t} + Ccost + Dsint.$$

Using the condition $x_1(0) = x_2(0) = 1$ and $\lambda_1(\pi) = \lambda_2(\pi) = 0$,we find

$$\begin{pmatrix} -1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 \\ e^{\pi} & e^{-\pi} & 0 & 1 \\ e^{\pi} & e^{-\pi} & 0 & -1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Approximate value are $\approx 0.0452, B = 0, C = D \approx 1.0452$, So that the optimal solution are

$$x^*(t) = x_1(t) = -0.0452e^t + 1.0452(\cos(t) + \sin(\text{t})),$$

$$u^*(t) = -\lambda_2(t) = -0.0452e^t - 1.0452(\cos(t) + \sin(\text{t})).$$

3.9 Linear Dependence on the Control

In the preceding section, we have examined increasingly more general optimal control problems. However, we now turn our attention to a special case, which often arises in applications. Specifically, we focus on problems that are linear in the control *u*. The method of solving such problems is sometimes quite different, and the optimal solution often involves discontinuities in u^*.

3.9.1 Bang-Bang Controls

Consider the optimal control problem

$$\max_u \int_{t_0}^{t_1} (f_1(t,x) + u(t)f_2(t,x))dt$$

$$\text{Subject to} \quad x'(t) = g_1(t,x) + u(t)g_2(t,x), x(0) = x_0$$

$$a \le u(t) \le b.$$

Notice the integrand function *f* and the right-hand side of the differential equation *g* are both linear functions of the variable *u*. Thus, the Hamiltonian is also a linear function of *u*, and can be written

$$H = [f_1(t,x) + \lambda(t)g_1(t,x)] + u(t)[f_2(t,x) + \lambda(t)g_2(t,x)]$$

The necessary condition $\lambda'(t) = -\frac{\partial H}{dx}$ is as normal. However, the optimality condition

$$\frac{\partial H}{du} = f_2(t,x) + \lambda(t) g_2(t,x)$$

Contains no information on the control. We must try to maximize the Hamiltonian *H* with respect to *u* using the sign of $\frac{\partial H}{du}$, but, when $f_2(t,x) + \lambda g_2(t,x) = 0$, we cannot immediately find a characterization of u^*.

Define $\varphi(t) = f_2(t, x(t)) + \lambda(t) g_2(t, x(t))$, usually calld the *switching function*. Our characterization of u^* is

$$u^*(t) = \begin{cases} a & if \varphi(t) < 0 \\ ? & if\ \varphi(t) = 0 \\ b & if\ \varphi(t) > 0 \end{cases}$$

If $\varphi(t) = 0$ cannot be sustained over an interval of time, but occurs only at finitely many points, then the control u^* is referred to as *bang-bang*. In this case, it is piecewise constant function, switching between only the upper and lower bounds. An example of such a control is given in Figure 3.15. The switches coincide with the places where φ switches signs (so that $\varphi(t) = 0$), hence the name switching function. The actual points where this occurs are called *switching times*.

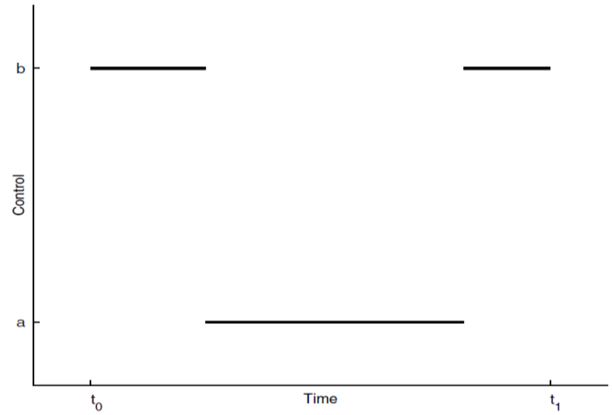


FIGURE 3.15 a typical bang-bang control.

If $\varphi(t) = 0$on some interval of time, we say u^* is *singular* on that interval. A characterization of u^* on this interval must be found using other information. The endpoints of this interval are sometimes called switching times as well. We postpone the discussion of singular controls until the next section.

To solve a bang-bang problem numerically, the forward-backward sweep method can be employed. First, it must be analytically proven that the problem is in fact bang-bang, i.e., $\varphi(t) \equiv 0$ over an interval is impossible. Once this is established, the code is written as usual, where the characterization of *u* is given by

```
for i=1:N+1
    temp = psi(t(i),x(i),lambda(i))
    if(temp < 0)
        u1(i) = a;
    else
        u1(i) = b;
    end
end
u = 0.5*(u1 + oldu);
```

Where $psi(t(i), x(i), lambda(i))$ in the second line is replaced by the actual value of the function φ in terms of *t*, *x*, and λ, according to the specific problem. Notice, even though it is irrelevant from an analytical standpoint, the value of the control at the switching times must be assigned in our MATLAB code. Here, we have arbitrarily assigned *u* = b when $\varphi = 0$. Assigning *u* = a would have been just as prudent. Defining *u* to be the average of *a* and *b* at these points is also used. It usually makes little difference. Finally, note the convex combination remains as before. This hastens the finding of the switching times. Here, we consider a few bang-bang examples which can be solved by hand.

Example (3.9.1)

$$\max_u \int_0^2 e^t(1-u(t))dt$$

$$\text{Subject to} \quad x'(t) = u(t)x(t), x(0) = 1,$$

$$0 \le u(t) \le 1$$

The objective functional here does not depend on *x*, and the state does not have a terminal time condition. Therefore, looking at the format of the integrand of the objective functional we see that the optimal control should be 0.

The Hamiltonian is

$$H = e^t(1-u) + \lambda ux$$

The adjoint and transeversality conditions are

$$\lambda' = -\frac{\partial H}{dx} = -\lambda u, \quad \lambda(2) = 0$$

And

$$\varphi(t) = \frac{\partial H}{du} = -e^t + \lambda(t)x(t).$$

Suppose u^* is singular on some interval, i.e., $0 < u^* < 1$. Then, $\varphi(t) = 0$ on this interval, so that

$$e^t = \lambda x.$$

As this holds on an interval, we can differentiate both sides,

$$e^t = (\lambda x)' = \lambda' x + \lambda x' = -\lambda ux + \lambda ux = 0.$$

This is clearly impossible, so u^* is nowhere singular, thus bang-bang. Considering both possible values for u^*,

$$u^* = 0 \Rightarrow x' = 0 = \lambda' \Rightarrow x, \lambda \; constant,$$

$$u^* = 1 \Rightarrow \lambda' = -\lambda \Rightarrow \lambda(t) = ke^{-t},$$

for some constant *k*. Note, for $\lambda(t) = ke^{-t}$ to satisfy ¸$\lambda(2) = 0$, *k* must be zero. In the other case, ¸λ is constant. Hence, regardless of what the control is near *t* = 2,

$\lambda \equiv 0$, on some interval including t = 2. However, we require λ to be continuous, and it is impossible for ¸ $\lambda = ke^{-t}$ be continuously joined with $\lambda \equiv 0$ for non-zero k. Thus, $\lambda \equiv 0$ everywhere. It follows

$$\frac{\partial H}{du} = -e^t < 0 \; for \; all \; t$$

So that

$$u^* \equiv 0 \quad and \quad x^* \equiv 1.$$

Example (3.9.2)

$$\max_u \int_0^2 (2x(t) - 3u(t))dt$$

$$\text{Subject to} \quad x^{'}(t) = x(t) + u(t), \; x(0) = 5,$$

$$0 \le u(t) \le 2.$$

If we view this as a simple population model with exponential growth, we seek to increase the population as much as possible, while keeping the cost of control down.

The Hamiltonian is

$$H = 2x - 3u + \lambda x + \lambda u.$$

Using the necessary conditions and transversality condition, ¸can be immediately solved:

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

The switching function

$$\varphi(t) = \frac{\partial H}{du} = \lambda - 3 = 2e^{2-t} - 5$$

is clearly nowhere constant, thus not identically 0 over an interval. So, u^* is bang-bang, and

$$u^*(t) = 0 \Leftrightarrow \varphi < 0 \Leftrightarrow e^{2-t} < \frac{5}{2} \Leftrightarrow t > 2 - \ln(\frac{5}{2})$$

$$u^*(t) = 2 \Leftrightarrow \varphi > 0 \Leftrightarrow e^{2-t} > \frac{5}{2} \Leftrightarrow t < 2 - \ln\left(\frac{5}{2}\right).$$

For $0 \leq t > 2 - \ln(\frac{5}{2})$

$$u = 2 \Rightarrow x' = x + 2.$$

Along with $x(0) = 5$,this gives $x(t) = -2 + 7e^t$.on $2 - \ln\left(\frac{5}{2}\right) < t \leq 2$,

$$u = 0 \Rightarrow x' = x \Rightarrow k_0 e^t$$

for some constant k_0. As x must be continuous, the expressions $-2 + 7e^t$ and $k_0 e^t$ must agree at t= $2 - \ln\left(\frac{5}{2}\right)$. this gives $k_0 = 7 - 5e^t$. Hence, the optimal solutions are

$$u^* = \begin{cases} 2 \quad when \quad t \leq 2 - \ln\left(\frac{5}{2}\right) \\ 0 \quad when \quad t > 2 - \ln\left(\frac{5}{2}\right) \end{cases}$$

And

$$x^* = \begin{cases} 7e^t - 2 \quad when \quad t < 2 - \ln\left(\frac{5}{2}\right) \\ 7e^t - 5e^{t-2} \quad when \quad t > 2 - \ln\left(\frac{5}{2}\right) \end{cases}$$

The optimal state is shown in Figure 3.16

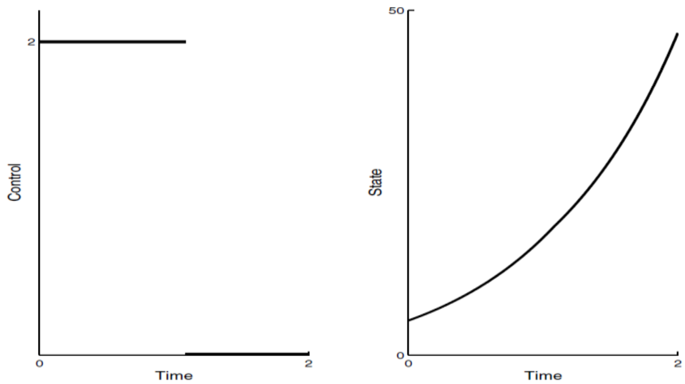


Figure 3.16: The optimal control and state for Example 3.9.2. The state appears differentiable here, but this is due to scale. It is in fact only continuous at $\boldsymbol{t = 2 - \ln\left(\frac{5}{2}\right)}$

3.9.2 Singular Controls

We now turn our attention towards singular controls, and in particular, a few examples. In the first example, the solution is relatively easy to guess. However, because it is singular, generating the optimal control via the necessary conditions is somewhat difficult.

Example (3.9.3)

$$\max_{u} \int_0^2 (x(t) - t^2)^2 dt$$

$$\text{Subject to} \quad x'(t) = u(t), \; x(0) = 1,$$

$$0 \le u(t) \le 4.$$

First, generate the necessary conditions as usual,

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

$$\lambda'(t) = -\frac{\partial H}{dx} = -2(x - t^2), \qquad \lambda(t) = 0,$$

$$\varphi = \frac{\partial H}{du} = \lambda.$$

If $\varphi \equiv 0$ on some interval, then

$$0 \equiv \lambda'(t) = -2(x - t^2) \implies x(t) = t^2,$$

So that on this interval

$$u = x' = 2t.$$

Hence, we obtain

$$u^* = \begin{cases} 0 & when\ \lambda > 0, \\ 2t & when\ \lambda = 0, \\ 4 & when\ \lambda < 0. \end{cases} \tag{3.9.1}$$

Our first goal is to establish that $x^*(t) \geq t^2$ on [0, 2]. Suppose not, i.e., suppose that

$x(t) < t^2$ Somewhere. Then, as $x(t) > t^2$ at $t = 0$, there must exist a $t_0 \in 2\ (0\ , 2)$ such that $x(t_0) \leq t_0^2$ and $u(t_0) = x'(t_0) < 2t_0$. Hence, from (3.9.1), it follows $u(t_0) = 0$ and $\lambda(t_0) > 0$.

Now, consider the points in time $t > t_0$ for which $\lambda(t) = 0$. We know at least one such point exits, namely $t = 2$. Let t_1 be the minimum of these points so that $\lambda(t_1) = 0$ but $\lambda(t) > 0$ for $t \in [t_0\ , t_1\ )$. Then, from (3.9.1), we see $u^* = 0$ on$[t_0\ , t_1\ )$. This implies $x^*(t) = x^*(t_0)$ on $[t_0\ , t_1\ )$. As we choose t_0 so that $x^*(t_0) \leq t_0^2$, we see $x^*(t) \leq t^2$ on $[t_0\ , t_1\ )$. Hence, by the adjoint equation, $\lambda'(t) \geq 0$ on $[t_0\ , t_1\ )$. But, if $\lambda(t_0) > 0$ and λ never decreases on this interval, then $\lambda(t_1) = 0$ is impossible. This gives our contradiction.

Thus, $x^*(t) \geq t^2$ on [0, 2].This immediately gives $\lambda' \leq 0$ on [0, 2]. As $\lambda(2) = 0$, we must have $\lambda \geq 0\ 0$ on [0, 2]. As ¸ is a non-negative, non-increasing function, there is some $k \in [0, 2]$ so that $\lambda > 0$ on $[0, k)$ and $\lambda = 0$ on $[k, 2]$.

Suppose $k = 0$. Then $\lambda = 0$ everywhere, so that ¸ $\lambda' = 0$ everywhere. But, $x^*(t) > t^2$ at t=0 so that ¸ $\lambda'(0) < 0$. Contradiction. Now suppose $k = 2$. Then, $u^* = 0$ everywhere, so that

$x^* = 1$ everywhere. This clearly contradicts $x^*(t) \geq t^2$. Hence, $0 < k < 2$, and we have

$$u^*(t) = \begin{cases} 0 & when \quad 0 \le t \le k, \\ 2t & when \quad 0 \le t \le 2 \end{cases}$$

$$x^*(t) = \begin{cases} 1 & when \quad 0 \le t \le k, \\ t^2 + (1 - k^2) & when \quad 0 \le t \le 2 \end{cases}$$

as x^* must be continuous. Finally, to find k, note that, $\lambda \equiv 0$ on $[k, 2]$, which implies

$$0 = \lambda'(t) = -2(x - k^2) \; on \; (k, 2) \Longrightarrow k = 1.$$

Hence, the optimal solution set (Figure 3.17) is

$$u^*(t) = \begin{cases} 0 & when \quad 0 \le t \le 1, \\ 2t & when \quad 0 \le t \le 2, \end{cases}$$

$$x^*(t) = \begin{cases} 1 & when \quad 0 \le t \le 1, \\ t^2 & when \quad 0 \le t \le 2. \end{cases}$$

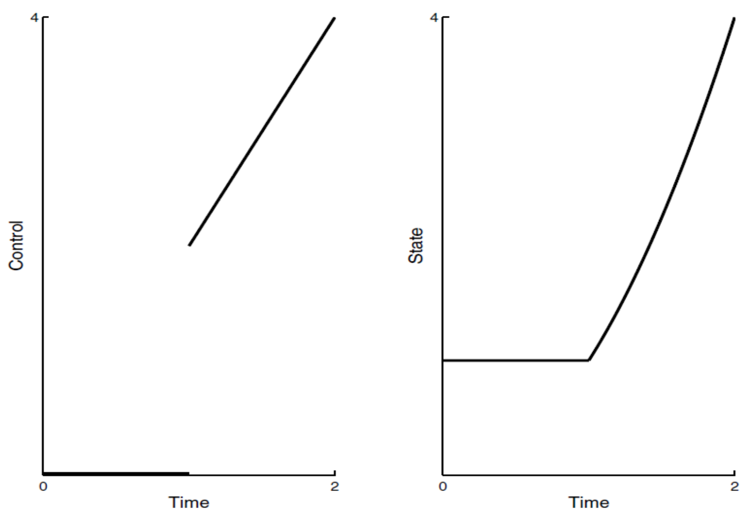


Figure 3.17 the optimal control and state for Example 3.9.3 Here, it is clear the state is not differentiable at *t* = 1.

3.10 Free Terminal Time Problems

In many applications, we are concerned with maximizing (or minimizing) an objective functional over a non-fixed time interval. If we return to our simple cancer example, Example 3.3, we could instead consider a slightly different problem. Before, we wanted to find a drug treatment over a given time frame [0, T] which would minimize the final tumor cell concentration and total harmful effects of the drug. Suppose, instead, we want to find a time frame and a control that produce an objective functional value minimum among all time frames and all controls. Namely,

$$\min_{u,T} x(t) + \int_0^T u(t)^2\, dt$$

$$\text{Subject to } x'(t) = \alpha x(t) - u(t)\ \ , x(0) = x_0$$

Notice that the minimization is now considered over the variables u and T. This is the standard way of writing an optimal control problem when T is free.

We now have more unknowns, with the optimal control and optimal terminal time both to be determined. To handle this problem, and other problems where the terminal time is free, we must redevelop the necessary conditions. As you will see, having given up information, in some sense, by allowing T to be free, we will gain new information in the way of a necessary condition we did not have before.

We note that we could just as easily allow the initial time, or both the initial and terminal times, to be free. In most applications, though, it is the final time which is allowed to move, so we handle this case.

Necessary Conditions

Let $f(t,x,u)$ *and* $g(t,x,u)$ be continuously differentiable functions in all three variables, and consider the free terminal time problem

$$\max_{u,T} \int_{t_0}^T f\big(t,x(t),u(t)\big)\, dt + \emptyset\big(T,x(T)\big)$$

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

As there are two unknowns here, we write the value of the objective functional as

$$J(u,T) = \int_0^T f\big(t, x(t), u(t)\big)\, dt + \emptyset\big(T, x(T)\big)$$

where, of course, *x* is the state corresponding to *u*. Let (u^* , T^*) be an optimal pair. Namely, u^* is a control on the nonempty, finite interval [t_0 , T^*] and $J(u\,,T) \leq J(u^*,T^*) < \infty$ for all other controls *u* and times *T*. Let x^* be the corresponding state. Let *h* be a piecewise continuous function and ϵ real number. Then,

$u^\epsilon(t) = u^\epsilon(t) + \epsilon h(t)$ is a control. As *J*(*u* , *T*) reaches a maximum at u^*, T^*, we have that

$$0 = \lim_{\epsilon \to 0} \frac{J(u^*,T^*) - J(u^\epsilon,T^*)}{\epsilon}$$

It follows from the same arguments used in section 3.4 and 3.7 that

$$0 = \frac{\partial H}{\partial u} \; at\; u^*,$$

$$\lambda = -\frac{\partial H}{\partial x} = -f_x - \lambda g_x\,,$$

$$\lambda(T^*) = \emptyset_x(T^*\,, x(T^*))\,,$$

where $\emptyset_x$ refers to the partial derivative of $\emptyset$ in the state variable or the second variable.

However, this still does not give any information about the optimal final time T^*. For this, we exploit the *T* variable of *J*. Consider real numbers $\delta \geq t_0 - T^*, so\ that\ T^* + \delta$ is an admissible terminal time. It is necessary to consider u^* and x^* on an interval larger than [t_0 ,T^*]. First, we can assume that u^* is left-continuous at T^*, by simply reassigning its value there if necessary. Then, set $u^*(t) = u^*(T^*)$ for all $t > T^*$, so that u^* will be continuous at T^*. Now, x^* is also defined for $t > T^*$. As *J*(*u*, *T*) reaches its maximum at u^* , T^*, we have

$$0 = \lim_{\delta \to 0} \frac{J(u^*,T^* + \delta) - J(u^\epsilon,T^*)}{\delta}$$

Or equivalently

$$0 = \lim_{\delta \to 0} \frac{1}{\delta} \left[\int_{t_0}^{T^*+\delta} f(t, x^*, u^*) dt + \emptyset\big(T^* + \delta, x^*(T^* + \delta)\big) - \int_{t_0}^{T^*} f(t, x^*, u^*) dt \right.$$

$$\left. + \emptyset\big(T^*, x^*(T^*)\big) \right]$$

$$= \lim_{\delta \to 0} \frac{1}{\delta} \int_{t_0}^{T^*+\delta} f(t, x^*, u^*) dt + \frac{\emptyset\big(T^* + \delta, x^*(T^* + \delta)\big) - \emptyset(T^*, x^*(T^*)}{\delta}$$

$$= f\big(T^*, x^*(T^*), u^*(T^*)\big) + \emptyset_t\big(T^*, x^*(T^*)\big) + \emptyset_x\big(T^*, x^*(T^*)\big) \frac{dx^*}{dt}(T^*)$$

$$= f\big(T^*, x^*(T^*), u^*(T^*)\big) + \lambda(T^*) g\big(T^*, x^*(T^*) u^*(T^*)\big) + \emptyset_t\big(T^*, x^*(T^*)\big)$$

$$= H\big(T^*, x^*(T^*), u^*(T^*)\ , \lambda(T^*)\big) + \emptyset_t\big(T^*, x^*(T^*)\big)\,.$$

We see the need for extending u^* and x^* in the first and second lines, as the values of t considered are greater than T^*, in the case when $\delta > 0$. The transition from the second to the third line follows via the Fundamental Theorem of Calculus and the product rule. This is due to our earlier assurance that u^* is continuous at T^*, and thus x^* is differentiable at T^*.

This gives the new necessary condition we promised. Namely,

$$H\big(T^*, x^*(T^*), u^*(T^*)\ , \lambda(T^*)\big) + \emptyset_t\big(T^*, x^*(T^*)\big) = 0.$$

In the case when $\emptyset$ is a function of $x(T)$ only, this says the Hamiltonian is 0 at the terminal time. This proof was done on a simplified problem for convenience. It should be clear, however, that the same necessary conditions would arise with bounds on the control and multiple states and controls as before, and that this new necessary condition would be unchanged. What is not clear, though, is how problems with a state fixed at both endpoints are affected. Because we did not provide a development of this case, it is not obvious how a free terminal time would alter the necessary conditions. In fact, the same new necessary condition arises. Stated formally,

if u^* is an optimal control on the finite, nonempty interval $[t_0, T^*]$, with control x^*, for the optimal control problem

$$\max_{u,T} \int_{t_0}^{T} f\big(t, x(t), u(t)\big)\, dt + \emptyset(T)$$

$$\text{Subject to } x'(t) = g\big(t, x(t), u(t)\big)\ , x(t_0) = x_0\ ,\ x(T) = x_1\ ,$$

then for some piecewise differentiable adjoint variable λ, the following necessary conditions are simultaneously satisfied by u^* , x^* , λ :

$$0 = \frac{\partial H}{\partial u}\ at\ u^*$$

$$\lambda' = -\frac{\partial H}{\partial x}$$

$$0 = H\big(T^*, x^*(T^*), u^*(T^*)\ , \lambda(T^*)\big) + \emptyset'^{(T^*)}.$$

Example (3.10.1)

$$\min_{u,T} \frac{1}{2} \int_0^T (u(t)^2 + 1)\, dt$$

$$\text{Subject to} \quad x'(t) = u(t), x(0) = 5\ , x(T) = 0.$$

$$-2 \le u(t) \le 2$$

Form the Hamiltonian

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

The adjoint equation is

$$\lambda' = -\frac{\partial H}{\partial x} = \ 0\ ,$$

So that $\lambda \equiv c$ for some constant c. further

$$\frac{\partial H}{\partial u} = u + c$$

$$0 > u^* + c \Longrightarrow u^* = 2 \Longrightarrow -2 > c\ ,$$

$$0 < u^* + c \Longrightarrow u^* = -2 \Longrightarrow 2 < c\ ,$$

$$0 = u^* + c \Longrightarrow -2 \leq u^* \leq 2 \Longrightarrow -2 \leq c \leq 2\ .$$

Clearly, as c is a constant, only one of these cases can be true. Thus, u^* is identically constant. Also, as the control must push the state from 5 down to 0, it is also clear that u^* must be negative. So, either $u^* = -2\ or\ u^* = c$.

If $u^* = -2$, then the Hamiltonian is

$$H = \frac{5}{2} - 2c.$$

As the terminal time is free, H must be zero at the final time. This allows us to solve for c giving $c = \frac{5}{4}$. However, this contradicts what we saw above, namely, when $u^* = -2$ we must have

$c > 2$. Thus, $u^* = -c$, and

$$0 = H\big(T^*, x^*(T^*), u^*(T^*)\ , \lambda(T^*)\big),$$

$$= \frac{1}{2}(u^2 + 1) - \lambda u,$$

$$= \frac{1}{2}(c^2 + 1) - c^2.$$

This yields $c = \pm 1$. As the control must be negative, we have $c = 1$ and $u^* \equiv -1$. This and

$x(0) = 5$ gives $x^*(t) = -t + 5$ so that $T^* = 5$.

3.10.1Time Optimal Control

Of particular interest is a specific type of free terminal time problems called minimal time problems, or time optimal control. The idea is simple: move a state (or states) from a given initial location to a specified final position in minimum time. It may not be immediately clear that this confirms to the form discussed above, but note that

$$T = \int_0^T 1 dt.$$

Therefore, the problem

$$\min_{u,T} \int_0^T 1\, dt$$

$$\text{Subject to} \quad x'(t) = g\big(t, x(t), u(t)\big)\ , x(t_0) = x_0\ ,\ x(T) = x_1\ ,$$

$$a \le u(t) \le b$$

is precisely what we want, namely, to find a control u which moves x from x_0 to x_1, subject to its dynamics, in minimal time. Of course, this is just as easily done with multiple states and controls. We make the note here that more complicated terminal state conditions or constraints can be used. Many times in applications, we are interested instead in moving the state or states from a specific initial condition to a certain region in minimal time. For example, we could only require $x(T)$ to be close to x_1, i.e $x(T') \le \delta$. Or, if we have two states, we could simply require they be equal, $x_1(T) = x_2(T)$, or be close, i.e, $x_1(T) - x_2(T) \le \delta$. In general, the constraints

$$k(T, x_1(T), \dots \dots, x_n(T) = 0 \ and$$

$$k(T, x_1(T), \dots \dots, x_n(T) \ge 0,$$

where k is a continuously differentiable function in all variables, can be considered. We do not treat these conditions, as such problems are generally a great deal more complicated. We refer the reader to [26, 24, 27, 16]. For examples of such problems, see [29, 28].

Example (3.10.2)

$$\min_{u,T} \int_0^T 1\, dt$$

Subject to $x'(t) = x(t)u(t) - \frac{1}{2}u(t)^2, x(0) = x_0 \in (0,1), x(T) = 1.$

We write the Hamiltonian

$$H = 1 + xu\lambda - \frac{1}{2}u^2\lambda.$$

The adjoint equation is

$$\lambda' = -\frac{\partial H}{\partial x} = -\lambda u$$

which gives

$$\lambda(t) = C \exp\left(-\int_0^t u(s)ds\right),$$

for some constant C. Note, if $C = 0$, then $\lambda \equiv 0$. This gives $H \equiv 1$, which contradicts the Hamiltonian being 0 at T^*. Thus, $C \neq 0$ so that is λ never zero. Hence, the optimality condition

$$0 = \frac{\partial H}{\partial u} = \lambda(x - u)$$

Gives

$$u^* = x^*.$$

Making this substitution in the state equation, we see x^* satisfies

$$x' = \frac{1}{2}x^2 \quad , x(0) = x_0 .$$

This gives the solutions

$$x^*(t) = \frac{2x_0}{2 - x_0 t} = u^*(t).$$

The condition $x(T) = 1$ gives

$$T^* = 2/x_0 - 2.$$

Example (3.10.3)

Let $x(t)$ represent the location of a particle at time t. Initially, it is at rest and is positioned at $x_0 > 0$. We can steer the particle by controlling its acceleration, within its designated limits. Find the acceleration which brings x to a rest at position 0 in minimum time. Specifically,

$$\min_{u,T} \int_0^T 1\, dt$$

$$\text{Subject to} \quad x''(t) = u(t)\,, x(0) = x_0 > 0, x(T) = 0.$$

$$x'(t) = 0, x'(T) = 0, -1 \le u(t) \le 1.$$

First, we recast this as a systems problem

$$\min_{u,T} \int_0^T 1\, dt$$

$$\text{Subject to} \quad {x_1}'(t) = x_2(t)\,, x_1(0) = x_0 > 0, x_1(T) = 0,$$

$${x_2}'(t) = u(t)\,, x_2(0) = 0, x_2(T) = 0,$$

$$-1 \le u(t) \le 1.$$

The Hamiltonian is

$$H = 1 + \lambda_1 x_2 + \lambda_2 u.$$

From the adjoint equations

$$\lambda'_1 = -\frac{\partial H}{\partial x_1} = 0,$$

$$\lambda'_2 = -\frac{\partial H}{\partial x_2} = -\lambda_1,$$

it is clear λ_1 is identically some constant, and λ_2 is a linear function in t. If $\lambda_2$were identically 0, then $\lambda_1 \equiv \lambda_2 \equiv 0$ from which we see $H \equiv 1$. This contradicts $H(T^*) = 0$, so λ_2 is not identically 0. Further,

$$\frac{\partial H}{\partial u} = \lambda_2.$$

As λ_2 is linear and not identically 0, it can be 0 only at a point and only once, so that u^* is bang-bang with at most one switch. Now, $x'_2 = u \; and \; x_2$ is to begin and end at 0. Therefore, it is clear that u cannot be identically $-1 \; or \; 1$, but must utilize the one allowed switch. It should also be clear that this switch occurs at the half-way point of the interval, $T^*/2$. The only thing to determine is which bound u^* begins with.

Suppose the optimal control is

$$u^*(t) = \begin{cases} 1 & when \; 0 \le t \le T^*/2 \\ -1 & when \; T^*/2 \le t \le T^*. \end{cases}$$

Using the state equation $x_2(0) = 0 = x_2(T^*)$ we can see

$$x_2{}^*(t) = \begin{cases} t & when \; 0 \le t \le T^*/2 \\ T^* - t & when \; T^*/2 \le t \le T^*. \end{cases}$$

Using $x'_1 = x_2, \; x_1(0) = x_0 \; and \; x_1(\,T^*) = 0$ it follows

$$x_1{}^*(t) = \begin{cases} \dfrac{1}{2}t^2 + x_0 & when \; 0 \le t \le T^*/2 \\ -\dfrac{1}{2}t^2 + T^*t - \dfrac{1}{2}(T^*)^2 & when \; T^*/2 \le t \le T^*. \end{cases}$$

Now, $x_1{}^*$ is continuous, so the two expressions must agree at $T^*/2$. This implies

$$\frac{1}{2}(T^*/2)^2 + x_0 = -\frac{1}{2}(T^*/2) + T^*(T^*/2) - \frac{1}{2}(T^*)^2$$

$$\Longrightarrow x_0 = -(T^*/2)^2 < 0.$$

This contradicts the original assumption of x_0. Therefore, the optimal control, and resulting optimal states, must be

$$u^*(t) = \begin{cases} 1 & when \; 0 \le t \le T^*/2, \\ -1 & when \; T^*/2 \le t \le T^*. \end{cases}$$

$$x_1{}^*(t) = \begin{cases} \frac{1}{2}t^2 + x_0 & when \; 0 \le t \le T^*/2, \\ -\frac{1}{2}t^2 + T^*t - \frac{1}{2}(T^*)^2 & when \; T^*/2 \le t \le T^*. \end{cases}$$

$$x_2{}^*(t) = \begin{cases} t \quad when \; 0 \le t \le T^*/2, \\ T^* - t \;\; when \;\; T^*/2 \le t \le T^*. \end{cases}$$

Using the fact that $x_1{}^*$ must be continuous, we find $x_0 = (T^*/2)^2$, so that

$T^* = 2\sqrt{x_0}$.

This can be substituted into the expressions above to finish the problem. The optimal states are shown in Figure 3.18

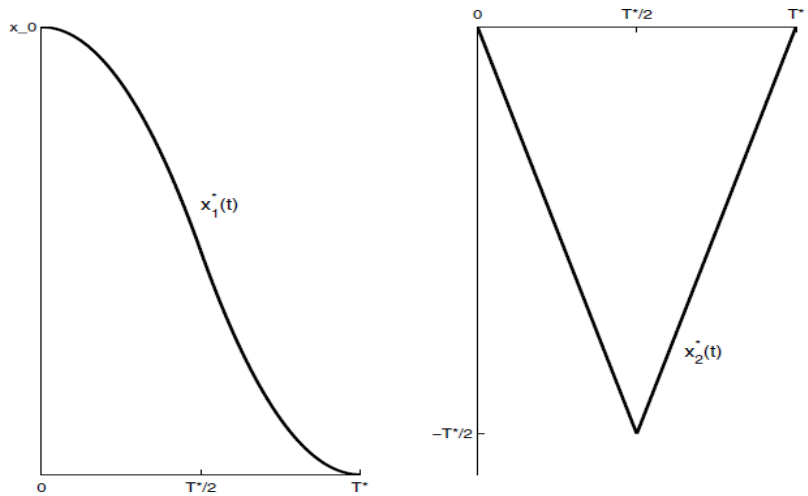


Figure 3.18 The Optimal states for example 3.10

Chapter 4

Biological Applications

4.1 Introduction:

We now begin working on the first few interactive lab programs. They will allow you to experiment with optimal control problems and see the solutions. Most of the labs are based on current applied mathematical research, dealing with an array of biological problems.

First, while MATLAB is needed to run the provided programs, it is certainly not needed to solve optimal control problems in general. Any mathematical programming language, such as FORTRAN or C++, is capable of the calculations needed. On that note, however, the programs used in this workbook are designed so that no knowledge of MATLAB is required. For each problem, there is a user-friendly interface that will guide you through. Each lab consists of two different MATLAB programs, *lab .m* and *code .m*. For example, there are two programs associated with Lab 1, *lab1.m* and *code1.m*. Here, **.m* is the extension given to all files intended for use in MATLAB. The file *code1.m* is the Runge-Kutta based, forward-backward sweep solver we will building in this chapter. It takes as input the values of the various parameters in the problem and outputs the solution to the optimality system. The file *lab1.m* is the user- friendly interface. It will ask you to enter the values of the parameters one by one, compile *code1.m* with these values, and plot the resulting solutions. All the files must be in the directory that MATLAB treats as the home directory.

This is usually the work directory. If you have experience with MATLAB, you may wish to not use the interface and instead use only the actual codes. They operate as standard MATLAB

Function files, with the parameters entered as input. This will allow you a little more freedom than the interface. However, the interface especially when going through the labs, is very convenient and will most likely save time. If you do choose to use only the *code* files, you will need to run the interface a few times before starting the labs in order to see exactly what they do, so that you can emulate them on your own.

4.2 Forward-Backward Sweep Method

Consider the optimal control problem

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

$$\text{Subject to } x'(t) = g\big(t, x(t), u(t)\big), x(t_0) = a.$$

We want to solve such problems numerically, that is, devise an algorithm that generates an approximation to an optimal piecewise continuous control u^*. We break the time interval [t_0 , t_1] into pieces with specific points of interest $t_0=b_1$,b_2 ,……………,b_N ,$b_{N+1}=t_1$. These points will usually be equally spaced. The approximation will be a vector $u^{\rightarrow} = (u_1\ , u_2, \dots \dots . \dots , u_{N+1})$ where $u_i \approx u(b_i)$. There are various methods of this type which can be employed to solve optimal control problems. For example, total-enumeration methods or linear programming techniques can be employed [30]. However, as we saw in the previous chapters, any solution to the above optimal control problem must also satisfy

$$x'(t) = g\big(t, x(t), u(t)\big), x(t_0) = a,$$

$$\lambda'(t) = -\frac{\partial H}{\partial x} = -\big(f_x(t,x,u) + \lambda(t) g_x(t,x,u)\big), \lambda(t_1) = 0,$$

$$0 = \frac{\partial H}{\partial u} = f_u(t,x,u) + \lambda(t) g_u(t,x,u) at\ u^*.$$

The third equation, the optimality condition, can usually be manipulated to find a representation of u^* in terms of t, x, and λ. If this representation is substituted back into the ODEs for x, λ, then the first two equations form a two-point boundary value problem. There exist many numerical methods to solve initial value problems, such as Runge-Kutta or adaptive schemes, and boundary value problems, such as shooting methods [27, 31]. Any of these methods could be used to solve the optimality system, and thus, the optimal control problem (if appropriate existence and uniqueness results are established).

We wish to take advantage of certain characteristics of the optimality system, however. First, we are given an initial condition for the state *x* but a final time condition for the adjoint ¸. Second, g is a function of *t*, *x*, and *u* only. Values for ¸ are not needed to solve the differential equation for *x* using a standard ODE solver. Taking this into account, the method we present here is very intuitive. It is generally referred to as the Forward-Backward Sweep method. Information about convergence and stability of this method can be found in [32]. A rough outline of the algorithm is given below. Here, $x^{\rightarrow} = (x_1, \dots \dots . x_{N+1})$ and $\lambda^{\rightarrow} = (\lambda_1, \lambda_2, \dots \dots, \lambda_{N+1})$ are the vector approximations for the state and adjoint.

Step1. Make an initial guess for $u^{\rightarrow}$ over the interval.

Step2. Using the initial condition $x_1 = x(t_0) = a$ and the values for $u^{\rightarrow}$, solve $x^{\rightarrow}$ forward in time according to its differential equation in the optimality system.

Step3. Using the transversality condition ¸$\lambda_{N+1} = \lambda(t_1) = 0$ and the values for $u^{\rightarrow}$ and $x^{\rightarrow}$, solve $\lambda^{\rightarrow}$¸ backward in time according to its differential equation in the optimality system.

Step4. Update $u^{\rightarrow}$ by entering the new $x^{\rightarrow}$ and $\lambda^{\rightarrow}$¸ values into the characterization of the optimal control.

Step5. Check convergence. If values of the variables in this iteration and the last iteration are negligibly close, output the current values as solutions. If values are not close, return to Step 2.

An example of successive control estimates is shown in Figure 4.1. We make a few notes about the algorithm. For the initial guess, $u^{\rightarrow} \equiv 0$ is almost always sufficient. In certain problems, where division by *u* occurs for example, a different initial guess must be used. Occasionally, the initial guess may require adjusting if the algorithm has problems converging. Often in Step 4, it is necessary to use a convex combination between the previous control values and values given by the current characterization. This often helps to speed the convergence. As you will see, this is done in the provided codes. For Steps 2 and 3, any standard ODE solver can be used. For the purposes, a Runge-Kutta4 routine is used. Specifically, given a step size *h* and an ODE $x'(t) = f(t, x(t))$, the approximation of $x\,(t + h)$ given $x\,(t)$ is

$$x(t+h) \approx x(t) + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{4.2.1}$$

Where

$$k_1 = f\big(t, x(t)\big)$$

$$k_2 = f(t + \frac{h}{2}\,, x(t) + \frac{h}{2}k_1)$$

$$k_3 = f(t + \frac{h}{2}\,, x(t) + \frac{h}{2}k_2) \tag{4.2.2}$$

$$k_4 = f(t + h\,, x(t) + hk_3).$$

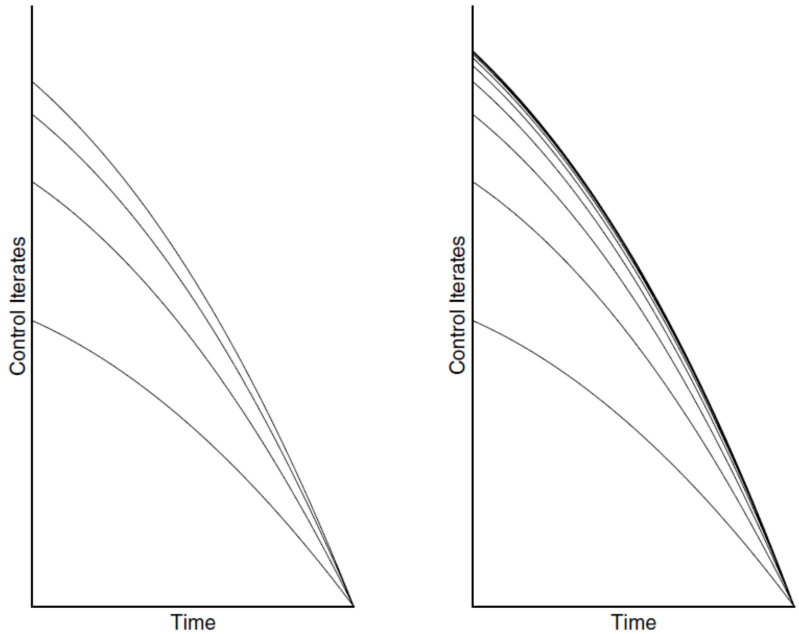


Figure 4.1 Control estimates are plotted. The first four iterations (after the initial guess) are plotted in the first graph, and the frst fifteen in the second. Note the graphs are converging to the correct control.

The error for Runge-Kutta 4 is $O(h^4)$. More information on the stability and accuracy of this and other Runge-Kutta routines is found in numerous texts. One of the classic references for these methods is Butcher [16, 17].

Many types of convergence tests exist for Step 5. Often times, it is sufficientto require $\|u - oldu\| = \sum_{i=1}^{N=1} |u_i - oldu_i|$ to be small, where $u^{\rightarrow}$ is the vector of estimated values of the control during the current iteration, and $oldu^{\rightarrow}$ is the vector of estimated values from the previous iteration. Here, $\| \,.\, \|$ refers to the l^1 norm for vectors, i.e., the sum of the absolute value of the terms. Both these vectors are of length $N + 1$, as there are N time steps. In this text, we use a slightly stricter convergence test. Namely, we will require the relative error to be negligibly small, i.e.,

$$\frac{\|u^{\rightarrow} - oldu^{\rightarrow}\|}{\|u^{\rightarrow}\|} \leq \delta$$

Where δ is the accepted tolerance. We must make one small adjustment; we must allow for zero controls. So, multiply both sides by$\|u^{\rightarrow}\|$ to remove it from the denominator. Therefore, our requirement is

$$\delta\|u^{\rightarrow}\| - \|u^{\rightarrow} - oldu^{\rightarrow}\| \geq 0,$$

$$\delta \sum_{i=1}^{N+1}|u_i| - \sum_{i=1}^{N+1}|u_i - oldu_i| \geq 0. \quad (4.2.3)$$

We will actually make this requirement of all variables, not just the control.

In the lab programs, we take $N = 1000 \; and \; \delta = 0.001$.

The remainder of this chapter will be devoted to further explanation of the Forward-Backward Sweep algorithm by way of example.

Example (4.2.1)

$$\min_{u} \int_0^1 (Ax(t) - Bu(t)^2)\, dt$$

$$\text{Subject to } x'(t) = -\tfrac{1}{2}x(t)^2 + Cu(t), x(0) = x_0 > -2,$$

$$A \geq 0, B > 0.$$

We require $B > 0$ so that this is a maximization problem. Before writing the code, we develop the optimality system of this problem by first noting the Hamiltonian is

$$H = Ax - Bu^2 - \frac{1}{2}\lambda x^2 + C\lambda u.$$

Using the optimality condition,

$$0 = \frac{\partial H}{\partial u} = -2Bu + C\lambda \Longrightarrow u^* = \frac{C\lambda}{2B}.$$

We also easily calculate the adjoint equation to find

$$x'(t) = -\frac{1}{2}x^2 + Cu, x(0) = x_0$$

$$\lambda'(t) = -A + x\lambda, \lambda(1) = 0.$$

Using these two differential equations and the representation of u^*, we generate the numerical code as described above, written in MATLAB [5]. The code can be viewed in its entirety in the file *code1.m*, and is also shown in increments below.

Code1.m

```
function y = code1(A,B,C,x0)
test = -1;
delta = 0.001;
N = 1000;
t = linspace(0,1,N+1);
h = 1/N;
h2 = h/2;
u = zeros(1,N+1);
x = zeros(1,N+1);
x(1) = x0;
lambda = zeros(1,N+1);
while(test < 0)
oldu = u;
oldx = x;
oldlambda = lambda;
for i = 1:N
```

```
k1 = -0.5*x(i)^2 + C*u(i);
k2 = -0.5*(x(i) + h2*k1)^2 + C*0.5*(u(i) + u(i+1));
k3 = -0.5*(x(i) + h2*k2)^2 + C*0.5*(u(i) + u(i+1));
k4 = -0.5*(x(i) + h*k3)^2 + C*u(i+1);
x(i+1) = x(i) + (h/6)*(k1 + 2*k2 + 2*k3 + k4);
end

for i = 1:N
j = N + 2 - i;
k1 = -A + lambda(j)*x(j);
k2 = -A + (lambda(j) - h2*k1)*0.5*(x(j)+x(j-1));
k3 = -A + (lambda(j) - h2*k2)*0.5*(x(j)+x(j-1));
k4 = -A + (lambda(j) - h*k3)*x(j-1);
lambda(j-1) = lambda(j) - ...
    (h/6)*(k1 + 2*k2 + 2*k3 + k4);
end

u1 = C*lambda/(2*B);
u = 0.5*(u1 + oldu);

temp1 = delta*sum(abs(u)) - sum(abs(oldu - u));
temp2 = delta*sum(abs(x)) - sum(abs(oldx - x));
temp3 = delta*sum(abs(lambda)) - ...
    sum(abs(oldlambda - lambda));
test = min(temp1, min(temp2, temp3));
end

```

```
y(1,:) = t;
y(2,:) = x;
y(3,:) = lambda;
y(4,:) = u;
```

Lab1

```
clear,close all
clc
y = code1(1,1,4,1);
figure
ax1 = subplot(3,1,1);
ax2 = subplot(3,1,2);
ax3 = subplot(3,1,3);
x = linspace(0,1,1001);
y1 = y(2,:);
y2 = y(3,:);
y3 = y(4,:);

plot(ax1,x,y1)
% title(ax1,'Top Subplot')
xlabel(ax1,'Time')
ylabel(ax1,'State')

plot(ax2,x,y2)
% title(ax2,'Second Subplot')
xlabel(ax2,'Time');
```

```
ylabel(ax2,'Adjoint')

plot(ax3,x,y3)
% title(ax2,'Bottom Subplot')
xlabel(ax3,'Time');
ylabel(ax3,'Control')
```

To begin the program, open MATLAB. At the prompt, type *lab1* and press enter. To become acquainted with the program, perform a few test runs. Enter values for the constants *A*, *B*, *C*, and x_0. At first, do not vary any parameters. The graphs of the resulting optimal solutions, i.e., the adjoint and the optimal control and state, will automatically appear. Run the program again, enter different values, and vary one of the parameters. Once you feel comfortable with the structure of the program.

This lab will focus on using the program to characterize the optimal control and resulting state and to ascertain how each parameter affects the solution. First, let us consider the goal of the problem. On one hand, we want to use the control *u* to maximize the integral of *x*. On the other hand, we also want to maximize the negative squared value of *u*. This, of course, is equivalent to minimizing the squared value of *u*. Thus, we must find the right balance of increasing *x* and keeping *u* as small as possible. Enter the values

$$A = 1 \quad B = 1 \quad C = 4 \quad x_0 = 1 \tag{4.2.4}$$

and do not vary any parameters, then look at the solutions. Your output should look something like Figure 4.2. We see *u* begins strongly, pushing *x* up but steadily decreasing to 0. This makes logical sense when we consider the differential equation of *x*. Undisturbed by *u*, the state *x* will decrease monotonically. So, we want to push *x* up early in the time period, so that the natural decay will be less significant. As we only care about minimizing the integral of u, and the distribution is irrelevant, the control should be highest early on. We see this is exactly what the optimal control is. Also, note that x begins to decrease at the end of the interval, as the control approaches zero.

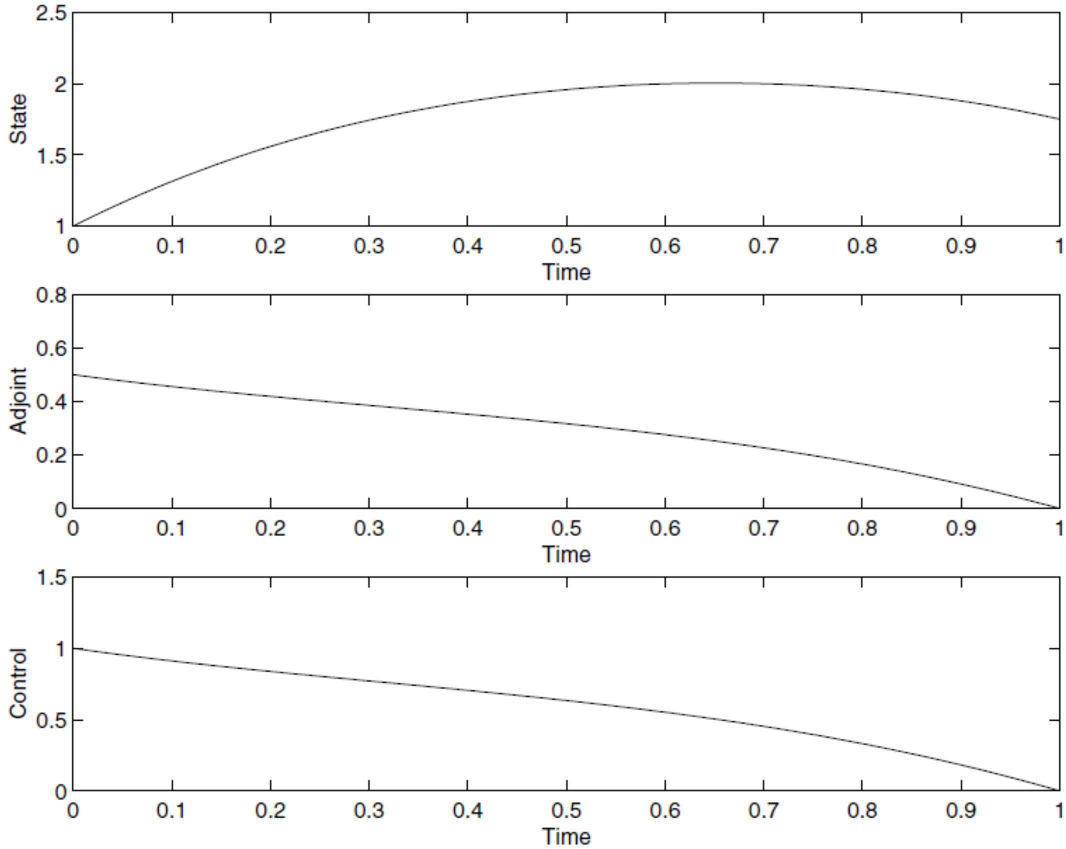


Figure 4.2.1 the optimal state, adjoint and control for the value (4.2.4)

Reenter the values in (4.2.4) and then vary the initial condition with $x_0 = 2$. As the second state begins higher, less control is needed to achieve a similar effect. Notice that the second control begins lower than the first, but they quickly approach each other and are almost identical by $t = 0.6$. This causes the two states to move towards each other as well, although they never actually meet. Now use $x_0 = -1$. This time, x begins below zero, so a greater control is needed to push the state up more quickly. Notice, however, we see the same effect as before, where the two controls eventually merge, although, much later than in the previous simulation. We mention here why the requirement $x_0 = -2$ is imposed. If you were to solve the state equation without u (i.e.,C = 0), you would find $x_0 = 2$is required, or division by 0 will occur and the state will blow-up in finite time. However, we know u will be used to increase x, so this condition is sufficient to give a finite state solution with the control.

Use the (4.4) values, varying C with C = 1*:* We have decreased the effect u has on the growth of the state. The optimal control in the second system is less than in the first. It is worth using a greater control in the first system, as it is more effective. Also, the second state, unlike the others we have seen, is decreasing over the whole interval. What little control is used does not increase the state, but only neutralizes some of the natural decay. It would now take far too much control to increase the state. Enter the same parameter values, this time varying with C = 8. The results are as you might expect. The second optimal control, now more effective, is greater than the first. The second state increases far more than the first, but still decreases as its control approaches zero. Finally, note that when C is varied, we do not have the two controls merging together. Reenter (4.4). Choose to vary A. Specifically, try A = 4 as your second value. In the second system, A = 4B, so maximizing $x(t)$ is four times as important as minimizing u^2. We see this playing out in the solutions. A greater u is used so that x can be increased appropriately. Conversely, enter (4.4) varying with B = 4. In this case, minimizing $u(t)^2$ is more important.

We see on the graph, $u(t)$ is pulled closer to zero, even though this causes $x(t)$ to increase much less at the beginning. The constants A and B are called *weight parameters*, as they determine the importance or weight of variables in the objective functional.

If you were to compare the graphs of the optimal solutions for

$$A = 1 \quad B = 2 \quad C = 4 \quad x_0 = 1 \tag{4.2.5}$$

to the solution for

$$A = 1 \quad B = 4 \quad C = 4 \quad x_0 = 1 \tag{4.2.6}$$

you would notice they were exactly the same. This is because the system is only influenced by the ratio of the constants A and B, not the actual values. We know$B \neq 0$, so we could divide it out of the integral.

4.3 Mosquitoes and Insecticide:

Mosquito- borne diseases, the best known of which is malaria, are among the leading causes of human deaths worldwide. Vector control is a very important part of the global strategy for management of mosquito-associated diseases, and insecticide application is the most important component in this effort. However, mosquito-borne diseases are now resurgent, largely because of the insecticide resistance that has developed in mosquito vectors and the drug resistance of pathogens.

Insecticides are a quick, powerful way to get rid of mosquitoes around the yard, but, unfortunately, they are only temporary. The effect usually lasts only as long as the insecticide is present, so as soon as it drifts away or dries out, the mosquitoes are back. Mosquito control officials use insecticides only when mosquitoes are especially thick and only in combination with other form of mosquito control. The same should apply to use around the house. By itself, insecticide is not a long-term solution. Two popular insecticides are:

Malathion: an organophosphate often used to treat crops against a wide array of insects. It can be sprayed directly onto vegetation, such as the bushes where mosquitoes like to rest, or used in a 5 percent solution to fog the yard. In the small amounts used for mosquito control it poses no threat to humans or wildlife. In fact, Malathion is also used to kill head lice.

Permethrin: one of a group of chemicals called pyrethroids, it is a synthetic form of a natural insecticide found in chrysanthemum flowers. It usually is mixed with oil or water and applied as a mist, about 1/100th of a pound per acre. Like malathion, permethrin kills mosquitoes by disrupting their central nervous systems. Not harmful to people and animals in small amounts, but it is toxic to fish and bees. There are three types of mosquito spraying with insecticides. Home and fog spraying, sprinkling ponds and swamps. The lesson will focus on sprinkling ponds and swamps.

Let *x* (*t*) be a population concentration at time *t*, and suppose we wish to reduce the population over a fixed time period. We will assume *x* has a growth rate *r* and carrying capacity *M*. The application of a substance is known to decrease the rate of change of *x*, by decreasing the rate in proportion to the amount of *u* and *x*. Let *u*(*t*) be the amount of this substance added at time *t*. For example, the population could be an infestation of an insect, or a harmful microbe in the body. Here we view *x*(*t*) as the concentration of a mosquitoes and *u*(*t*) an insecticide known to kill it. The differential equation representing the mold is given by

$$x'(t) = r\big(M - x(t)\big) - u(t)x(t), x(0) = x_0 \tag{4.3.1}$$

Where $x_0 > 0$ is the given initial population size. Note the term *u*(*t*)*x*(*t*) pulls down the rate of growth of the mosquitoes. The effects of both the mosquitoes and insecticide are negative for individuals around them, so we wish to minimize both. Further, while a small amount of either is acceptable, we wish to penalize for amounts too large. Hence, our problem is as follows

$$\min_u \int_0^T (Ax + u(t)^2)dt \tag{4.3.2}$$

$$\text{Subject to } x'(t) = r\big(M - x(t)\big) - u(t)x(t), x(0) = x_0.$$

The coefficient *A* is the weight parameter, balancing the relative importance of the two terms in the objective functional.

Before writing the code we develop the optimality of this problem by first noting the Hamiltonian is

$$H = Ax + u^2 + \lambda r(M - x) - \lambda xu$$

Using the optimality condition

$$0 = \frac{\partial H}{\partial u} = 2u - \lambda x \; at \; u^* \Rightarrow u^* = \frac{\lambda x}{2}$$

The adjoint equation is

$$\lambda'(t) = -\frac{\partial H}{\partial x} = -A + \lambda r + \lambda u$$

$$= -A + \lambda r + 0.5\lambda^2 x$$

$$x'(t) = Mr - x(r + u) \quad , \qquad x(0) = x_0$$

$$\lambda'(t) = -A + \lambda r + 0.5\lambda^2 x \quad , \qquad \lambda(T) = 0$$

Using these two differential equations and the representation of u^*, we generate the numerical code as described above, written in MATLAB [5]. The code can be viewed in its entirety in the file *code2.m*, and is also shown in increments below.

```
function y = code2(r,M,A,x0)

test = -1;

delta = 0.001;
N = 1000;
t = linspace(0,7,N+1);
h = 1/N;
h2 = h/2;

u = zeros (1,N+1);
```

```
x = zeros (1,N+1);
x(1) = x0;
lambda = zeros (1,N+1);
while(test < 0)
oldu = u;
oldx = x;
oldlambda = lambda;
for i = 1:N
k1 = M*r - x(i)*(r + u(i));
k2 = M*r-(x(i) + h2*k1)*(r + 0.5*(u(i) + u(i+1)));
k3 =  M*r-(x(i) + h2*k2)*(r + 0.5*(u(i) + u(i+1)));
k4 = M*r - (x(i) + h*k3)*(r + u(i+1));
x(i+1) = x(i) + (h/6)*(k1 + 2*k2 + 2*k3 + k4);
end
for i = 1:N
j = N + 2 - i;
k1 = -A - lambda(j)*r +0.5*(lambda(j))^2 * x(j);
k2=-A -(lambda(j)-h2*k1)*r + 0.5*(lambda(j)- h2*k1)^2*0.5*(x(j)+x(j-1)) ;
k3=-A-(lambda(j)-h2*k2)*r + 0.5*(lambda(j)- h2*k2)^2 * 0.5*(x(j)+x(j-1)) ;
k4 = (lambda(j)-h*k3)*r + 0.5*(lambda(j) - h*k3)^2 * x(j-1) ;
lambda(j-1) = lambda(j) - ...
                                    (h/6)*(k1 + 2*k2 + 2*k3 + k4);
```

```
end

u1 = (lambda/2)*x(i);
u = 0.5*(u1 + oldu);

temp1 = delta*sum(abs(u)) - sum(abs(oldu - u));
temp2 = delta*sum(abs(x)) - sum(abs(oldx - x));
temp3 = delta*sum(abs(lambda)) - ...
 sum(abs(oldlambda - lambda));
test = min(temp1, min(temp2, temp3));
end

y(1,:) = t;
y(2,:) = x;
y(3,:) = lambda;
y(4,:) = u;
```

4.4 Fish Harvesting

The area of internal water in Sudan (rivers - reservoirs - lakes) in the range of 2 million hectares with a total length of the Nile and its branches in the range of 6400 km and dams represent about half of this area (million hectares), next to all the above Nile sources there are many bays and canals in the common Agricultural projects p with areas of a few thousand square meters and the volume of water in the tens of billions and depths from 11-20 meters. We find that the irrigation canals in the El-Gezira project extend 5.649 km and a depth of 7.50050 meters, along with other projects' channels such as: El-Manqal - El-Rahad - El-Junaid - Khashm Al-Qirba - Al-Suki. From what mentioned it is obvious that the availability of areas and stimulated water for the establishment and success of fish farming is related agricultural projects. Fish breeding activity started with the establishment of the experimental tree farm since 1953 with the idea of the Jonglei project and its expected impact on fish resources in the region (dams), as well as the contribution to fish production by compensating areas that complain about lack of production and access to fish such as agricultural areas in El-Gezira and El- Managl and dry areas in the east and the West.

The contribution of fisheries to the GDP of Sudan is currently marginal. However, the country is endowed with water resources (by way of the Nile river system) and lands that can support vigorous capture fisheries and aquaculture. Sudan’s capture fisheries production was almost 38400 tonnes in 2017, 35100 tonnes from inland water catches and 3300 from marine catches. In 2017, there were an estimated 2330 small boats and 605 powered boats. A total of 13 686 people was reported as engaged in inland fishing in 2017, with 11% women.The aquaculture sector showed an increasing trend in the past few years, reaching 9000 tonnes in 2017. Capture fisheries activities are centered around the

River Nile and its tributaries, seasonal flood plains and four major reservoirs as well as the territorial waters of Sudan on the Red Sea. Freshwater fish culture is primarily based on the pond culture of the Nile tilapia Oreochromis niloticus and African catfish. The country is also dependant on imports of fish and fishery products (estimated at about USD 5.3 million in 2017) to satisfy the limited per capita fish consumption (about 1.1 kg in 2017). Exports are very small and were estimated at USD 1.5 million in 2017.

A large cross-section of contemporary problems in applied mathematics, related to Biology is concerned with the analysis and synthesis of dynamic processes. Fish is one of the major sources of human diet and the main source of protein and fat. Recently, consumers have become more conscious of fish as a healthier alternative meat. This is particularly due to the problems with overweight and cardiovascular diseases that have turn into one of the major problems in human health. Awareness of fish as nutritious diet has caused the demand of fish for food consumption to increase. The dynamics of the fish population may be dependent of various ecological variables, Such as the population size of natural predators, climate change and access of food. In order to determine optimal fishery policies, a suitable model for the ecosystem must be obtained. Apart from complex dynamical system of the ecosystem, defining optimal fishery is a complex problem in bioeconomics. In order to obtain a suitable definition, both ecological and economical viewpoints need to be taken into consideration. In this study optimal result is considered as the outcome which is achieved Increase the profit to the maximum extent of the harvest during a specific time period.

Suppose $x(\mathrm{t})$ represents a unit of harvested fish or population level at time t, where $x(0) = x_0 > 0$ is the initial concentration. Then, p is the price of one unit, *q* is the catchability of the fish, and *c* is cost of harvesting one unit. The control *u(t)* is the effort put into harvesting at time t. The profit at time t is

$$(pqx(t) - c)u(t) \tag{4.4.1}$$

And the total profit is

$$\int_0^T (pqx(t) - c)u(t)dt \tag{4.4.2}$$

The optimal control problem is

$$\max_u \int_0^T (pqx(t) - c)u(t)dt$$

Subject to: $x'(t) = x(t)\big(1 - x(t)\big) - qu(t)x(t)\,, x(0) = x_0 > 0$ (4.4.3)

$$0 \leq u(t) \leq M$$

The Hamiltonian is

$$H = (pqx - c)u + \lambda x - \lambda x^2 - \lambda qux \tag{4.4.4}$$

Using the necessary conditions and transversality condition

$$\lambda'(t) = -\frac{\partial H}{dx} = -pqu - \lambda + 2\lambda x + \lambda qu\,, \lambda(t) = 0 \tag{4.4.5}$$

The switching function

$$\varphi = \frac{\partial H}{du} = pqx - c - qx\lambda \tag{4.4.6}$$

Consider the singular case, i.e., suppose $\varphi = 0$ on some interval. Assuming $c > 0$, this means $x \neq 0$. So, solving for ¸ we find

$$\lambda(t) = \frac{pqx - c}{qx} = p - \frac{c}{qx} \tag{4.4.7}$$

Differentiating this expression, and using the state equation for x', it follows

$$\lambda'(t) = \frac{c}{qx^2}x' = \frac{c}{qx}(1 - x - qu) \tag{4.4.8}$$

By plugging the ¸ expression (7) into the adjoint equation (5), we get

$$\lambda'(t) = -pqu - (\,p - \frac{c}{qx})(1 - 2x - qu) \tag{4.4.9}$$

Setting the expressions (8) and (9) equal to each other and doing some simple algebra, we find the *u* terms will cancel, and we arrive at the constant expression for the state below. Noting that $x_0 = 0$ during the singular interval, by plugging x^* into state equation, we can find u^*

$$x^* = \frac{c+pq}{2pq} \tag{4.4.10}$$

$$u^* = \frac{pq-c}{pq^2} \tag{4.4.11}$$

This problem bang-bang control. The singular value for the control is a constant. If this constant lies outside the bounds on the control, i.e., is less than 0 or greater than M, then the singular control is not achievable. This forces the optimal control to be bang-bang. Suppose, however, the singular control is possible, namely, that

$$0 \le \frac{pq-c}{pq^2} \le M.$$

Using these two differential equations (4.4.3), (4.4.5) and the representation of u^*, we generate the numerical code as described above, written in MATLAB [5].

Algorithm:

Using the Runge- kutta sweep method solving $x^{\rightarrow}$ forward in time

```
for i = 1:N
k1 = x(i)*(1-x(i)-q*u(i));
k2 = (x(i) + h2*k1)*(1-(x(i) + h2*k1)-q*0.5*(u(i) + u(i+1)));
k3 = (x(i) + h2*k2)*(1-(x(i) + h2*k2)-q*0.5*(u(i) + u(i+1)));
k4 = (x(i) + h*k3)*(1-(x(i) + h2*k3)-q*u(i+1));
x(i+1) = x(i) + (h/6)*(k1 + 2*k2 + 2*k3 + k4);
```

Using the Runge- kutta sweep method solving $u^{\rightarrow}$ backward in time

```
for i = 1:N
j = N + 2 - i;
k1 = -p*q*u(j)+lambda(j)*(2*x(j)+q*u(j)-1);
k2 = -p*q*u(j)+(lambda(j)-h2*k1)*(x(j)+x(j-1)+q*u(j)-1);
```

```
k3 = -p*q*u(j)+(lambda(j)-h2*k2)*(x(j)+x(j-1)+q*u(j)-1);
k4 = -p*q*u(j)+(lambda(j)-h*k3)*(x(j-1)+q*u(j)-1);
lambda(j-1) = lambda(j) - ...
        (h/6)*(k1 + 2*k2 + 2*k3 + k4);
```

u^* based on the value of the switching function

```
for i=1:N+1
    temp = p*q*x(i) - c - q*x(i)*lambda(i)
    if(temp < 0)
            u1(i) = 0;
    elseif(temp == 0)
            u1(i) = (p*q - c)/(2*p*q^2);
else
            u1(i) = M;
    end
end
u = 0.5*(u1 + oldu);
```

Chapter 5

Results and Conclusions

5.1 Results:

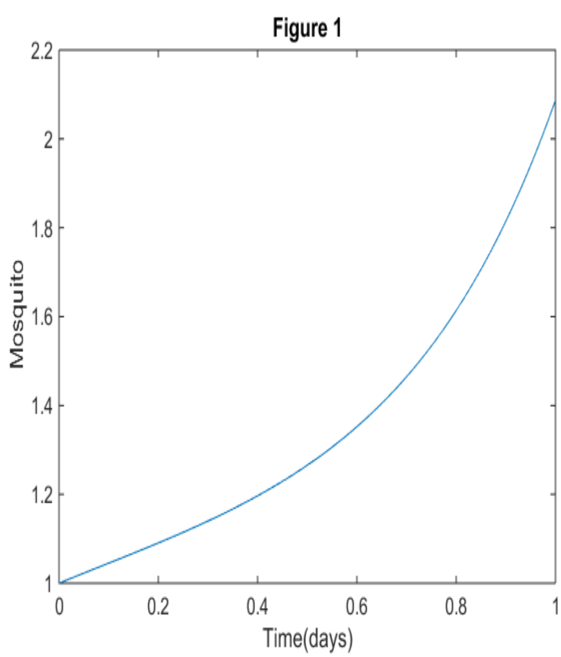
Here we consider a general mosquito and insecticide model and all the parameter values are chosen hypothetically. Enter the values

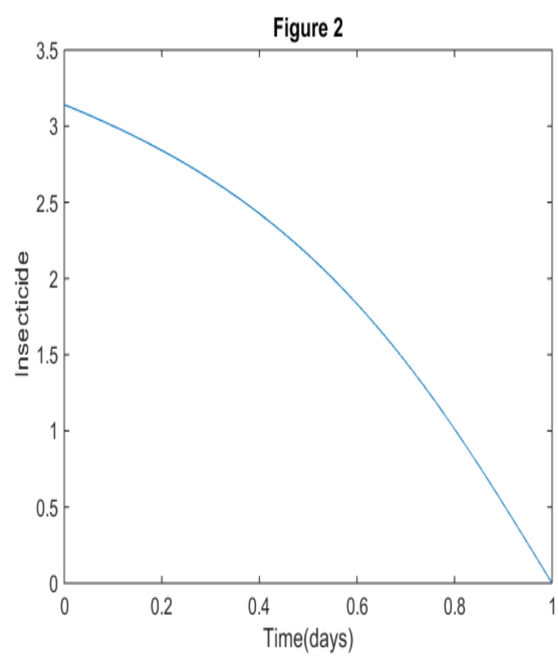
$$r = 0.4, \quad M = 10, \quad A = 6, \quad X_0 = 1.$$

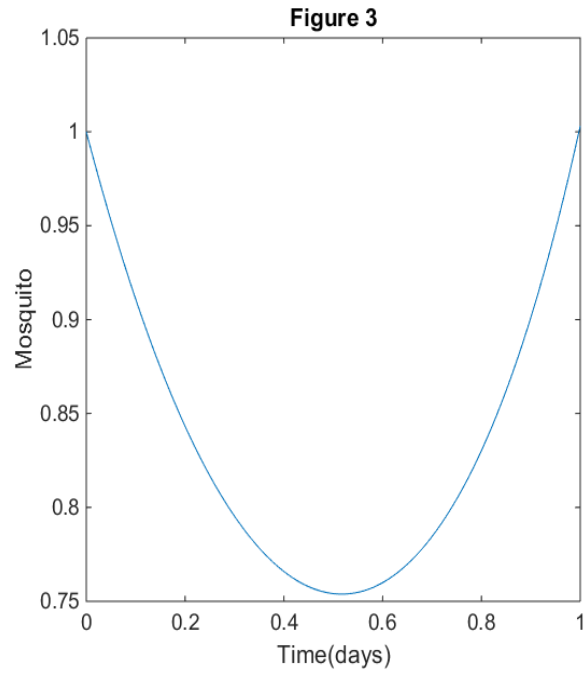
The state never decreases, with growth at the beginning and end of the interval Figure1. The control initially increases then decreasing to zero Figure2. Enter the values

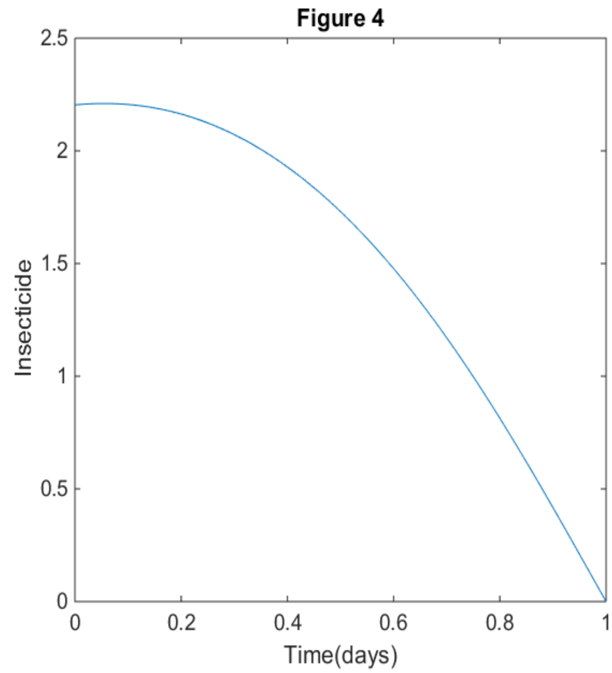
$$r = 0.3, \quad M = 5, \quad A = 10, \quad X_0 = 1.$$

The state decreases, since beginning and constant in the middle with growth at the end Figure3. The control initially increases, and then levels off to become constant. The control eventually begins decreasing again, going all the way to 0 Figure4.









Here we consider fish harvesting model and all the parameter values are chosen hypothetically. Enter the values

$$\mathrm{p} = 15 \quad \mathrm{q} = 7 \quad \mathrm{c} = 2 \quad \mathrm{M} = 1 \quad x_0 = 0.2$$

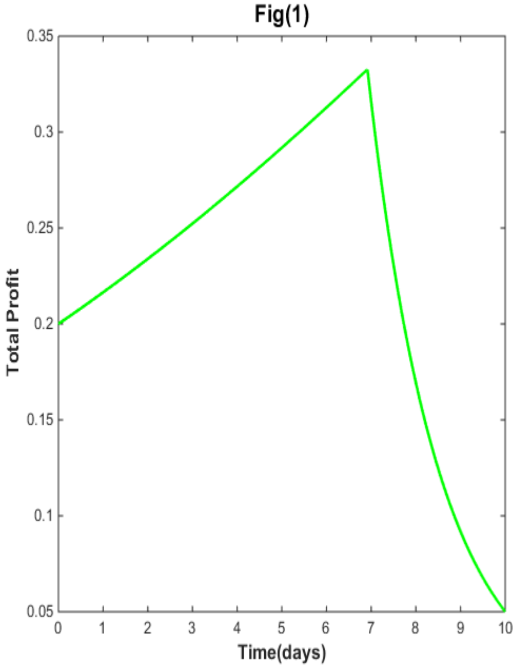
We note that the state is increasing until it reaches the maximum limit 0.36 when t=7and then begin to decline as in the figure (1) and figure (2).

Enter the value

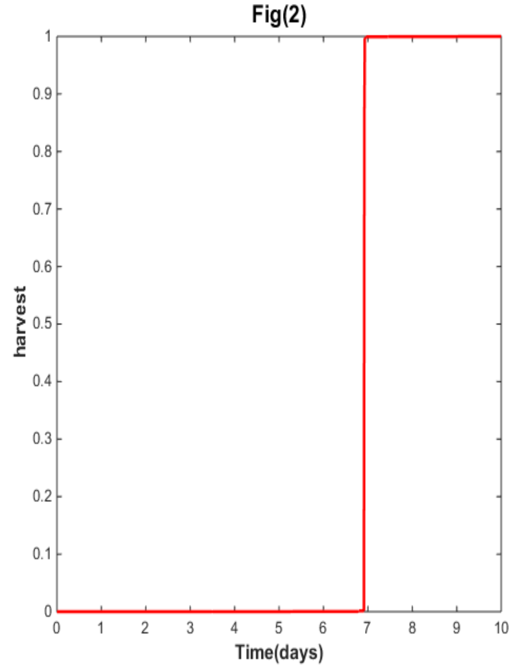
$$\mathrm{p} = 40 \quad \mathrm{q} = 15 \quad \mathrm{c} = 4 \quad \mathrm{M} = 1 \quad x_0 = 0.2$$

The optimum harvest strategy for getting the maximum profit here is harvesting in a specific time period, followed by the increase, then the decrease. Can be obtained the maximum return 0.36 when t = 8 in figure (3) and figure (4).

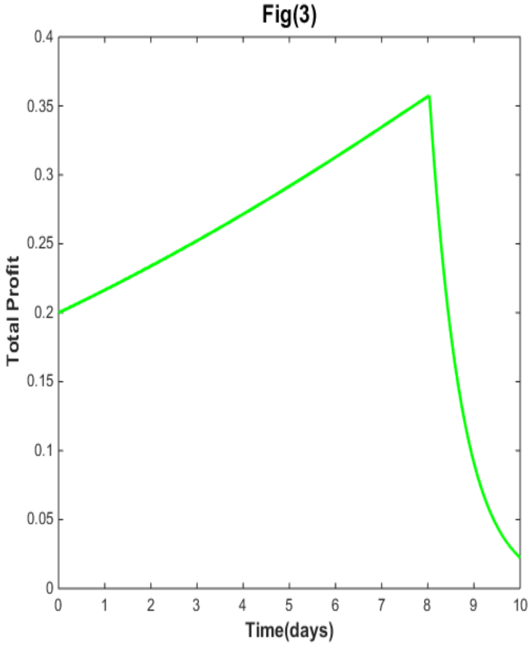
Fig(1)



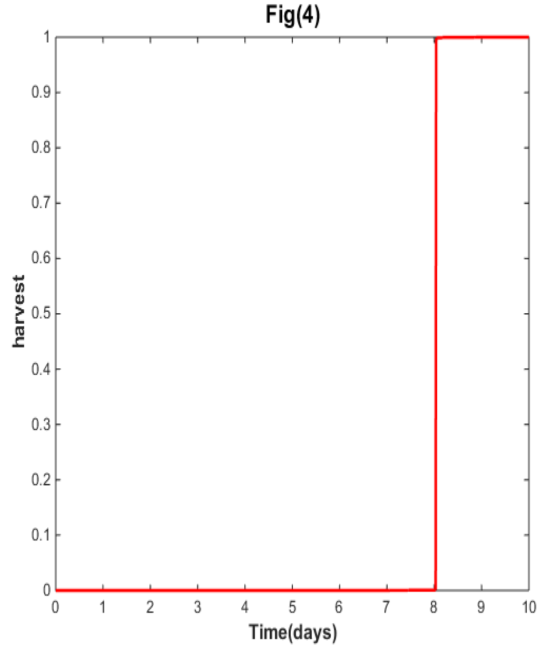
Fig(2)



Fig(3)



Fig(4)



5.2 conclusions:

In this research:

First, it is not possible to completely get rid of mosquitoes and their different phases in ponds and swamps, but by optimal control, we will reduce it to a large extent and therefore we have reduced the spread of disease malaria.

Secondly, using the maximal principle our goal is to get the best profit of fish harvesting without loss through by applying the Pontryagin's Maximum Principle of optimal control.

References

[1] Morton I. Kamien and Nancy L. Schwartz. Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management. North-Holland, New York, 1991.

[2] John. T. Betts. *Practical Methods for Optimal Control Using Nonlinear Programming*. SIAM, Philadelphia, 2001.

[3] Richard L. Burden and J. Douglas Faires. *Numerical Analysis*. PWS- KENT Publishing Company, Boston, 2004.

[4] John Butcher. On the convergence of numerical solutions of ordinary differential equations. *Math. Comp.*, 20:1{10, 1966.

[5] Matlab. *Matlab Release 13*. The Mathworks, Inc., Natich, MA, 2003.

[6] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelize, and E. F. Mishchenko. *The Mathematical Theory of Optimal Processes*. Wiley, New York, 1962.

[7] Suzanne lenhart, johnT.workman, Optimal control applied to Biological Models, 2007.

[8] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill Book Com- pany, New York, 1987.

[9] Wayne M. Getz. Optimal control and principles in population management. Proceedings of Symposia in Applied Mathematics, 30:63,82, 1984.

[10] H. J. Sussman and V. Jurdjevic. Controllability of nonlinear systems. Journal of Differential Equations, 12, 1972.

[11] H. L. Royden. *Real Analysis*. McMillan Publishing Company, Inc., New York, 1968.

[12] C. W. Clark. Mathematical Bioeconomics: The Optimal Management of Renewable Resources. Wiley, New York, 1990.

[13] L. D. Berkovitz. *Optimal Control Theory*. Springer-Verlag, New York,1974.

[14] W. H. Fleming and R. W. Rishel. Deterministic and Stochastic Optimal Control. Springer-Verlag, New York, 1975.
[15] H. Behncke. Optimal control of deterministic epidemics. Optimal Control Applications and Methods, 21:269{85, 2000
[16] John Butcher. On the convergence of numerical solutions of ordinary differential equations. Math. Comp., 20:1{10, 1966.
[17] John Butcher. A multistep generalization of Runge-Kutta methods with four or five stages. Journal of the ACM, 14:84{99, 1967.
[18] S. Butler, D. Kirschner, and S. Lenhart. Optimal control of chemotherapy affecting the infectivity of HIV. Advances in Mathematical Population Dynamics - Molecules, Cells and Man, 6:557{69, 1997.
[19] A. F. Bennett. Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press, Cambridge, 2002.
[20] H. T. Banks, H-D. Kwon, J. Toivanen, and H. T. Tran. A state dependent Riccati equation-based estimator approach for HIV feedback control. Optimal Control Applications and Methods, 27, 2006.
[21] H. Smith and P. Waltman. The theory of the chemostat. Cambridge University Press, Cambridge, 1995.
[22] Peter Dorato, Chaoki Abdallah, and Vito Cerone. Linear Quadratic Control: An Introduction. Prentice Hall, Englewood Cliffs, New Jersey, 2000.
[23] Frank L. Lewis. Optimal Control. Wiley-Interscience, New York, 1986.
[24] Arthur E. Bryson, Jr. and Yu-chi Ho. Applied Optimal Control. Ginn and Company, Waltham, Massachusetts, 1969.
[25] J. Canada, J. L. Gamez, and J. A. Montero. A study of an optimal control problem for diffusive nonlinear elliptic equations of logistic type. SIAM J. of Control Optim., 36:1171{89, 1998.
[26] R. W. Brockett. Finite Dimensional Linear Systems. Wiley, New York, 1970.

[27] Richard L. Burden and J. Douglas Faires. Numerical Analysis. PWS- KENT Publishing Company, Boston, 2004.

[28] Hal Caswell. Matrix Population Models: Construction, Analysis and Interpretation. Sinauer Press, Sunderland, MA., 2001.

[29] Vincent D. Blondel and Alexandre Megretski. Unsolved Problems in Mathematical Systems and Control Theory. Princeton University Press, Princeton, 2004.

[30] John. T. Betts. Practical Methods for Optimal Control Using Nonlinear Programming. SIAM, Philadelphia, 2001.

[31] Ward Cheney and David Kincaid. Numerical Mathematics and Computing. Thomson, Belmont, California, 2004.

[32] W. Hackbush. A numerical method for solving parabolic equations with opposite orientations. Computing, 20(3):229{40, 1978.

[33] R.W. Brockett. Finite Dimensional Linear Systems. John Wiley, 1970.

[34] D.N. Burghes and A. Graham. Introduction to Control Theory, Including Optimal Control. John Wiley, 1980.

[35] I.M. Gelfand and S.V. Fomin. Calculus of Variations. Dover, 1963.

[36] H.J. Sussmann and J.C. Willems. 300 years of optimal control: from the brachystochrone problem to the maximum principle. IEEE Control Systems, 17:32–44, 1997.

[37] J.L. Troutman. Variational Calculus and Optimal Control: Optimization with Elementary Convexity. 2nd Edition, Springer, 1996.

[38] L. Elsgolts. Differential Equations and Calculus of Variations. Mir Publishers, Moscow, Russia, 1970.

[39] L. E. Elsgolts. Calculus of Variations. Addison-Wesley, Rading, MA,1962.

[40] W.F. Feehery and P.I. Barton. Dynamic optimization with state variable path constraints. Computers in Chemical Engineering, 22:1241-1256, 1998.

[41] D. E. Kirk. Optimal Control Theory. Prentice Hall, Englewood Cliffs, NJ, 1970.
[42] D. R. Smith. Variational Methods in Optimization. Prentice Hall, Englewood Cliffs, NJ, 1974.
[43] N. I. Akhiezer. The Calculus of Variations. Blaisdell Publishing Company, Boston, MA, 1962.
[44] B. D. O. Anderson and J. B. Moore. Linear system optimization with prescribed degree of stability. Proceedings of the lEE, 116(12):2083-2087, 1969.
[45] B. D. O. Anderson and J. B. Moore. Optimal Control: Linear Quadratic Methods. Prentice Hall, Englewood Cliffs, NJ, 1990.
[46] P. J. Antsaklis and A. N. Michel. Linear Systems. The McGraw Hill Companies, Inc., New York, NY, 1997.
[47] K. Asatani. Sub-optimal control of fixed-end-point minimum energy problem via singular perturbation theory. Journal of Mathematical Analysis and Applications, 45:684-697, 1974.
[48] M. Athans and P. Falb. Optima Control: An Introduction to The Theory and Its Applications. McGraw Hill Book Company, New York, NY, 1966.
[49] T. B8.§ar and P. Bernhard. HOC-Optimal Control and Related Minimax Design Problems: Second Edition. Birkhauser, Boston, MA,1995.
[50] A. V. Balakrishnan. Control Theory and the Calculus of Variations. Academic Press, New York, NY, 1969.
[51] J. S. Bay. Fundamentals of Linear State Space Systems. The McGraw Hill Companies, Inc., New York, NY, 1999.