NEED-BASED FEEDBACK: AN OPTIMIZATION APPROACH

By

DEBRAJ CHAKRABORTY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2007

To my family.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Jacob Hammer for his kind and patient guidance during my stay at University of Florida. Our long discussions, whether technical or not were intellectually stimulating and extremely enjoyable. I thank all the faculty members of my department who helped me learn the basics of Electrical Engineering. Especially I would like to thank Dr. Pramod Khargonekar, Dr. Haniph Latchman and Dr. John Schueller for serving on my PhD committee and for improving my thesis through some insightful suggestions and corrections. Lastly I would like to thank my friends Yoav Levinbook, Saravanan Vijayakumaran, Niranjan Venkatraman and Jun Peng for comments, suggestions and especially for listening to my half-baked ideas and theories at various stages of their development.

TABLE OF CONTENTS

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

NEED BASED FEEDBACK: AN OPTIMIZATION APPROACH

By

Debraj Chakraborty

August 2007

Chair: Jacob Hammer
Major: Electrical and Computer Engineering

Feedback is often used to overcome the adverse effects of perturbations and uncertainties on the performance of engineering systems. However, failures of the feedback channel cannot be completely avoided. This dissertation addresses the questions of how and for how long can desirable performance of a perturbed system be maintained after a failure of the feedback channel.

Let Σ_ε be a system that is subject to a perturbation ε in its parameters. The exact value of the perturbation ε is not known; it is only known that ε is bounded by a given constant δ. Now, let $u(t)$ be an input function of Σ, and let $\Sigma_\varepsilon u$ be the response of the perturbed system to the signal $u(t)$. The nominal system is Σ_0, and the nominal response to the signal u is $\Sigma_0 u$. Therefore, the deviation in the response caused by the perturbation is $||\Sigma_\varepsilon u - \Sigma_0 u||$. To reduce the perturbation, add a "correction signal" $v(t)$ to the input signal, so that the perturbed response becomes $\Sigma_\varepsilon(u+v)$. Then, the new deviation between the perturbed and nominal cases becomes $||\Sigma_\varepsilon(u+v) - \Sigma_0 u||$. The correction signal $v(t)$ must be independent of perturbation value ε, as the latter is not known.

Let M be the maximal deviation allowed for the response, and let t_f be the time for which $||\Sigma_\varepsilon(u+v) - \Sigma_0 u|| \le M$. Then, the objective is to find a correction signal $v(t)$ that

maximizes t_f, given only that the perturbation ε is bounded by δ. Euler-Lagrange type first-order conditions for calculating the optimal correction signal $v(t)$ is presented. It is shown that, under rather broad conditions, the optimal correction signal $v(t)$ is either a bang-bang signal or can be arbitrarily closely approximated by a bang-bang signal.

CHAPTER 1
INTRODUCTION

In this work we reduce the need to communicate between the controller and the sensor measuring the system output by maximizing the time during which the feedback loop can remain open. This is motivated from the fact that in certain applications, feedback is not continually available or cannot be continually measured. Sometimes it is advantageous to temporarily stop the transmission of the feedback signal from the sensors at the system output to the controller. In other situations an unpredictable failure of the feedback channel can occur.

For example, in controlling space vehicles, obstacles may accidentally disrupt the line of vision between the spacecraft and earth for varying time periods. In telemetry, the need to conserve battery life may motivate a planned reduction of the time of transmission of feedback signal. Moreover stealth applications, that hide systems from detection, prefer to minimize transmission between the controlled object and the remote controller. In agriculture, measurements about soil parameters like moisture, etc., must be done manually and are consequently expensive. Usually such measurements are carried out intermittently after long intervals. In medicine, mathematical modeling and control of certain diseases have become quite common (see Panetta 2003 and references therein). However the feedback available is inherently of an intermittent nature, since measurements on the patients can only be made after long intervals. Lastly network based control systems use a common unreliable network on time-shared basis with other users and hence the feedback signal may be available to the controller only intermittently. In such applications, it is relevant to ask:

- **Research Question 1:** How long can the feedback loop be kept open, while maintaining desirable performance specifications?
- **Research Question 2:** What is the best way to control the system when we do not have the feedback signal?

Feedback is necessary due to the uncertainty inherent in any system. The main reason for this need for feedback is the lack of perfect knowledge about the system. This includes modeling inaccuracies, unmodeled nonlinearities, parametric uncertainties, spurious disturbances, input noise and measurement noise. Robustness and immunity to all this features are usually achieved through feedback strategies. Hence, it is evident that no uncertain system can be controlled *indefinitely* if the feedback signal is not available.

However, assuming some amount of tolerance for error, there will be an interval of time in which the system will perform acceptably even without any feedback. Let us assume that the system is allowed to operate within some specified level of error and the loss of feedback occurs when the error is within this specified limit. Then the system trajectory will usually require at least a finite amount of time to exceed the specified error level. This is true for most real systems and in particular, for linear time invariant systems. As we illustrated in the above applications, it is interesting to know, how and for how long this period of acceptable operation without feedback may be extended. The question we ask is: what is the maximal time until which the feedback loop can be kept open so that the error remains within the specified limit. We will show that under suitable assumptions, this period is always finite. Hence the disruption in the feedback signal, whether intended or accidental, will have to end or must be ended after this period of time. Otherwise, effective control is not possible, and the system trajectory may exceed the tolerated error.

If the disruption was accidental, this maximal time describes the upper bound of guaranteed safe operation. If the disruption carries on beyond this point the system can potentially fail. The a priori knowledge of this critical maximal time interval can help in decision-making. If the suspension of feedback was intentional, at that point of time the feedback is reconnected and the knowledge of the current outputs/states are utilized to bring down the system error. Once the error is reduced to near zero, the feedback may be disconnected again and the previous cycle may resume.

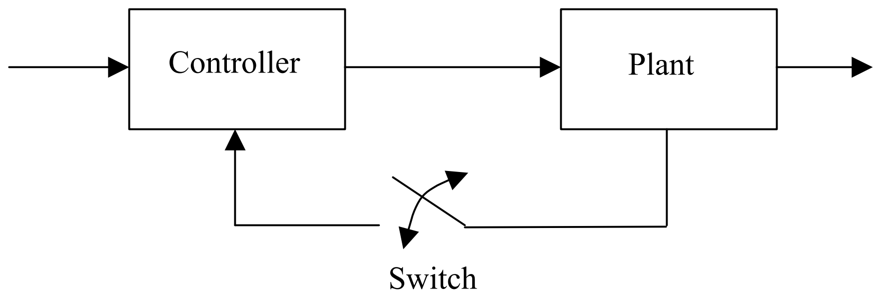

Figure 1-1: Schematic of intermittent feedback control

In this study we consider linear time invariant systems with bounded inputs and the states are assumed to be available as outputs. As in most common situations, we assume that the system parameters are uncertain but are known to lie within some bounds. The Euclidean norm of the state is taken as a measure of the system error at any time instant, which is required to be always less than a pre-specified upper bound. The norm of the initial condition is assumed to satisfy this bound. Under these assumptions the objectives for the open loop operation may be outlined briefly as follows. Find, for any permissible uncertainty, the maximal time interval for which the system error does not exceed the allowed limit. The system error is not monitored during open loop operation. Hence, such a worst-case optimization is considered to guarantee acceptable performance. A special input function is calculated to achieve maximal duration of the open loop period. Note that this is legitimate even when the loop is open, as long as no knowledge of the feedback signal is used for its computation. Existence of the time optimal controller is proved and Euler-Lagrange type conditions are derived for calculating the time optimal open loop input. It is shown that this time optimal input signal is either purely bang-bang or can be uniformly approximated by a bang-bang signal.

Finally, it should be noted that, corresponding to the optimal open loop input, an infinite number of system trajectories might be realized depending on the particular value the uncertainty takes within the allowed set The main emphasis of this formulation is to guarantee that none of the trajectories, from the infinite number possible, can exceed the allowed error bound for the maximal time interval of operation.

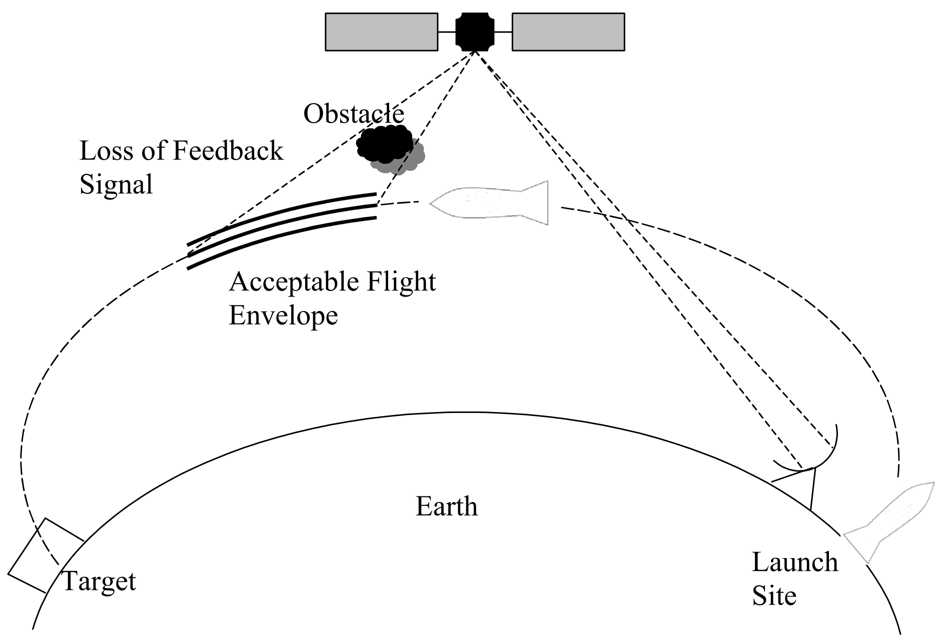

Figure 1-2: Loss of feedback due to communication channel disruption

We conclude this introductory chapter with a hypothetical example to illustrate our problem formulation: Consider the simulated firing of a ground-to-ground missile as illustrated in the figure. The flight path of the missile is predetermined and the missile is regulated carefully so that it follows the prescribed flight path. The feedback about the current coordinates, velocity, angles, etc., of the missile is transmitted to the ground controller via a satellite. The controller uses this information to determine the angle and amount of thrust needed to keep the missile on track. Under this scenario even a temporary loss of signal from the satellite may lead to the catastrophic failure of the control and a subsequent loss of the missile. We are proposing a solution to the problem by defining the best way to design the control of the missile for the duration of no feedback. Our method also guarantees that the missile can stay within the acceptable flight envelope for a maximal time after loss of signal from the satellite occurs. We prove that the best control under these circumstances is either bang-bang or can be approximated by a bang-bang signal.

CHAPTER 2
LITERATURE REVIEW

Related Work in Min-Max Optimal Control Problems

The objective of time optimality, for every possible uncertainty, in open or closed loop operation embeds our problem within the framework of min-max optimal control. This area of differential game theory has been extensively researched and encompasses a wide variety of interesting results. We review a few papers directly related to our work.

Isaacs (1954) was the first investigation to formulate a control problem in which two players with conflicting interests interact with each other in a game theoretical setup. These results were later published as a book in Isaacs (1965). The so-called "main" equation was derived, which can be viewed as a game theoretical extension of the Hamilton-Jacobi-Bellman equation and the dynamic programming approach (Bellman 1957). Independently, Kelendzheridze (1961) solved a pursuit-evasion problem, which in turn spawned a body of research of which Pontryagin et al. (1962), Pontryagin (1966, 1967a and b), Mishchenko and Pontryagin (1967) and Mischenko (1971) were early contributors. These papers generally derived Euler-Lagrange type necessary conditions for the equilibrium solutions to the pursuit-evasion problem. Berkovitz (1964 and 1967) formalized Isaacs' results in a classical calculus of variations setting and also derived necessary conditions for general systems under weak assumptions. Other early contributors were Fleming (1961 and 1964), Ho et al. (1965), Friedman (1971), Elliott and Kalton (1972) and Elliott et al. (1973).

Numerous authors developed different definitions of the equilibrium solution and various payoff functions of which Linear Quadratic Games have been extensively developed. A complete set of references can be found in Basar (1982). However the majority of interesting results have been concentrated for differential games with a saddle point solution. The problem of worst-case optimization without the saddle point assumption was posed by Feldbaum (1961 and 1965). This problem was addressed by, among others, Koivuniemi (1966), Howard and Rekasius (1964) and Bellanger (1964). Witsenhausen (1968) solved the same problem with a convex cost functional for sampled data systems. As discussed next, the problem was mathematically solved by Warga (1965a).

The concept of relaxed solutions in optimal control theory was introduced by Young (1937) and later in the book Young (1969). In a series of papers Warga (1965a and b, 1970 and 1971a and b) solved the min-max control problem using relaxed solutions with quite general assumptions. The existence of relaxed solutions (in the sense of Young 1937) to the min-max problems were guaranteed and the solutions were found to satisfy variants of Euler-Lagrange type necessary conditions. The techniques used were similar to those of Neustadt (1966 and 1967) and Gamkrelidze (1965). These results were reworked and accumulated into a book (Warga 1972), from which we use a result on conflicting controls (Theorem IX.1.2) to derive necessary conditions for our problem. We quote a simplified version of the actual theorem in Chapter 4.

It should be noted that with the exception of papers related to the Isaacs "main" equation, almost all the work reviewed above looked for open loop solutions. Apparently little has been done in search of closed loop necessary conditions characterizing solutions to the min-max problem. Conditions for existence of a stable solution were formalized with the help of concepts like stable bridges in Krasovskii and Subbotin (1974). This book contains a detailed exposition of "positional", i.e., feedback solution to differential games. In a series of papers, Ledyaev and Mishchenko (1986, 1987 and 1988) and Ledyaev (1989 and 1994) derived necessary conditions for min-max control problems of a fixed duration. This has been an area of intensive research since but most of the advances have been related to the dynamic programming approach. See Vinter (2000) and the references therein. However efforts to derive Euler-Lagrange type necessary conditions have been limited.

Throughout this study, results and ideas from standard optimal control and mathematics references have been used. Some of them are Kolmogorov and Fomin (1957 and 1961), Liusternik and Sobolev (1961), Rudin (1966), Bryson and Ho (1969), Luenberger (1969), Balakrishnan (1971), Hirsch and Smale (1974), Halmos (1982) and Zeidler (1985).

Other Approaches for Reducing Feedback Requirements

As outlined in the introduction, the main objective of this work is to reduce the duration during which the feedback loop has to be closed for controlling an uncertain plant. As far as we are aware of, such a robust time optimal formulation for reducing feedback requirements has not been dealt with in literature. However, in the context of Network Control Systems (see Nair et al. 2007), model based approaches have been developed for controlling systems with intermittent feedback. Design methodologies, in which ideal models of the system were used to guess the output of the system during the open loop, were proposed by Zhivogyladov and Middleton (2003) and Montestruque and Antsaklis (2004). Estimates of the maximum time the feedback signal may be delayed were calculated in Walsh et al. (1999). However in these works, the problem of finding the best open loop input was not dealt with and consequently, the maximum time for acceptable open loop operation has not been calculated. In conclusion, our work may have some implications for the problem of bandwidth reduction or control under communication constraints usually addressed in Network Control Systems (see Nair et al. 2007).

Residence Time Control

The problems of *pointing* and the related concept of *residence time control* are similar to our formulation of the problem. (See Meerkov and Runolfsson 1988 and the references therein).

The pointing problem was investigated among others by Skelton (1973), Eng (1979), Cannon and Schmitz (1984) and Halyo (1983). In Meerkov and Runolfsson (1988), the residence time control problem was formulated as that of choosing a feedback control law, so as to force the system states to remain, at least on the average, within pre-specified limits during some minimal period of time in spite of the disturbances that are acting on the system. The objective is very similar to that treated in this work, but the use of feedback creates a fundamental difference. Recall that in our hypothesis the feedback signal is completely absent over the period of interest. Moreover the maximum residence time was also not calculated in these papers.

CHAPTER 3
MATHEMATICAL PRELIMINARIES

In this chapter we will cover some of the well-known results from functional analysis, measure theory and linear algebra that have been used repeatedly in this dissertation. This chapter is meant as an easy reference for the results derived in the later chapters, and is an attempt to make this thesis tolerably self-contained. We start by a selection of relevant results from the theory of normed linear spaces. Most of these standard results are taken from Liusternik and Sobolev (1961), Halmos (1982) and Zeidler (1984).

Normed Linear Spaces

Weak Convergence

First we recall the definition of weak convergence in a normed linear space.

Definition 3.1: Let E be a normed linear space, $\{x_n\}$ a sequence of elements of E and $x_o \in E$. If for all functionals $f \in E^*$ (where E^* is the space conjugate to E), the sequence $f(x_n) \rightarrow f(x_o)$ as $n \rightarrow \infty$, then we say that $\{x_n\}$ converges weakly to x_o and we write $x_n \xrightarrow{w} x_o$. Here x_o is the weak limit of the sequence $\{x_n\}$.

Since Hilbert spaces are self-conjugate, the above definition may be specialized to the following form for Hilbert spaces:

Definition 3.2: A sequence $\{x_n\}$ in a Hilbert space H converges weakly to $x_o \in H$ if $\langle x_n,y \rangle \rightarrow \langle x_o,y \rangle$ as $n \rightarrow \infty$ for all $y \in H$. Then we write $x_n \xrightarrow{w} x_o$, where x_o is the weak limit of the sequence $\{x_n\}$.

A continuous functional on a compact set in a normed linear space is bounded and achieves is maximum and minimum. However, in this dissertation, most of the sets we would be interested in are not compact in the sense of strong convergence. Weak convergence and weak compactness are much less severe requirements and will shown to hold for the sets of interest in this work. Weak compactness is defined next.

Definition 3.3: A set $X \subset E$ is weakly compact (i.e., sequentially weakly compact) if every infinite sequence $\{x_n\} \in X$ has a subsequence $\{x_{n_k}\}$ such that $x_{n_k} \xrightarrow{w} x_o$ and $x_o \in X$.

The existence of solutions to optimal control problems is very much dependent on the following very important theorem.

Alaoglu's theorem

Theorem 3.4: Every bounded sequence of elements in a Hilbert space contains a weakly convergent subsequence.

This theorem guarantees the weak compactness of bounded sets in the Hilbert space. However for compactness in itself, the set of interest must be closed in the weak topology. The following result is useful for checking the sequential weak closure of convex sets:

Theorem 3.5: A bounded strongly closed convex set in a Hilbert space is also weakly closed.

In the analysis of existence of optimal solutions to dynamic optimization problems, we are actually interested in the functionals defined on compact or weakly compact sets. We define the concepts of weakly continuous and weakly upper-semicontinuous functionals.

Weak upper semicontinuity

Definition 3.6: A functional f (possibly nonlinear) on a Hilbert space H is defined to be weakly continuous at $x_0 \in H$, if for any sequence $x_n \in H$ and $x_n \xrightarrow{w} x_o$, we have $f(x_n) \to f(x_o)$.

Definition 3.7: A functional f (possibly nonlinear) defined on a Hilbert space H is defined to be weakly upper semicontinuous at $x_0 \in H$, if for any sequence $x_n \in H$ and $x_n \xrightarrow{w} x_o$, implies $\limsup_n f(x_n) \le f(x_o)$.

A generalization of the classical theorem of Weierstrass to semicontinuous functionals is often used to prove existence of solutions in optimization problems. A simplified statement is listed next:

Theorem 3.8: Let f be a weakly upper semicontinuous real valued functional on a weakly compact subset S of a Hilbert space H. The functional f is bounded above on S and achieves its maximum on S.

In differential game theory it is often not enough to study a single upper/lower semicontinuous functional but a family of such functionals. Next we state a very useful property of a family of upper semicontinuous functionals:

Theorem 3.9: (Willard 1970) Let X and A be two topological spaces. If f_α is a upper semi-continuous real valued function on X for each $\alpha \in A$, and if $\inf_\alpha f_\alpha(x)$ exists at each $x \in X$, then the function $f(x) = \inf_\alpha f_\alpha(x)$ is upper semi-continuous on X.

Definition of Borel, Lebesgue and Radon Measures

The interplay between measure theory and functional analysis is extremely common in dynamic optimization and especially differential game theory. We briefly define a few standard and well-known concepts for easy reference. For detailed exposition on these concepts see Kolmogorov and Fomin (1957) and Rudin (1966).

If S be a topological space, then the smallest σ-field containing every open set in S is called the Borel field of sets, and denoted $\Sigma(S)$, and the elements of $\Sigma(S)$ are called Borel sets. A measure defined on $\Sigma(S)$ is called a Borel measure. Let μ be a Borel measure. A Borel set $E \in \Sigma(S)$ is regular if both of the following properties hold:

- $\mu(E) = \inf\{\mu(V): E \subset V, V \text{ open}\}$
- $\mu(E) = \sup\{\mu(K): K \subset E, K \text{ compact}\}$ whenever E is open or $\mu(E) < \infty$.

If every Borel set in S is regular then μ is called regular. A finite regular Borel measure is called a Radon measure. Radon measures are important from the point of this work, because of their appearance in the Riesz representation theorem for characterizing linear functionals. Finally, the Lebesgue measure on the real line R can be defined simply as follows: Let $\Sigma(R)$ be the Borel field defined on the real line and consisting of intervals. Then the set function that assigns the interval $[a,b]$, the measure $(b-a)$ is called *the* Borel measure on $\Sigma(R)$. The completion of $\Sigma(R)$ relative to the Borel measure is called the class of Lebesgue measurable sets. The extension of the Borel measure to the completion of $\Sigma(R)$ is called the Lebesgue measure.

Riesz Representation Theorem

We present here a special form of the well-known Riesz Representation theorem, which will be used in the later chapters. This specific form is a slight simplification of that presented in Evans and Gariepy (1992).

Theorem 3.10: Let P be a compact subset of R^m. Denote the space of continuous functions mapping P to R by $C(P,R)$ and each element in this space by f. Let L be a bounded linear functional on the space $C(P,R)$. Specifically $L: C(P,R) \rightarrow R$. Then there exists a positive Radon measure μ on P and a μ-measurable function $\lambda: P \rightarrow R$ such that:

- $|\lambda(x)| = 1$ for μ-a.e. $x \in P$.
- $L(f) = \int_P \lambda\, f\, d\mu$ for all $f \in C(P,R)$.

Fubini's Theorem

Theorem 3.11: Let (X,Σ_X,μ) and (Y,Σ_Y,λ) be σ-finite measure spaces, and let f be an $(\Sigma_X \times \Sigma_Y)$-measurable function on $X \times Y$. If $0 \le f \le \infty$, and if

$$\phi(x) = \int_Y f_x \, d\lambda \quad \text{and} \quad \psi(y) = \int_X f_y \, d\mu \quad (x\in X,\, y\in Y)$$

then ϕ is Σ_X-measurable, Ψ is Σ_Y measurable, and $\int_X \phi \, d\mu = \int_{X\times Y} f \, d(\mu\times\lambda) = \int_Y \psi \, d\lambda$.

Separation of Convex Sets

Let H be a topological vector space. A set $A \subset H$ is convex if $\{\alpha a + (1-\alpha)b\} | \; 0 \le \alpha \le 1\} \subset A$ whenever $a,b\in A$. We say that a linear functional $\ell \in H^*$ (where H^* denotes the space conjugate to H), separates subset A and B of H if $\ell \neq 0$ and either $\ell(x) \le \alpha \le \ell(y)$ or $\ell(x) \ge \alpha \ge \ell(y)$ $(x\in A, y\in B)$ for some $\alpha \in R$.

Theorem 3.12: Let K and M be non-empty disjoint convex subsets of a topological vector space H and the interior of K is non-empty. Then there exists $\ell \in H^*$ that separates K and M.

Directional (Gateaux) Derivative

Definition 3.13: Let X be a vector space, Y a normed space, and T a (possibly nonlinear) transformation defined on a domain $D\subset X$ and having range $R\subset Y$. Let $x_0 \in D \subset X$ and let h be arbitrary in X. If the limit:

$$\mathcal{D}T(x;h) = \lim_{\alpha\to 0} \frac{T(x_0+\alpha h) - T(x_0)}{\alpha}$$

exists, it is called the Gateaux differential of T at x_0 with increment h.

If the above limit exists for each $h\in X$, the transformation T is said to be Gateaux differentiable at x_0. We note that this definition makes sense only when $x_0+\alpha h \in D$ for all α sufficiently small and that the limit is taken in the usual sense of norm convergence in Y. If Y is the real line and the transformation T is a real valued functional on X, then the following definition may be used for the Gateaux differential:

$$\mathcal{D}T(x;h) = \frac{d}{d\alpha} T(x_0 + \alpha h)|_{\alpha=0} \text{ for each fixed } x_0 \in X.$$

Existence and Uniqueness of Solution to a LTI System

In this section we quote a widely known theorem about the existence and uniqueness of the solution to a differential equation with measurable right hand sides. In this dissertation, we consider linear systems of the form $\dot{x}(t) = Ax(t) + Bu(t)$ where $u(t)$ is Lebesgue measurable. This theorem is used throughout to guarantee existence and uniqueness of solution for the systems under consideration. This particular version of the theorem is taken from Young (1969).

Theorem 3.14: Let $f(t,x)$ be a vector valued function with values in x-space, and suppose that in some neighborhood of (t_0,x_0), $f(t,x)$ is continuous in x for each t, measurable in t for each x, and uniformly bounded in (t,x). Then there exists an absolutely continuous function $x(t)$, defined in some neighborhood of t_0, such that $x(t_0) = x_0$ and that, almost everywhere in that neighborhood: $\dot{x}(t) = f(t,x(t))$

Suppose in addition that for some constant L, the function $f(t,x)$ satisfies, whenever (t,x_1) and (t,x_2) lie in some neighborhood N of (t_0,x_0), the Lipschitz condition $|f(t,x_1)-f(t,x_2)| \le L|x_1-x_2|$, then in some neighborhood of t_0 there exists one and only one absolutely continuous function $x(t)$ such that:

$$x(t) = x_0 + \int_{t_0}^{t} f(\tau,x(\tau))\, d\tau .$$

Jordan Canonical Representation

Theorem 3.15 (Hirsch and Smale 1974)**:** Let A be a $n\times n$ matrix with real elements. Then there exists a $n\times n$ real matrix P such that $A_d = PAP^{-1}$ is composed of diagonal blocks of the following two forms:

$$S_1 = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ldots\ldots & & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix} \text{ or } S_2 = \begin{bmatrix} F & I_2 & & & \\ & F & I_2 & & \\ & & \ldots\ldots & & \\ & & & F & I_2 \\ & & & & F \end{bmatrix}$$

The diagonal elements denoted by λ is a real eigenvalue of A and appears as many times as the multiplicity of the minimal polynomial. In the expression for S_2,

$$F = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \text{ and } I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

where $(a \pm ib)$ are the complex eigenvalues of A. Each block F is repeated as many times as the multiplicity of the eigenvalue $a + ib$ in the minimal polynomial. A_d is called the Jordan real canonical form of the matrix A.

CHAPTER 4
NOTATION AND PROBLEM FORMULATION

Notation

The systems we consider are linear time invariant continuous-time systems given by a realization of the form

(4.1) $\dot{x}(t) = A'x(t) + B'u(t) \qquad x(0) = x_0.$

Where A' and B' are constant real matrices of dimensions $n\times n$ and $n\times m$ respectively. We assume that the state x of the system is available as output. The initial condition of the system is denoted by x_0, and is given. In addition, we assume that there is an uncertainty about the values of the entries of the matrix A' and B'. To represent these uncertainties, we introduce two matrices D_A and D_B. We denote

(4.2) $A' = A + D_A$ and $B' = B + D_B$,

where A, B, D_A and D_B are given constant matrices of appropriate dimensions. While A and B are known, the matrices D_A and D_B can take any values within the sets Δ_A and Δ_B respectively. We define Δ_A and Δ_B to be bounded subsets of the spaces of real $(n\times n)$ and $(n\times m)$ matrices such that $\Delta_A = \{D_A: ||D_A|| \le d\}$ and $\Delta_B = \{D_B: ||D_B|| \le d\}$. Here the norm $||D|| = \sup_{D_{ij}} |D_{ij}|$ where D_{ij} denotes the ij^{th} element of the matrix D and d is a fixed positive real number. Note that the sets Δ_A and Δ_B are closed, bounded and hence compact. The particular values D_A and D_B takes in the sets Δ_A and Δ_B are not known a priori and hence represents the uncertainty in the system matrices A' and B'. Such an assumption is realistic because though the exact values of the system parameters are never known, usually the sets in which they belong are known a priori. For compactness of expression we group the uncertainties together and define $D = (D_A,D_B)$ belonging to the set $\Delta = \Delta_A\times\Delta_B$. It is further assumed that at least one eigenvalue of the nominal system matrix A has a non-negative real part and that the pair (A',B') is stabilizable for every value of the uncertainty set Δ. We will show in Chapter 5 that the assumption that the nominal system is unstable guarantees the existence of a finite solution to the open loop problem.

We assume that the system is performing acceptably if the Euclidean norm of the states, representing the system error, stays below a pre-specified positive bound (say M). In other words, the inequality $x^T(t)x(t) \le M$, must hold as long as the system is operating. However once the state norm exceeds the bound, the system must either be stopped or some action must be taken to prevent this from happening. Here we assume that at the point in time when the feedback signal stops, the system norm denoted as the initial condition in our problem, is less than M.

(4.3) $\quad x_0^T x_0 < M$

Then the open loop is allowed to run as long as $x^T(t)x(t) \le M$, and the loop must be closed when this inequality can be no longer made to hold.

As outlined in the introduction, we assume that there is a special input during the time when the feedback signal is absent. This input is used to keep the system error within the tolerated level for a maximal amount of time. We denote this input by $u(t)$. For $u(t)$ we will use weighted norm and inner product as described below. Here $u(t)$ is a vector valued function $(u_1(t), ..., u_m(t))^T$, where each component is real valued and Lebesgue measurable. Each $u(t)$ is assumed to lie in the Hilbert space $L_{2m}^{\rho(t)}$ with the following inner product: let $x(t)$ and $y(t)$ be two elements of $L_{2m}^{\rho(t)}$. Then the inner product $<x,y> = \int_0^\infty \rho(t)\, x(t)^T y(t)dt$, where $\rho(t) = e^{-\alpha t}$ for some fixed constant $\alpha > 0$. The norm and the metric are defined correspondingly. Let U be a set of bounded functions in $L_{2m}^{\rho(t)}$ defined as $U \equiv \{u \in L_{2m}^{\rho(t)} : \max_i |u_i(t)| \le K, t \in [0,\infty]\}$ where K is a fixed positive number. This set U of bounded measurable functions define the prospective inputs we shall consider for system (4.1). As will be shown in Chapter 4, the optimal solution is bang-bang with possibly an infinite number of switches. Hence a smaller set, e.g., the set of piecewise continuous bounded functions, would not suffice. The set U thus defined will be shown to be compact in the weak topology and thus will facilitate the proof of existence of the optimal solution. For engineering purposes, any element in U can be approximated to any arbitrary accuracy by a piecewise continuous implementable function.

Statement of the Problem

We next state the mathematical problem formulation. During the open loop period the objective is to maximize the time during which the system error is guaranteed to stay below certain bound in the face of uncertainty. This objective is achieved with the help of the special input $u(t)$ used to correct the system during the time the feedback is absent. In the introduction we described that how in various applications it is important to know both the correcting open-loop input $u(t)$, and the maximal amount of time for which the system can be left running in open loop safely. Keeping this motivation in mind we pose the following problem:

Problem Statement: Using the notation of the last section the open loop problem may be formulated as follows.

Problem 4.4: Find $\max\limits_{u \in U} t_f$ subject to the following constraints:

$$\begin{cases} \dot{x}(t) = A'x(t) + B'u(t) & 0 \le t \le t_f \\ x(0) = x_0 & x_0^T x_0 \le M \\ x^T(t)x(t) \le M & \text{for } 0 \le t \le t_f \text{ and for all } (A',B') \in (A+\Delta_A)\times(B+\Delta_A) \end{cases}$$

where $(A+\Delta_A) \equiv \{A + D_A : D_A \in \Delta_A\}$ and $(B+\Delta_B) \equiv \{B+D_B : D_B \in \Delta_B\}$.

Among the main objectives of this work is to show that such a maximum t_f exists, is finite and that it can be achieved by an input $u(t) \in U$. Assuming they exist, the interval $[0,t_f]$ and the optimal input solution $u(t)$ to the above problem has the following feature. If $u(t)$ is applied over the interval $[0,t_f]$, irrespective of the value the system matrices (A',B') takes in the sets $(A+\Delta_A)$ and $(B+\Delta_B)$, the system trajectory stays within the allowed error level, i.e., $x^T(t)x(t) \leq M$ for the entire interval $[0,t_f]$. The second objective is to find the nature of the optimal input $u(t)$ that achieves the maximum t_f. This will be achieved through the derivation of first order necessary conditions characterizing the optimal solution.

CHAPTER 5
EXISTENCE OF A SOLUTION

In this chapter we prove the existence of a solution to Problem 4.4. This will be achieved through the application of the generalized Weierstrass theorem stated in Chapter 3 (Theorem 3.8). Before going into the mathematical details we describe a brief outline of this chapter. The main idea is to repose Problem 4.4 in turns of a suitably defined upper semicontinuous functional. This functional is nothing but the minimum possible time (corresponding to the worst uncertainty) of open loop operation for any given $u \in U$. We prove that the set U defined in chapter 4 is compact in the weak topology and then we show that the functional mentioned above is weakly upper semicontinuous. Thus we are ready to use a generalized Weierstrass theorem to prove that a maximal time for the open loop operation exists, is finite, and can be achieved with an input u from the set U we defined in Chapter 4.

Weak Compactness of U

We start with a few properties of the input set U defined in Chapter 4. Most of the standard results of this section are taken from Liusternik and Sobolev (1961), Halmos (1982) and Zeidler (1984) and has been listed for reference in Chapter 3.

We prove the weak compactness property of the set U of inputs, in the form of a lemma . This property, which is a consequence of the Alaoglu's theorem (Theorem 3.4)), along with a few well-known properties of semi-continuous functions on compact sets, is essential for proving the existence of a solution to Problem 4.4.

Lemma 5.1: The set U is weakly compact in the Hilbert space $L_{2m}^{\rho(t)}$.

Proof: The set U is obviously bounded and hence by Theorem 3.4 every infinite sequence in U has a weakly convergent subsequence. However we need to show that the set U is weakly closed. Since U is convex by definition, by Theorem 3.5 we just need to show that U is strongly closed. Consider a sequence of functions $u_n(t) \to u_o(t)$ where $u_o(t) \notin U$. Hence for some set $\delta t' \subset [0,\infty)$ of nonzero measure, $u_{oi}(t) > K$, where $t \in \delta t'$ and u_{oi} denotes the i^{th} element of $u_o(t)$. Hence the following inequality must hold:

$$\int_0^\infty e^{-\alpha t} \|u_n(t) - u_o(t)\|^2 dt = \int_{\delta t'} e^{-\alpha t} \|u_n(t) - u_o(t)\|^2 dt + C_n \quad \text{where } C_n \text{ is positive real.}$$

$$\geq \int_{\delta t'} e^{-\alpha t} (K - u_{oi}(t))^2 \, dt + C_n$$

As $n \to \infty$, while C_n can tend to zero, the first term remains constant and finite. So $u_n(t)$ does not converge to $u_o(t)$. ♦

Alternative Problem

We now define an alternative formulation for Problem 4.4 using the functional defined next. It is easier to show the existence of the optimal solution in terms of this functional.

Definition 5.2: Let $J(t) = x(t)^T x(t)$ where $t \in [0,\infty)$. Define

$$T(u,D) = \begin{cases} \inf t : J(t) > M & \text{if } J(t) > M \text{ for some } t < \infty \\ \infty & \text{otherwise} \end{cases}.$$

Here $x(t)$ is related to $u(t)$ and D through Equation 4.1.

This functional may be interpreted as the time when the system error exits the allowed envelope of operation for the first time. This functional is identical to the quantity t_f defined in Problem 4.4, for trajectories obeying the constraints. In other words, for a fixed (u,D) pair, the relation $x(t)^T x(t) \leq M$ is satisfied for the interval $[0, T(u,D)]$. Moreover for a fixed $u(t)$, the relation $x(t)^T x(t) \leq M$ is satisfied for every value of the uncertainty $D \in \Delta$ within the interval $\left[0, \inf_{D \in \Delta} T(u,D)\right]$. Hence it can be easily seen that Problem 4.4 is equivalent to finding $\sup_{u \in U} \inf_{D \in \Delta} T(u,D)$, if it is finite. In addition we need to show that the supremum over U can be achieved. We summarize in the following restatement of Problem 4.4:

Problem 5.3: Show that $\sup_{u\in U} \inf_{D\in\Delta} T(u,D) < \infty$. Find $u^*\in U$ such that $\inf_{D\in\Delta} T(u^*,D) = \sup_{u\in U} \inf_{D\in\Delta} T(u,D)$.

Well posed: For this problem to be well posed, the first step is to show that $\inf_{D\in\Delta} T(u,D) < \infty$. This observation is crucial to us since we are interested in running the system on open loop as long as possible. We realize that no matter what input we create for the system, nature can choose a value for $D\in\Delta$ that will make the system trajectory leave the set of acceptable operation in finite time.

First we prove a preliminary lemma:

Lemma 5.4: Let A and D be a real $n\times n$ matrices such that A has repeated eigenvalues and $D\in\Delta$. Also let x_0 be a fixed non-zero vector in R^n. Then for any $d > 0$, there exists a D such that $\|D\| < d$ and $(A+D)$ has distinct eigenvalues. Moreover let T be a $n\times n$ real matrix such that $Q=T(A+D)T^{-1}$ results in a Q which has the following block diagonal form:

$$(5.5)\quad Q = \begin{bmatrix} Q_1 & & & & \\ & Q_2 & & & \\ & & \ddots & & \\ & & & Q_r & \\ & & & & Q_{r+1} \end{bmatrix} \text{ where } Q_i = \begin{bmatrix} a_i & b_i \\ -b_i & a_i \end{bmatrix} \text{ for } i=1,\ldots,r \text{ and } Q_{r+1} = \begin{bmatrix} a_{2r+1} & & & \\ & a_{2r+2} & & \\ & & \ldots & \\ & & & a_n \end{bmatrix}.$$

Here $(a_i \pm ib_i)$, i=1,…,r represents the complex eigenvalues and a_i (i=2r+1,….,n) represents the real eigenvalues of A. Then D can be chosen such that

$$\sum_{j=1}^{n} T_{1j}\, x_{0j} \neq 0$$

where T_{ij} is the ij^{th} element of T and x_{0j} is the j^{th} element of x_0.

Proof: Let $A_d = PAP^{-1}$ where A_d is the real Jordan canonical form of A. Recall the A_d is composed of diagonal blocks of the following two forms:

$$S_1 = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ldots & & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix} \text{ or } S_2 = \begin{bmatrix} F & I_2 & & & \\ & F & I_2 & & \\ & & \ldots & & \\ & & & F & I_2 \\ & & & & F \end{bmatrix}$$ where λ is a real eigenvalue

(possibly repeated) of A. In the expression for S_2, $F = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ where $(a \pm ib)$ are the complex roots of A (possibly repeated). (Theorem 3.15)

Let E be a diagonal matrix that we add onto A_d such that blocks S_1 and S_2 are changed as follows:

$$S_1' = \begin{bmatrix} \lambda+\varepsilon_1 & 1 & & & \\ & \lambda+\varepsilon_2 & 1 & & \\ & \ldots & & & \\ & & & \lambda+\varepsilon_{q-1} & 1 \\ & & & & \lambda+\varepsilon_q \end{bmatrix}$$

Where let q be the multiplicity of the eigenvalue λ.

$$S_2' = \begin{bmatrix} F_1 & I_2 & & & \\ & F_2 & I_2 & & \\ & \ldots & & & \\ & & & F_{s-1} & I_2 \\ & & & & F_s \end{bmatrix}$$

Where let s be the multiplicity of the eigenvalue $(a \pm ib)$ and $F_i = \begin{bmatrix} a+\gamma_i & b \\ -b & a+\gamma_i \end{bmatrix}$ i=1,2,…s.

Now the eigenvalues corresponding to block S_1' and S_2' are distinct and of the form $(\lambda + \varepsilon_i)$ and $(a + \gamma_j \pm ib)$ where i=1,2,…q and j=1,2,…s.

(Note: $\det(\lambda I - S2') = \det(\lambda I - F_1)\det(\lambda I - F_2)\ldots\ldots\det(\lambda I - F_s)$)

Thus $(A_d + E)$ has distinct eigenvalues. From the construction above it can be seen that it is possible to choose such an E such that $||E|| < r$ for any $r > 0$.

Now consider the reverse transformation:

$$P^{-1}(A_d + E)P = A + P^{-1}EP$$

$$\| P^{-1}EP\| \le k\| P^{-1}EP\|_2 \le k\,\|P^{-1}\|_2\,\|E\|_2\,\|P\|_2 = k\|E\|_2 .$$

The first inequality follows from the equivalence of norms where k is a finite constant and $\|\cdot\|_2$ is the spectral norm of the matrix. Clearly by making $\|E\|_2 < \frac{d}{k}$ we get that $\| P^{-1}EP\| < d$. Hence by defining $D = P^{-1}EP$ we achieve our objective.

Now let $\sum_{j=1}^{n} T_{1j}\, x_{0j} = 0$ with a certain choice of A with distinct eigenvalues and with a T that transforms A according to Equation 5.5. Clearly, the vector $T_1 \equiv [T_{11}\ T_{12}\ \ldots.\ T_{1n}]$ is an eigenvector of the matrix A and hence must satisfy

$\lambda_1 T_1 = AT_1$ where λ_1 is the eigenvalue corresponding to T_1.

Now from the continuity of the eigenvalues of A w.r.t. the entries of A, we can change A to $(A+D)$ arbitrarily keeping the eigenvalues distinct. Then the following equation must hold:

$\lambda_1'\, T_1' = (A+D)T_1'$.

Since D is arbitrary, T_1' can be made non-collinear to T_1 keeping $\|T_1 - T_1'\|$ arbitrarily small. It follows that T_1' and T_1 cannot be both orthogonal to x_0. ♦

Theorem 5.6: Consider the system (4.1) with a fixed initial condition x_0 in the set $\{x_0 : 0 < x_0^T x_0 \le M\}$. Then for any fixed $u(t) \in U$ and $M < \infty$ there exists a $D \in \Delta$ such that $T(u,D) < \infty$.

Proof: Assume that for some fixed $u(t) \in U$, $T(u,D) = \infty$ for all $D \in \Delta$. This implies that

$$x(t)^T x(t) \le M \quad \forall D \in \Delta \text{ and } \forall t \in [0,\infty]$$
$$\Rightarrow \| e^{A't}\, [x_0 + \int_0^t e^{-A'\tau} B' u(\tau) d\tau] \| \le M.$$

Now fix A' at some arbitrary value in $(A + \Delta_A)$ such that A' has distinct eigenvalues. By Lemma 5.4 this can be done without loss of generality. Then we do a similarity transform on A'

such that $Q = T^{-1}A'T$ is in the block diagonal form shown in Equation 5.5. Note that in this form Q and Q^T are commutative.

Now we consider the system transformed by T and name the new state vector z(t). It is related to x(t) by the following relation: z(t) = Tx(t).

We further assume that T is such that $z_{01} = \sum_{j=1}^{n} T_{1j}\, x_{0j} \neq 0$. By Lemma 5.4 this also can be done without loss of generality. We seek to prove that $\|x(t)\|$ can be made divergent with a proper choice of $D\in\Delta$. We claim that it is enough to prove $\|z(t)\|$ diverges.

$$\|z(t)\| = \|Tx(t)\| \leq \|T\|\|x(t)\| .$$

Hence $\lim_{t\to\infty} \|z(t)\| = \infty \Rightarrow \lim_{t\to\infty} \|x(t)\| = \infty$

In terms of $z(t)$ our hypothesis implies:

$$z(t)^T z(t) \leq \|T\|\, M = M' \quad \forall D'\in\Delta \text{ and } \forall t\in[0,\infty]$$

$$\Rightarrow \| e^{Qt} [z_0 + \int_0^t e^{-Q\tau} B''u(\tau)d\tau] \| \leq M'$$

where $M' = \|T\|\, M$ and $B'' = TB'$

Let $F(t) = [z_0 + \int_0^t e^{-Q\tau} B''u(\tau)d\tau]$.

Therefore,

$$[\, e^{Qt} F(t)\,]^T [\, e^{Qt} F(t)\,] \leq M'$$

$$\Rightarrow F(t)^T\ e^{(Q^T+Q)t} F(t) \leq M'$$

$$\Rightarrow F(t)^T e^{\Sigma t} F(t) \leq M'$$

Where $\Sigma = Q^T + Q$ is a diagonal matrix with at least one of the elements positive, by hypothesis. For simplicity assume that the 1st diagonal element denoted by a_1 is positive. Let us denote each entry of $F(t)$ by $f_i(t)$, i.e., $F(t)=\begin{bmatrix} f_1(t) \\ f_2(t) \\ \cdots \\ \cdots \\ f_n(t) \end{bmatrix}$. Thus for the last equation to hold at least the following must be true.

(5.7) $\lim\limits_{t\to\infty} f_1(t) = 0$

Now $f_1(t)$ may be written in a general form in the following way:

$$f_1(t) = z_{01} + \int_0^t e^{-a_1 t}[\cos(b_1\tau) \sum_{j=1}^{m} B_{1j}u_j(\tau) + \sin(b_1\tau) \sum_{j=1}^{m} B_{2j}u_j(\tau)]d\tau$$

where $(a_1 + ib_1)$ is the first eigenvalue of Q and B_{ij} is the ijth element of the matrix B''.

Then Equation 5.7 implies that

$$\lim_{t\to\infty} z_{01} + \int_0^t e^{-a_1\tau}[\cos(b_1\tau) \sum_{j=1}^{m} B_{1j}u_j(\tau) + \sin(b_1\tau) \sum_{j=1}^{m} B_{2j}u_j(\tau)]d\tau = 0$$

$$\Rightarrow z_{01} + \int_0^\infty e^{-a_1\tau} [\cos(b_1\tau) \sum_{j=1}^{m} B_{1j}u_j(\tau) + \sin(b_1\tau) \sum_{j=1}^{m} B_{2j}u_j(\tau)]d\tau = 0$$

But this is a linear equation in B_{ij} and hence cannot hold over all values of B'' in the set $T(B + \Delta_B) \equiv \{ T(B + D_B) : D_B \in \Delta_B \}$ unless $u_j(t) \equiv 0$ for every $j=1,\ldots,m$ and for all $t\in[0,\infty]$. However if this is true then $z_{01} = 0$. Hence we have a contradiction. ♦

In the above theorem we showed that, under the assumptions made in Chapter 4, the system norm $\|x(t)\|$ can diverge for any given input $u(t)$ for a proper choice of the uncertainty D. This is true for arbitrarily small uncertainty sets. Now define $T'(u) \equiv \inf\limits_{D\in\Delta} T(u,D)$, i.e., $T'(u)$ is the minimum time corresponding to $u(t)$ in which the system norm $\|x(t)\|$ can be made to escape the set $[0,M]$. According to Theorem 5.6 for any fixed $u(t) \in U$ the quantity $T'(u)$ is finite.

Corollary 5.8: If the conditions of Theorem 5.6 are satisfied then for any fixed $u(t) \in U$, the inequality $\inf_{D \in \Delta} T(u,D) < \infty$ holds.

Weak Upper Semicontinuity of $T(u,D)$

Recall that according to Problem 5.3 our objective is to show that $\sup_{u \in U} \inf_{D \in \Delta} T(u,D) < \infty$.

Define $T'(u) \equiv \inf_{D \in \Delta} T(u,D)$. Thus it would be enough to show that $T'(u)$ is weakly upper semi-continuous in $u(t)$. Along with the fact that the set U is weakly compact (Lemma 5.1)), this would effectively prove that the functional $T'(u)$ is bounded over the set U and that the supremum is achieved. (Theorem 3.8). The upper semi-continuity of the functional $T(u,D)$ in $u(t)$, for each fixed $D \in \Delta$ is demonstrated next. Denote the set of solutions for Equation 4.1 for all $u \in U$ and $D \in \Delta$ by $X(U,\Delta)$. By Theorem 3.14, each element of $X(U,\Delta)$ is an absolutely continuous function in t and unique corresponding to each (u,D) pair.

Lemma 5.9: For a fixed $D \in \Delta$, $T(u,D)$ is weakly upper semi-continuous in $u(t)$, i.e., as a sequence of functions $u_n(t) \xrightarrow{w} u_o(t)$, the functional $T(u,D)$ obeys the following relation: for any $\varepsilon > 0$, and $T(u_o,D) < \infty$ we can choose an integer N such that $T(u_n,D) - T(u_o,D) < \varepsilon$ for $n > N$.

Proof: For a fixed $D \in \Delta$ and $t < \infty$ consider the function $x(u):U \rightarrow X(U,\Delta)$, where $x(t;u)$ is the solution to Equation 4.1. As $u_n \xrightarrow{w} u_o$ (i.e., weakly converges), then for each $t \in [0,\infty)$, $x(t;u_n) \rightarrow x(t;u_o)$, i.e., $x(u_n)$ converges pointwise to $x(u_o)$. This is because for each $t \in [0,\infty)$, $x(t;u)$ is a functional linear in u.

This can be seen by defining $\rho(\tau) = \begin{cases} 1 & \text{if } \tau \le t \\ 0 & \textit{otherwise} \end{cases}$ and noting that

$$x(t) = e^{A't}\,[x_0 + \int_0^{\infty} \rho(\tau) e^{-A'\tau} B'u(\tau)d\tau].$$

Now define a functional T_p: $X(U,\Delta) \to [0,\infty]$ as follows:

$$T_p(x(t)) = \begin{cases} \inf t: x^T x > M & \text{if } x^T x > M \text{ for some } t < \infty \\ \infty & \textit{otherwise} \end{cases}.$$

Next consider a sequence of function $x_n(t) \to x_o(t)$ pointwise. We show that for any $\varepsilon > 0$ there exists an integer N such that for $n > N$ $T_p(x_n) - T_p(x_o) < \varepsilon$.

If $T_p(x_n) < T_p(x_o)$ then the claim is true. So we assume that $T_p(x_n) > T_p(x_o)$.

Let $T_p(x_n) = t_n$ and $T_p(x_o) = t_o$.

Now we assumed that $x_n \to x_o$ pointwise. By the definition of t_0, the following is true for every $\varepsilon > 0$: there is a $t_1 \in (t_0, t_0+\varepsilon)$ such that $x^T{}_0(t_1)x_0(t_1) > M$. Consequently, there is N such that $x^T{}_n(t_1)x^T{}_n(t_1) - M \ge [x^T{}_0(t_1)x_0(t_1) - M]/2 > 0$ for all $n > N$. Therefore, $t_n < t_0 + \varepsilon$ which implies that $t_n - t_o < \varepsilon$. Thus we have shown that $T_p(x_n) - T_p(x_o) < \varepsilon$.

Now, for a fixed $D \in \Delta$, consider the composition map $T(u,D) : U \xrightarrow{\Sigma} X \xrightarrow{T_p} [0,\infty]$.

From the above arguments it is clear that as $u_n \xrightarrow{w} u_o$, $x_n \to x_o$ pointwise and for any $\varepsilon > 0$ there exists N such that for $n > N$, $T_p(x_n) - T_p(x_o) < \varepsilon$. Hence $T(u_n) - T(u_o) < \varepsilon$. Thus for a fixed D, T(u,D) is weakly upper semi continuous in u(t). Hence proved. ♦

Now we need to prove that the functional $T'(u)$ is also weakly upper semi-continuous. This is true due to Lemma 5.9 and Theorem 3.9.

Corollary 5.10: The functional $T'(u) \equiv \inf_{D \in \Delta} T(u,D)$ is weakly upper semi-continuous in u(t).

Proof: The proof amounts to checking the conditions of Theorem 3.9. By Lemma 5.9), for each $D \in \Delta$, the functional $T(u,D)$ is weakly upper semi-continuous in u(t). Also for each $u \in U$, the functional $T'(u) = \inf_{D \in \Delta} T(u,D) > 0$. Hence the infimum exists always, and by the property above $T'(u)$ is weakly upper semi-continuous in $u(t)$. ♦

Existence of u^*

The weak upper semicontinuity of $T'(u)$ along with the weak compactness of the set of inputs leave us posed for applying the general Weierstrass theorem (Theorem 3.8). Hence we can now conclude the following:

Theorem 5.11: Let U and Δ be as defined above and $T'(u) \equiv \inf_{D \in \Delta} T(u,D)$ and let $\theta = \sup_{u \in U} T'(u) = \sup_{u \in U} \inf_{D \in \Delta} T(u,D)$. Then $\theta < \infty$. Moreover there exists some $u^*(t) \in U$ such that $T'(u^*) = \theta$.

Proof: This follows directly by noting that the set U is weakly compact and the functional $T'(u)$ is weakly upper semi-continuous in $u(t)$ over U. (Theorem 3.8). Hence $T'(u)$ is bounded above and it achieves it upper bound over U. ♦

In conclusion, we showed in this chapter that a solution to Problem 4.4 exists within the set of inputs U. The set U was shown to be compact in the weak topology. Then an alternative formulation of the Problem 4.4 was stated using a functional $T(u,D)$. A proof of the well-posed ness of this new statement was presented next. In the process it was shown that under the assumptions made in Chapter 4, the system trajectory can leave the envelope of acceptable operation in finite time no matter what input we apply to the system. This inbuilt finiteness of the functional $T(u,D)$, along with particular features of the formulation was utilized to prove weak upper semicontinuity of the functional $T(u,D)$ for any fixed $D \in \Delta$. Another functional, namely

$T'(u)$, denoting the worst possible time for the input u was introduced and it too was shown to be weakly upper semicontinuous in u. Lastly Weierstrass theorem was applied to prove the existence of the best input u^* and the finiteness of $\theta = \sup_{u \in U} T'(u)$. The application of u^* guarantees that the system error does not leave the allowed envelope for any value of the uncertainty for a maximal duration of time. However, it should be noted that even for this optimal input, the system error can leave the allowed envelope in finite time for certain values of the uncertainty.

CHAPTER 6
GENERALIZED FIRST ORDER CONDITIONS

In Chapter 5 we proved that the optimal solution to Problem 4.4 existed among the set of bounded and measurable input functions. Evidently it would be useful to know more about the optimal solution, with a view to facilitate it's calculation for specific problems. The conventional method for calculating solution to optimal control problems has been through characterizing first order necessary conditions. Here however the situation is complicated by the game-theoretic formulation, in that the maximization of open loop time has to be done over every possible value of the uncertainty. We draw on the considerable amount of previous research in the area of differential game theory and min-max optimal control. The relevant literature was reviewed in Chapter 2. In particular, we note that Problem 4.4 is an example of min-max or conflicting control problems studied previously among others, by Warga (1970). We use Theorem IX.1.2 (Warga 1970), for finding the first order necessary conditions characterizing the optimal solution. However we present a modified version of the theorem, which has been simplified to suit the problem we are solving in this dissertation.

Simplified Theorem from Warga (1970)

In this section we state and prove the theorem along with some required modifications. We require some new notation in this chapter:

Notation 6.1: Let S be a compact set on R^n and Σ be the Borel field of subsets of S. Then by $rpm(S)$ we denote the set of all Radon probability measures defined on Σ. Moreover we denote the Banach space of continuous functions from $S \rightarrow R$ with the supremum (L_1) norm by the notation $C(S,R)$. $\bar{C}$ denotes the closure of C and $int(C)$ denotes the interior of C.

First we require the following lemma. Let H be a topological vector space and W be a convex subset of $R \times H$. We define $W_0 \equiv \{ w_0 \in R: (w_0, w_H) \in W\}$ and $W' = \{w_H: (w_0, w_H) \in W, w_0 < 0\}$. Let H^* denote that space conjugate to H.

Lemma 6.2: Let C' be an open convex subset of H, $0 \in \text{int}(W_0)$ and $0 \in \overline{C}'$. Then either there exists $\ell_H \in H^*$ such that $\ell_H \neq 0$, $\ell_H(w_H) \geq 0 \geq \ell_H(c)$ for all $w_H \in W'$ and $c \in \overline{C}'$ or there exists $w \in W$ such that $w_0 < 0$ and $w_H \in C'$.

Proof: For every $\xi \in W$ let $\xi_0 < 0$ and $\xi_H \in C'$ be outside W. Recall that $W' = \{w_H: (w_0, w_H) \in W, w_0 < 0\}$. Then W' is a nonempty convex subset of H and $W' \cap C' = \phi$. Since C' is an open convex set, by the separation property(Theorem 3.12), there exists $\ell_H \in H^*$ and $\alpha \in R$ such that $\ell_H \neq 0$ such that

$$\ell_H (w_H) \geq \alpha \geq \ell_H (c) \quad \text{where } w_H \in W' \text{ and } c \in C'.$$

But $0 \in \overline{C}' \cap \overline{W}'$ and hence $\alpha = 0$. Thus we have

$$(6.3) \quad \ell_H(w_H) \geq 0 \geq \ell_H(c) \qquad (w_H \in W' \text{ and } c \in C')$$

This is the first alternative of the lemma.

The remaining possibility is when there exists $w \in W$ such that $w_0 < 0$ and $w_H \in C'$. Hence we have the second alternative of the lemma. ♦

The above lemma will be used in the proof of the following theorem. Next we state and prove the simplified version of Theorem IX.1.2 from Warga (1970).

Theorem 6.4: Let Q be a convex set, F be a compact set in R and P be any compact finite dimensional metric space. Consider two functions $T_1: Q \times F \to R$ and $T_2: Q \times F \times P \to R$. Now assume the following:

1. Let $\mathcal{H} = \{(q,f)\in Q\times F : T_2(q,f,p)\in[-\varepsilon,M]$ for every $p\in P\}$ where $\varepsilon,M>0$. Let there be $(q^*,f^*)\in\mathcal{H}$ such that $T_1(q^*,f^*) = \min_{(q,f)\in\mathcal{H}} T_1(q,f)$.
2. The functions T_1 and T_2 have convex Gateaux derivatives.
3. For each pair $(q,f)\in Q\times F$ the function $T_2(q,f,p):P\to R$ bounded and continuous in p.

Then there exists $\omega\in rpm(P)$ and an ω-integrable $\tilde{\omega} : P \to R$ such that

i. $|\tilde{\omega}(p)|=1$ for all $p\in P$ and $\omega(P) > 0$

ii. $\int\tilde{\omega}(p)\cdot\mathcal{D}T_2((q^*,f^*),p;(q,f)-(q^*,f^*))\omega(dp) \ge 0$

iii. $\tilde{\omega}(p)\cdot T_2(q^*,f^*,p) = \max_{a\in[-\varepsilon,M]} \tilde{\omega}(p)\cdot a$ for ω-a.a. $p\in P$

where $\mathcal{D}T_j\big((q^*,f^*);(q,f) - (q^*,f^*)\big)$ denotes the directional derivative of T_j at (q^*,f^*).

Proof: Define the set

$$\mathcal{W}(q,f) = \{\ \mathcal{D}T_1\big((q^*,f^*);(q,f) - (q^*,f^*)\big),\ \mathcal{D}T_2\big((q^*,f^*),p;(q,f) - (q^*,f^*)\big) :(q,f)\in Q\times F\ \}$$

This set is convex by hypothesis. Let $C_2 = \{\ c(p)\in \mathcal{C}(P,R) : -\varepsilon \le c(p) \le M\ \}$ where $\varepsilon > 0$. Note that C_2 is a closed convex subset of $\mathcal{C}(P,R)$ and let $C' = \mathrm{int}(C_2) - T_2(q^*,f^*)$.

Then we can apply Lemma 6.2 and the first alternative of the lemma yields the following: there exists $\ell \in \mathcal{C}^*(P,R)$ such that $\ell \ne 0$, and for all $(q,f)\in Q\times F$

(6.5) $\quad \ell\big(\mathcal{D}T_2\big((q^*,f^*),p;(q,f) - (q^*,f^*)\big)\big) \ge 0$

(6.6) $\quad \ell(c) \le 0 \ \ (c\in\overline{C}')$

In the above expressions note that $\ell\in C^*(P,R)$ and hence by the Riesz representation theorem (Theorem 3.10) we can specify an integral form of ℓ. In particular, there exists a positive Radon measure ω on P and an ω-integrable $\tilde{\omega} : P \to R$ such that

$|\tilde{\omega}(p)| = 1$ for all $p\in P$ and $\omega(P) > 0$

$$\ell(c) \equiv \int_P \tilde{\omega}(p)c(p)\omega(dp) \qquad c \in \mathcal{C}(P,R)$$

Next we investigate the inequality (6.6). In view of the definition of C' this implies

$$\ell(\, c - T_2(q^*,f^*)\,) \le 0 \quad (c \in \overline{C}')$$
$$\Rightarrow \int_P \tilde{\omega}(p)\,[c^*(p) - c(p)]\omega(dp) \ge 0 \quad (c \in C_2)$$

We claim that this equation implies claim (iii) of the theorem. To see this assume that there exists some $c_o \in C_2$ such that $\tilde{\omega}(p)c^*(p) < \tilde{\omega}(p)c_o(p)$ for $p \in \delta P \subset P$ where $\omega(\delta P) > 0$. Then we can form $c' \in C_2$ such that

$$c' \equiv \begin{cases} c_0 & \text{when } p \in \delta P \\ c^* & \text{otherwise} \end{cases} .$$

But, the integral inequality is not satisfied with $c'(p)$ since:

$$\int_P \tilde{\omega}(p)\cdot[c^*(p) - c'(p)]\omega(dp) = \int_{\delta P} \tilde{\omega}(p)\cdot[c^*(p) - c'(p)]\omega(dp) < 0.$$

Now note that $\mathcal{C}(P,R)$ is separable and hence C_2 contains a dense denumerable subset $\{c_1, c_2, \ldots.\}$. Moreover the set $\{c(p) | c \in C_2\}$ is dense in $[-\varepsilon, M]$ for all $p \in P$ and hence it follows that the set $\{c_1(p), c_2(p), \ldots\}$ is dense in $[-\varepsilon, M]$ for all $p \in P$. So we can write the following:

$$\begin{aligned} \tilde{\omega}(p)c^*(p) &\ge \sup_{j \in N} \tilde{\omega}(p)c_j(p) \\ &= \sup_{a \in [-\varepsilon, M]} \tilde{\omega}(p)a \qquad (p \in P) \\ &= \max_{a \in [-\varepsilon, M]} \tilde{\omega}(p)a \qquad (p \in P) \ \ (\text{Since } c^*(p) \in [-\varepsilon, M]) \end{aligned}$$

Thus we have proved that the statement of the theorem is a result of the first alternative of the lemma. We now prove that the other alternative of the lemma implies a contradiction with the assumption in that (q^*,f^*) is found to be not minimizing. To prove this consider the set $\mathcal{W}(q,f)$

and C' as defined above. From the lemma , there exists $w \in \mathcal{W}(q,f)$ such that $w_1 < 0$ and $w_2 \in C'$.

For each choice of $q \in Q$ and $f \in F$, define the function $h_1:[0,1] \to R$ and $h_2: [0,1] \to \mathcal{C}(P,R)$ as follows:

$$(6.7) \quad \begin{aligned} h_1(\theta) &= T_1(q^* + \theta(q-q^*), f^* + \theta(f-f^*)) \\ h_2(\theta) &= T_2(q^* + \theta(q-q^*), f^* + \theta(f-f^*),p) \quad \text{where } \theta \in [0,1]. \end{aligned}$$

It can be easily verified that the derivatives h_1' and h_2' at (0,0) can be expressed in terms of the directional derivatives of T_1 and T_2 at (q^*,f^*).

$$(6.8) \quad h_i'(0)\delta\theta = \mathcal{D}T_i\big((q^*,f^*);(q,f) - (q^*,f^*)\big)\delta\theta \quad (i=1,2)$$

Now for sufficiently small $\delta\theta > 0$,

$$\begin{aligned} h_1(\delta\theta) &= h_1(0) + h_1'(0)\delta\theta + o(\|\delta\theta\|) \\ &= h_1(0) + w_1\,\delta\theta + o(\|\delta\theta\|) < h_1(0) = T_1(q^*,f^*) \end{aligned}$$

Also we know that

$$\begin{aligned} &\lim_{\|\delta\theta\| \to 0} \|\delta\theta\|^{-1}[h_2(\delta\theta) - h_2(0) + h_2'(0)\delta\theta\,] = 0 \\ \Rightarrow &\lim_{\|\delta\theta\| \to 0} \|\delta\theta\|^{-1}\,[h_2(\delta\theta) - h_2(0)] = w_2 \in C' \end{aligned}$$

Hence for sufficiently small $\|\delta\theta\|$ we can conclude that

$$\begin{aligned} &[h_2(\delta\theta) - h_2(0)] \in \|\delta\theta\|\, C' \subset C' \quad (\text{Since } 0 \in C') \\ &h_2(\delta\theta) - T_2(q^*,f^*) \in \operatorname{int}(C_2) - T_2(q^*,f^*). \end{aligned}$$

Hence (q^*,f^*) is not a minimizing solution as hypothesized in assumption (1) of the theorem. ♦

The Scaled System

The theorem proved above is the main tool that we will use to derive first order conditions for the solution of Problem 4.4. The functions T_1 and T_2 will be appropriately defined for our problem and it will be shown that all the assumptions of Theorem 6.4 are satisfied. Hence a solution to Problem 4.4 must satisfy the conclusions of Theorem 6.4. It will be shown that some useful features of the optimal solution of Problem 4.4 may be derived from careful consideration of conclusion (ii) of Theorem 6.4. With these objectives in mind we redefine Problem 4.4 for a scaled system described next. We express system (4.1) in terms of scaled variables $y(s)$ and $w(s)$ as follows:

$$y(s) \equiv x(\beta s) \quad s\in[0,1]$$
$$v(s) \equiv u(\beta s) \quad s\in[0,1] \quad \text{where } \beta \le \theta \text{ and } \theta = \sup_{u\in U} \inf_{D\in\Delta} T(u,D).$$

In terms of these variables Problem 4.4 can be written as:

Problem 6.9: Find $\min(-\beta)$ subject to the following constraints:

$$\begin{cases} \dot{y}(s) = \beta(A'y(s) + B'v(s)) & 0 \le s \le 1 \\ y(0) = x_0 & x_0^T x_0 < M \\ y^T(s)y(s) \le M & \text{for } 0 \le s \le 1 \text{ and for all } (A',B') \in (A+\Delta_A)\times(B+\Delta_B) \end{cases}$$

where $(A+\Delta_A) \equiv \{A+D_A : D_A\in\Delta_A\}$ and $(B+\Delta_B) \equiv \{B+D_B : D_B\in\Delta_B\}$.

It can be easily checked that Problem 4.4 and 6.9 are identical. Using standard methods, $y(s)$ can be expressed as:

$$y(s) = e^{\beta A's}\left[x_0 + \int_0^s e^{-\beta A'\tau}\,\beta B'v(\tau)d\tau\right]$$

Definition 6.10: Now we identify the relevant sets in Problem 6.9 with those defined in Theorem 6.4 as follows:

$$Q \equiv U_{[0,1]} \quad \text{where } U_{[0,1]} \equiv \{v(t)\in U : v(t) = 0 \text{ for } t > 1\}$$
$$F \equiv [0,\theta+1]$$
$$P \equiv \{(A',B',s)\in (A+\Delta_A)\times(B+\Delta_B)\times[0,1]\}$$

And we define the functions T_1 and T_2 as follows:

$$T_1(v(s),\beta) = -\beta$$
$$T_2(v(s),\beta,(A',B',s)) = y^T y$$

In the following Theorem $P \equiv (A+\Delta_A)\times(B+\Delta_B)\times[0,1]$. Hence if $\omega \in rpm(P)$ then conditional probabilities $\omega(A',B'|s)$ as well as the marginal probability $\omega(s)$ can be defined in the usual way so that for any $(A',B',s)\in\Sigma$, the joint probability $\omega(A',B',s) = \omega(A',B'|s)\,\omega(s)$. Let us denote the Lebesgue measure on $[0,1]$ by μ. Also note that in the following theorem and thereafter the variables τ and s are both used to denote the scaled time and lies in [0,1]. They are used interchangeably as appropriate.

Necessary Conditions

We are now ready to apply Theorem 6.4 to Problem 6.9 and thus derive necessary conditions for the optimal solution (v^*,β^*). From this theorem we expect to gain some insight about the characteristics of the optimal input v^*. Indeed, from the theorem stated below it will become apparent that the optimal input v^* may have a bang-bang control feature. The result introduces the quantity $z(s)$ over the scaled interval time [0,1], which is similar to the classical switching function for the bang-bang control input v^*.(Pontryagin et al. 1962). However the exact characteristics of $z(s)$ will have to be clarified through some further investigations. The solution to Problem 6.9 and hence the solution to Problem 4.4 is summarized in the following theorem:

Theorem 6.11: Let $(v^*(s),\beta^*)$ be the solution to Problem 6.9. Then there exists $\omega\in rpm((A+\Delta_A)\times(B+\Delta_B)\times[0,1])$ and a μ-measurable function $z(s):[0,1]\to R^m$ such that the following are satisfied:

i. $$z(s) = \int_s^1 \int_P y^T(\tau) e^{\beta^* A'(\tau-s)} B' \omega(dA'\times dB'|\tau)\,\omega(d\tau)$$

ii. $z(s)\,v(s) \ge z(s)\,v^*(s)$ for μ-a.a. $s\in[0,1]$ and for any $v\in U_{[0,1]}$

Where the support of ω is given by the set: $\Omega = \{ (A',B',s)\in(A+\Delta_A)\times(B+\Delta_B)\times[0,1]: y^T y = M \}$

Proof: We can apply Theorem 6.4 since all the assumptions are clearly met We first calculate the directional derivatives of T_1 and T_2. It is more convenient to use the variable τ as the time variable and s as the running variable in the following expressions. The advantage will be apparent from the subsequent derivation.

$$\mathcal{D}T_1\big((v^*,\beta^*);(v,\beta)-(v^*,\beta^*)\big) = (\beta-\beta^*)$$
$$\mathcal{D}T_2\big((v^*,\beta^*);(v-v^*)\big)_{\beta=\beta^*} = y^T(\tau)\left\{\int_0^\tau e^{\beta^* A'(\tau-s)}\,\beta^* B'(v(s)-v^*(s))ds\right\}$$

By Theorem 6.4 there exists $\omega\in rpm(P)$ and an ω-integrable integrable $\tilde{\omega}: P\to R$ such that

$$|\tilde{\omega}(p)| = 1 \text{ for } \omega\text{-a.a } p\in P \text{ and } \omega(P) > 0$$
$$\int \tilde{\omega}(p)\cdot\mathcal{D}T_2((v^*,\beta^*);(v,\beta)-(v^*,\beta^*))\omega(dp) \ge 0$$

If we set $\beta = \beta^*$ in the above inequality and refer to the expression for the directional derivatives of T_1 and T_2 we obtain

$$(6.12)\quad \int \tilde{\omega}(p)\cdot\mathcal{D}T_2((v^*,\beta^*);(v-v^*))\omega(dp) \ge 0.$$

$$\Rightarrow \int_P \tilde{\omega}(p)y^T(\tau)\left\{\int_0^\tau e^{\beta^* A'(\tau-s)}\,\beta^* B'(v(s)-v^*(s))\mu(ds)\right\}\omega(dp) \ge 0$$

$$\Rightarrow \int_P \tilde{\omega}(p)y^T(\tau)\left\{\int_0^1 e^{\beta^* A'(\tau-s)}\,\beta^* B'\Psi_{[0,\tau]}(s)\big(v(s)-v^*(s)\big)\mu(ds)\right\}\omega(dp) \ge 0$$

(Where $\Psi_{[0,\tau]}(s) = \begin{cases} 1 & \text{when } 0\le s\le\tau \\ 0 & \text{otherwise}\end{cases}$).

$$\Rightarrow \int_0^1 \left\{\int_P \tilde{\omega}(p)y^T(\tau)e^{\beta^* A'(\tau-s)}\,B'\Psi_{[0,\tau]}(s)\omega(dp)\right\}(v(s)-v^*(s))\mu(ds) \ge 0$$ (By Fubini's Theorem (3.11))

$$\Rightarrow \int_0^1 z(s)(v(s)-v^*(s))\mu(ds) \ge 0 \quad \text{where } z(s) \equiv \int_P \tilde{\omega}(p)y^T(\tau)e^{\beta^* A'(\tau-s)}\,B'\Psi_{[0,\tau]}(s)\omega(dp)$$

Now by hypothesis, $P \equiv \{(A+\Delta_A)\times(B+\Delta_B)\times[0,1]\}$ and ω is a Radon probability measure on P. Hence we can write $\omega(dA',dB',d\tau)$ in terms of the appropriately defined conditional and marginal probabilities.

The expression for $z(s)$ can be rewritten in terms of the conditionals as follows:

$$z(s) \equiv \int_0^1 \int_\Delta \tilde{\omega}(A',B',\tau)y^T(\tau)e^{\beta^* A'(\tau-s)}\, B'\omega(dA',dB'/\tau)\, \Psi_{[0,\tau]}(s)\, \omega(d\tau)$$

$$= \int_s^1 \int_\Delta \tilde{\omega}(A',B',\tau)y^T(\tau)e^{\beta^* A'(\tau-s)}\, B'\omega(dA',dB'/\tau)\, \omega(d\tau)$$

We claim that:

$$\int_0^1 z(s)(v(s)-v^*(s))\mu(ds) \geq 0 \;\Rightarrow\; z(s)v(s) \geq z(s)v^*(s) \;\text{ for } \mu\text{-a.a. } s\in[0,1].$$

To see this assume that there exists some $v_o\in U$ such that $z(s)v_o(s) < z(s)v^*(s)$ for $s\in\delta T \subset [0,1]$ where $\mu(\delta T) > 0$. Then we can form $v'\in U$ such that

$$v' \equiv \begin{cases} v_o & \text{when } s\in\delta T \\ v^* & \text{otherwise} \end{cases}.$$

But, the integral inequality is then not satisfied with $v'(s)$ since:

$$\int_0^1 z(s)(v'(s)-v^*(s))\mu(ds) = \int_{\delta T} z(s)(v'(s)-v^*(s))\mu(ds) \; < 0.$$

Now the third conclusion of Theorem 6.4 directly determines the form of the support set of the measure ω. From Theorem 6.4):

$$\tilde{\omega}(p)\cdot y^T(p;v^*,\beta^*)y(p;v^*,\beta^*) = \max_{a\in[-\varepsilon,M]} \tilde{\omega}(p)\cdot a \qquad \text{for } \omega\text{-a.a. } p\in P.$$

Recall that $\tilde{\omega}(p) = \pm 1$. Hence: $y^T(p;v^*,\beta^*)y(p;v^*,\beta^*) = M$ when $\tilde{\omega} = 1$. If $\tilde{\omega} = -1$ the equality above cannot be satisfied. Hence $\tilde{\omega} = 1$ for ω-a.a. $p \in P$ and the measure ω has the support set defined as follows:

$$\Omega = \{p \in P: y^T(p;v^*,\beta^*)y(p;v^*,\beta^*) = M\}.$$

The expression for $z(s)$ simplifies to:

$$z(s) = \int_s^1 \int_\Delta y^T(\tau) e^{\beta^* A'(\tau - s)} B' \omega(dA', dB'/\tau)\, \omega(d\tau). \qquad \blacklozenge$$

The above theorem derives necessary conditions for the optimal solution (v^*,β^*) to Problem 6.9. Now consider conclusion (ii) of the above theorem.

$$z(s)\, v(s) \geq z(s)\, v^*(s) \text{ for } \mu\text{-a.a. } s \in [0,1] \text{ and for any } v \in U_{[0,1]}$$

It follows that when each component of $z(s) \neq 0$ the optimal input solution $v^*(s)$ is bang-bang and alternates between the maximum allowed input bounds (namely $\pm K$). In other words when $z^j(s) \neq 0$ for some $j \in \{1,2,\ldots m\}$, this implies that for a.a. $s \in [0,1]$:

$$v^{*j}(s) = \begin{cases} -K & \text{when } z^j(s) > 0 \\ K & \text{when } z^j(s) < 0 \end{cases}.$$

This partially solves our questions about the characteristics of the optimal solution v^*. However we cannot conclude anything about the corresponding components of the solution over intervals when some components of the function $z(s)$ are zero. A completely bang-bang solution would be extremely favorable from an engineering point of view, for its ease of implementation and numerical computation. However as we show in the following example, this may not always be true in this case. We will show that some components of the switching function $z(s)$ can turn out to be zero over non-zero intervals of time.

We would like to know more about the function z(s), in particular, whether some components of $z(s)$ could be zero over contiguous subintervals of $[0,1]$. Recall that for the solution to be purely bang-bang each component of the switching function $z(s)$ needs to be non-zero almost everywhere on [0,1]. We note that our problem is linear time optimal and in most such problems studied in literature the solution turned out to be bang-bang. (See Pontryagin et al. 1962). However it is interesting to note that in this case, it does *not* always hold. In general, it is not true that the optimal solution is bang-bang and hence some components of $z(s)$ are zero over subintervals of non-zero measure in [0,1]. We present the following example:

Example 6.13: Find $\max_{u \in U} t_f$ subject to the following constraints:

$$\begin{cases} \dot{x}(t) = ax(t) + u(t) & 0 \le t \le t_f \text{ and } |u(t)| \le 2 \\ x(0) = 1 & \\ x^T(t)x(t) \le M & \text{for } 0 \le t \le t_f \text{ and for all } a \in [1.2, 1.4] \end{cases}$$

We provide a solution to the problem for $M = 1.96$.

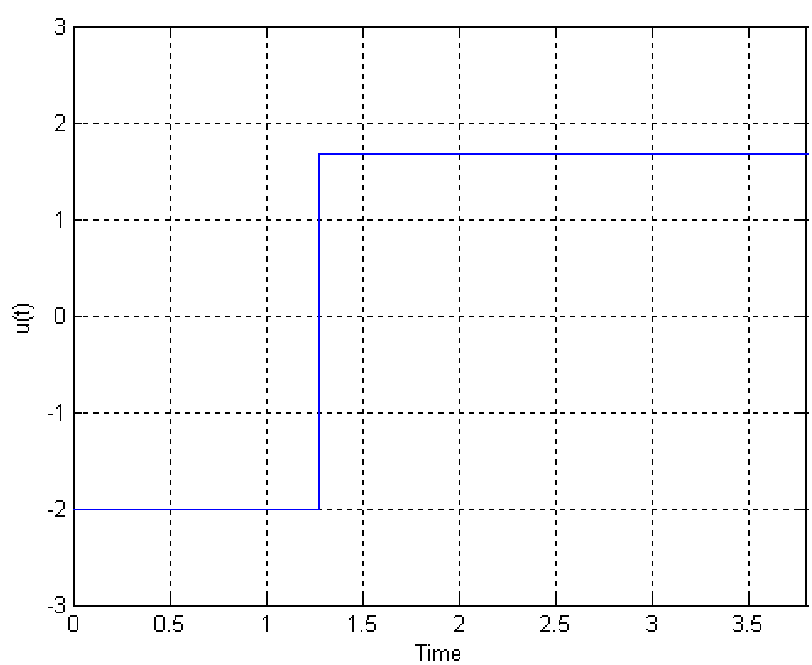

Figure 6-1: Optimal input has one switch: $M = 1.96$, $t_f = 3.7$

The solution is bang-bang only over the first part [0,1.27] of the interval. Over the rest of the interval until t_f the optimal input though constant (=1.67) is not ±2. For a clear understanding of the behavior of the system for different values of the uncertainty we have plotted the corresponding trajectories for ten different uncertainty values of the pole 'a'. The solution to this problem was calculated using brute force techniques, where both the time axis (0 to 4 seconds) and the amplitude of the input [-2,2] were discretized thereby forming a grid. Every input was checked on this grid to find the best solution. The continuous time solution was found by interpolation of the discrete optimal solution.

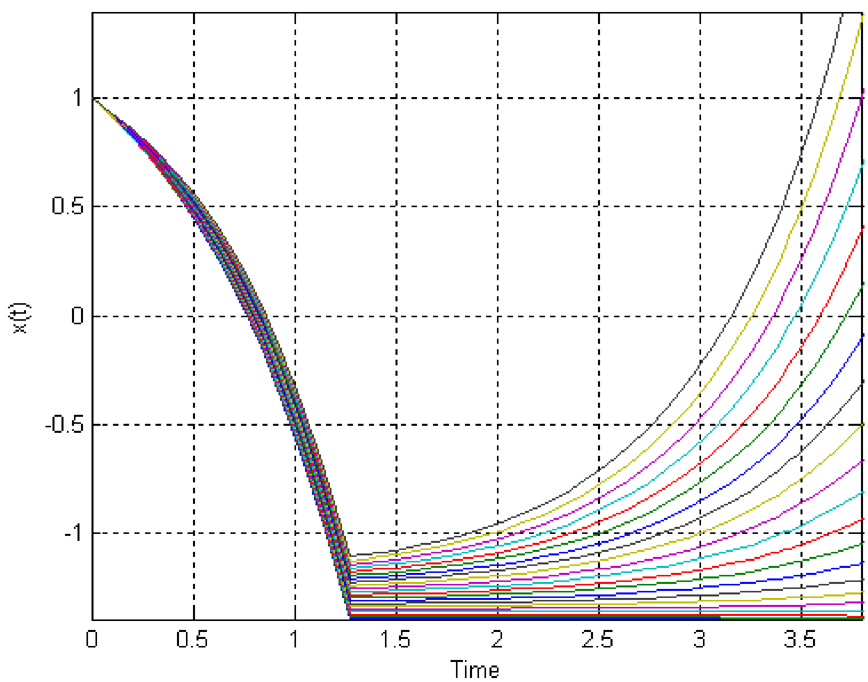

Figure 6-2: Trajectories for ten different uncertainty values of 'a'; M = 1.96, t_f = 3.7

While this feature of the optimal solution v^* is interesting because of its anomaly with other known solutions to linear time optimal problems, studied in literature, it is inconvenient from an engineering point of view. It is much easier to compute a bang-bang solution than otherwise, since for a bang-bang solution, only the switching instants need to be computed to effectively know the solution. Hence for a bang-bang solution the dynamic optimization with high complexity reduces to a simple static optimization over the space of switching instants. In the next section we show, that even over intervals where some components of $z(s)$ are zero and consequently the optimal solution v^* is possibly not bang-bang, there exists a bang-bang input v' which approximates v^* in an appropriate sense.

Bang-Bang Approximation to v^*

As we discussed in the last section, it would be interesting to know if the optimal solution, which may not be bang-bang, could be approximated by a bang-bang input in any sense. Recall that the objective of the optimal input is to maintain the inequality $x^T x \le M$ for every trajectory of the error, corresponding to any uncertainty value, during the maximal time interval t_f. We show that the optimal input u^* can be approximated by a bang-bang input u', in the sense that $x^T(u',D)x(u',D) \le M + \varepsilon$ for every value of the uncertainty $D \in \Delta$, for at least t_f seconds. Here ε can be made arbitrarily small by increasing the number of switches in u'.

Lemma 6.14: There exists a bang-bang function $u' \in U_{[0,\beta^*]}$ with a finite number of switches, such that $x^T(u';D)x(u';D) \le M + \varepsilon$ for every $D \in \Delta$ and for at least t_f seconds. Here, ε can be made arbitrarily small by increasing the number of switches in u'.

Proof: Let $x^*(t,D)$ and $x'(t,D)$ denote the state vectors, corresponding to uncertainty value $D \equiv (D_A, D_B) \in \Delta$ at instance t when the optimal control u^* and u' are applied respectively between [0,t]. Since they start from the same initial condition,

$$x^*(t,D) - x'(t,D) = e^{A't}\int_0^t e^{-A'\tau}B'[u^* - u']d\tau$$

We claim that for small enough t it is possible to switch u' appropriately so that

$$e^{A't}\int_0^t e^{-A'\tau}B'[u^*-u']d\tau = 0 \quad \forall(A',B')\in((A+\Delta_A)\times(B+\Delta_B))$$

Since e^{At} is invertible always, it is enough to prove that

$$\int_0^t e^{-A'\tau}B'[u^*-u']d\tau = 0.$$

Let t be small enough so that e^{-At} can be considered constant over $[0,t]$. We now propose to switch each component of $u'\equiv\{u_1',u_2',\ldots,u_m'\}$ once between $\pm K$ during $[0,t]$. The interval before the switch we denote by t_{i1} and the interval after the switch by t_{i2} corresponding to the i^{th} input u_i. The switching times have to add up to the total time t, i.e., $t_{i1} + t_{i2} = t$.

Hence:

$$\int_0^t e^{-A'\tau}B'[u^*-u']d\tau = e^{-A't}B'\begin{bmatrix} \Delta u_{11}t_{11} + \Delta u_{12}t_{12} \\ \Delta u_{21}t_{21} + \Delta u_{22}t_{22} \\ \cdots \\ \cdots \\ \cdots \\ \Delta u_{m1}t_{m1} + \Delta u_{m2}t_{m2} \end{bmatrix}$$

Where $\Delta u_{ij} = u_i - u_i'$ $(i = 1,2,\ldots m)$ and $j = 1$ denoted the time interval before the switch and $j=2$ implies the time after the switch.

Clearly, we can choose the switching times so as to make

$$\Delta u_{i1}t_{i1} + \Delta u_{i2}t_{i2} = 0 \text{ for every } i\in\{1,2,\ldots m\}.$$

This proves that at time instant t we can make

$$x^*(t,D) - x'(t,D) = 0$$

irrespective of the value of the uncertainty $D\equiv(D_A,D_B)$, by switching each input once. The total number of switches required for all the inputs is exactly m. This does not guarantee that $x^*(\tau,D) = x'(\tau,D)$ for all $\tau\in[0,t]$ $\forall D\in\Delta$. However because of the absolute continuity of $x(t,D)$, by choosing t small enough we can guarantee that $||x^*(t,D) - x'(t,D)|| \le \varepsilon$ for any given ε.

Recall that the optimal open loop interval is always finite. Hence the optimal interval $[0,t_f]$ can be split into arbitrarily small sub-intervals to form a partition $P\equiv\{0,t_1,t_2,\ldots,t_p,t_f\}$. We can use the above method for each subintervals $[t_i,t_{i+1}]$, $(i=1,\ldots,p)$ so that $x^*(t_i,D) = x'(t_i,D)\ \forall D\in\Delta$ for each $t_i\in P$. Also by making the partition P fine enough, we can make $\|x^*(\tau,D) - x'(\tau,D)\| \le \varepsilon$ $(\tau\in(t_i,t_{i+1}))$ for any given ε. It follows easily that $\|x'(t,D)\| \le M + \varepsilon$ for $t\in[0,t_f]$ and $\forall D\in\Delta$. ♦

Over intervals where some components of $z(s)$ are 0, the above lemma provides an approximation of the optimal solution with a bang-bang input. The approximation holds irrespective of the value of the uncertainty and hence guarantees that every state trajectory remains within the allowed error bound for the maximal period. Moreover, since the approximation is being done by a bang-bang function, all the computational advantages corresponding to a bang-bang solution are inherited by u'. In general, for an accurate approximation, the number of switches required for u' may be high. However for practical purposes, the required number of switches may be computed by repeatedly calculating the maximal time for increasing number of switches. The iteration should stop when no appreciable improvement occurs with the increase in the number of switches.

We consider Example 6.13 again to show that the non-bang-bang optimal input can be effectively approximated by a bang-bang signal, as is predicted by Lemma 6.14. The approximating bang-bang input has a total of sixteen switches. Corresponding to each uncertainty value of the pole a, the state trajectory with the actual optimal input is approximated by the trajectory resulting from the bang-bang input. To see this the reader need to compare the following figures with Figures 6-1 and 6-2. A slight relaxation of the error bound beyond $M=1.96$ allows for an identical maximal time of 3.7 seconds.

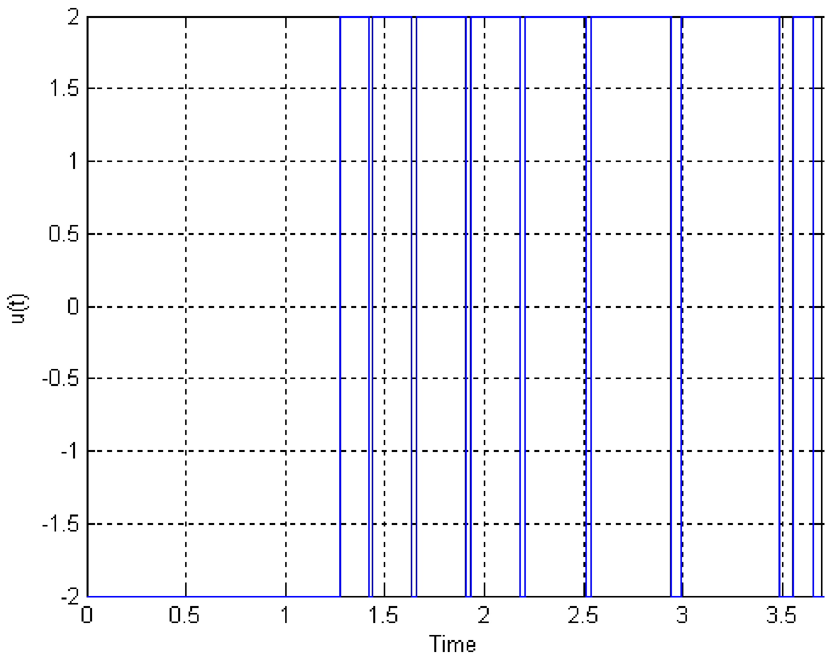

Figure 6-3: Approximate bang-bang input: 16 switches

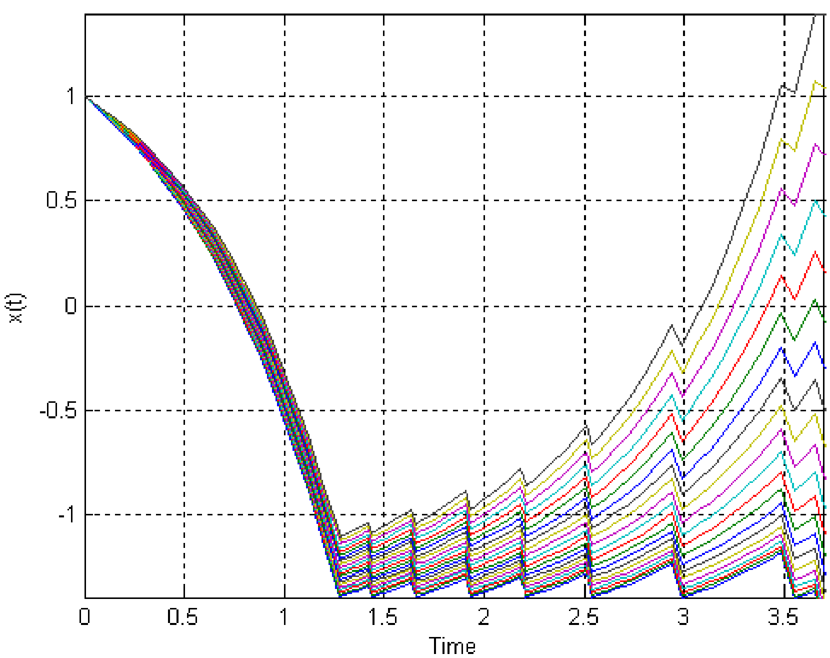

Figure 6-4: Trajectories for ten different uncertainty values of 'a'; $M = 1.96$, $t_f = 3.69$

CHAPTER 7
CONDITIONS FOR A PURELY BANG-BANG SOLUTION

Though the necessary condition derived in Theorem 6.11 hints that the optimal solution may be bang-bang, we could not rule out the possibility that some components of the switching function $z(s)$ could be zero over sub-intervals of [0,1]. In fact we discussed a case where this is the case and the optimal solution is clearly not bang-bang. Furthermore, we proposed a method in which a non-bang-bang optimal input may be approximated by a high frequency bang-bang input.

However it would be interesting to know whether there are conditions under which the optimal solution itself is purely bang-bang. For this we need to investigate two important questions about the switching function $z(s)$. Firstly, we need to find conditions under which each component of the switching function $z(s)$ is non-zero almost everywhere. Moreover we need to estimate number of zero crossings of the components of the function $z(s)$. The following assumptions were found sufficient to guarantee a purely bang-bang solution to Problem 4.4.

Assumptions 7.1:

- The uncertainty bound d is small.
- The input bound K is sufficiently small with respect to the error bound M.
- The system is controllable from each input for some value of the uncertain pair (A',B').

We first prove that all the components of the function $z(s)$ cannot be identically zero over the entire interval [0,1] for small uncertainties on the A and B matrices.

Lemma 7.2: If the maximum disturbance d is small, the function $z^j(s):[0,1]\to R$ cannot be zero μ-a.e. on [0,1] for all $j\in\{1,\dots,m\}$.

Proof: According to Theorem 6.11, for $\beta = \beta^*$,

$$\int \mathcal{D}T_2((v^*,\beta^*);(v-v^*))\omega(dp) \;=\; \int_0^1 z(s)(v(s)-v^*(s))\mu(ds).$$

Assume that $z(s) = 0$ a.e. on $[0,1]$. Then $\int \mathcal{D}T_2((v^*,\beta^*);(v-v^*))\omega(dp) = 0$ for any μ-measurable $v(s)$. Now, for $\beta = \beta^*$ the expression for $\mathcal{D}T_2((v^*,\beta^*);(v-v^*))$ may be simplified as follows :

$$\begin{aligned}
&\mathcal{D}T_2((v^*,\beta^*);(v-v^*))\\
&= y^T(\tau,A',B';\beta^*,v^*)\left\{ \int_0^\tau e^{\beta^* A'(\tau-s)}\,\beta^* B'(v(s)-v^*(s))ds \right\}\\
&= y^T(\tau,A',B';\beta^*,v^*)\left\{ \int_0^\tau e^{\beta^* A'(\tau-s)}\,\beta^* B'v(s)ds - \int_0^\tau e^{\beta^* A'(\tau-s)}\,\beta^* B'v^*(s)ds \right\}\\
&= y^T(\tau,A',B';\beta^*,v^*)\left\{ e^{\beta^* A'\tau}\,x_0 + \int_0^\tau e^{\beta^* A'(\tau-s)}\,\beta^* B'v(s)ds - y(\tau,A',B';\beta^*,v^*) \right\}\\
&= y^T(\tau,A',B';\beta^*,v^*)\ \{\, y(\tau,A',B';\beta^*,v) - y(\tau,A',B';\beta^*,v^*) \}
\end{aligned}$$

Now the support of ω was determined to be the set

$$\Omega = \{(\tau,A',B')\in P\colon y^T(\tau,A',B';v^*,\beta^*)y(\tau,A',B';v^*,\beta^*) = M\}.$$

Hence the integral expression simplifies to:

$$\begin{aligned}
&\int \mathcal{D}T_2((v^*,\beta^*);(v-v^*))\omega(dp)\\
&= \int y^T(\tau,A',B';\beta^*,v^*)y(\tau,A',B';\beta^*,v)\omega(d\tau,dA',dB') - M\int \omega(d\tau,dA',dB')\\
&= \int y^T(\tau,A',B';\beta^*,v^*)y(\tau,A',B';\beta^*,v)\omega(d\tau,dA',dB') \ - M
\end{aligned}$$

According to our assumption:

(7.3) $\quad \int y^T(\tau,A',B';\beta^*,v^*)y(\tau,A',B';\beta^*,v)\omega(d\tau,dA',dB') = M$ for every μ-measurable $v(t)$.

Hence for $v(t) = 0$ for $t\in[0,1]$ we have:

(7.4) $\quad \int y^T(\tau,A',B';\beta^*,v^*)e^{A\beta^* t}\,x_0\ \omega(d\tau,dA',dB') = M.$

Next we use some Dirac delta and Dirac delta derivatives for $v(t)$ to try to bring every trajectory to near zero. While these functions are not strictly μ-measurable, they can be thought of as the limits of very high amplitude measurable functions and the proof is considerably simplified as a result. Moreover, note that under the assumption of $x_0^T x_0 < M$, and by the definition of Ω, $z(s)$ can be expressed as the following power series in s in a small enough neighborhood about $s = 0$.

$$z(s) = \int_\Omega y^T(\tau,A',B')e^{\beta^* A'} B'\omega(dA',dB',d\tau) + s\int_\Omega y^T(\tau,A',B')e^{\beta^* A'} \beta^* A'B'\omega(dA',dB',d\tau) +$$
$$+ \frac{s^2}{2!}\int_\Omega y^T(\tau,A',B')e^{\beta^* A'} A'^2B'\omega(dA',dB',d\tau) + \ldots\ldots\ldots\ldots\ldots\ldots..$$

It follows that $z(s)$ is analytic in a small enough neighborhood about $s = 0$ and hence the assumption that $z(s) = 0$ μ-a.e. implies that $z(0) = 0$.

Choose $v(t) = -[\delta(0)\ \ \delta'(0) \ldots\ \delta^n(0)]\, C^{-1}x_0$

Since $z(0) = 0$, (7.3) is valid for such a v(t). Then:

$$\begin{aligned} y(0^+) &= x_0 - [\,(B+D_B)\ \ (A+D_A)(B+D_B) \ldots..\ (A+D_A)^n(B+D_B)\,]\, C^{-1} x_0 \\ &= [D_B\ \ AD_B + D_AB\ \ \ldots\ldots\ldots]C^{-1}x_0 \end{aligned}$$

Hence with this input (7.3) yields:

$$(7.5) \qquad \int y^T(\beta^*,v^*)e^{A\beta^* t}\, [D_B\ \ AD_B + D_AB\ \ \ldots\ldots\ldots]C^{-1}\, x_0\, \omega = M$$

Assuming that D_A and D_B are small the left hand side of the above equation can be approximated by:

$\left\{\int y^T(\beta^*,v^*)e^{A\beta^* t}\,\omega\right\} [D'_B\ \ AD'_B + D'_AB\ \ \ldots\ldots\ldots]C^{-1}\, x_0$ where D_A' and D_B' are average values with respect to the measure ω.

If we divide the above expression by the maximum allowed scalar disturbance, d (recall $||D_A|| \le d$ and $||D_B|| \le d \ \ \forall\ (D_A,D_B)\in\Delta$), then

$[D'_B\ \ AD'_B + D'_AB\ \ \ldots\ldots\ldots]C^{-1}\, x_0 = d\,[\,h_1\ \ Ah_2 + h_3B\ \ldots..]\, C^{-1}\, x_0$

where each element of the matrices h_i lies between [-1,1]. Hence for any measure ω and independent of the maximum disturbance d, the following upper bound holds:

$$[\, h_1 \;\; Ah_2 + h_3B \;\ldots..]\, C^{-1}\, x_0 \;\leq\; k_1\, x_0 \quad \text{where } k_1 \text{ is a large enough scalar.}$$

$$\Rightarrow \left\{ \int y^T(\beta^*,v^*) e^{A\beta^* t}\, \omega \right\} [D'_B \;\; AD'_B + D'_A B \;\ldots\ldots\ldots] C^{-1}\, x_0$$

$$\leq dk_1 \int y^T(\beta^*,v^*) e^{A\beta^* t}\, x_0\, \omega \;\leq dk_1\, M \quad \text{(by Equation 7.4)}$$

Now as $d \to 0$ Equation 7.5 cannot be satisfied. Hence we have a contradiction. ♦

Lemma 7.2 shows that at least one component $z^j(s)$ of the function $z(s)$ is non-zero over some part of [0,1]. The optimal solution is bang-bang over those subintervals.

In the following lemma we prove that if the input bound K is small enough compared to the allowed error bound M (by Assumption 7.1), then the optimal trajectories hit the boundary only at the end of the optimal interval $[0,t_f]$. We will show that this condition is not only enough to make optimal solution always bang-bang, but also provides more information about the zeros of the function $z(s)$. $\|.\|_2$ denotes the Euclidean norm below.

Lemma 7.6: If for some constants k_2, d_2 and d_1 defined below,

$$(M)^{1/2}\, (\|A\|_2 - d_1) > k_2\, (\|B\|_2 + d_2)\, K$$

holds, then $x^T(t)x(t) < M$ for all $D \in \Delta$, $t \in [0,\theta)$ and $x^T(\theta)x(\theta) = M$ for some $D \in \Delta$.

Proof: Recall that $J(t) = x^T(t)x(t)$. When some input $u(t)$ is applied to the system, the open loop operation must stop when for some disturbance $D=(\Delta_A,\Delta_B)$, the inequality $J(t) > M$ is satisfied. Now $J(t)$ is a continuous and differentiable function of t on any interval in $[0,\infty)$. Let $J(t_0) = M$ for some t_0 corresponding to some $x(t)$ and D. If for every permissible $u(t_0+)$, $\dot{J}(t)|_{t_0+} > 0$, then for every $\varepsilon > 0$, $J(t_0 + \varepsilon) > M$ and hence $\theta = t_0$. We derive conditions so that $\dot{J}(t)|_{t_0+} > 0$ for any permissible input and any trajectory satisfying $J(t_0) = M$.

$$\dot{J}(t)|_{t_0^+} > 0 \Leftrightarrow x^T\dot{x} > 0 \Leftrightarrow x^T(A'x + B'u) > 0.$$

We assumed A' to be unstable and hence, by the Lyapunov theorem, $x^TA'x > 0$. If $x^TB'u > 0$, then the inequality above is trivially satisfied. Hence we assume that $x^TB'u < 0$ and the last inequality becomes:

$$\begin{aligned} &x^T(A'x - B'u) > 0 \\ &\Leftarrow \| x^TA'x \|_2 > \| x^TB'u \|_2 . \\ &\Leftarrow \| x^TA'x \|_2 > \|x\|_2.\|B'\|_2.\|u\|_2 \end{aligned}$$

Now there exists $0 < k_1 \le 1$ such that $\| x^TA'x \|_2 = k_1\|x\|_2^2.\|A'\|_2$. Let k_2 be the lower bound of the values of k_1 for any value of $x(t_0)$ such that $x^T(t_0)x(t_0) = M$ and A'. Since by assumption $\|x^TA'x\|_2 > 0$, we claim that $k_2 > 0$. This can be seen by noting that $\|x^TA'x\|_2 \le M\,(\|A\|_2 + d_2)$.

Hence $k_2 = \inf\limits_{x^Tx=M,A'} \dfrac{\|x^TA'x\|}{M\,(\|A\|_2+d_2)}$. Since the right hand side is continuous in x and A', the infimum is achieved and hence $k_2 > 0$.

Hence the following implies the last inequality:

$$k_2\|x\|_2^2\,\|A'\|_2 > \|x\|_2\,\|B'\|_2\,\|u\|_2$$

$$\Leftarrow \frac{(M)^{1/2}\|A'\|_2}{\|B'\|_2} > k_3\,.\,K$$ (where k_3 is a constant resulting from the equivalence of norms: recall that $\|u\|_2 = k_2k_3\,\|u\| \le k_2k_3\,K$)

$$\Leftarrow \frac{(M)^{1/2}(\|A\|_2 - d_1)}{\|B\|_2 + d_2} > k_3\,.\,K$$

The last implication follows from:

$$\begin{aligned} &\|A'\|_2 = \| A + D_A \|_2 > \|A\|_2 - \|D_A\|_2 > \|A\|_2 - d_1 \\ &\|B'\|_2 = \| B + D_B \|_2 \le \|B\|_2 + \|D_B\|_2 \le \|B\|_2 + d_2. \end{aligned}$$

where d_1,d_2 are positive constants resulting from the equivalence of norms and the bound d on the uncertainties: $\|D_A\|, \|D_B\| \le d$. ♦

The significance of the above lemma is that it provides us with a simple condition when the optimal trajectories will hit the boundary only at one point in time and that time is necessarily at the end of the optimal time interval $[0,t_f]$. It is however evident that more than one trajectory corresponding to different values of the uncertainty (D_A,D_B), can hit the boundary at the same time.

Lemma 7.7: Under Assumption 7.1, each $z^j(s) \neq 0$ almost everywhere on (0,1) and the set of zeros of $z(s)$ do not have a limit point in (0,1).

Proof: Consider the expression for $z(s)$:

$$z(s) = \int_s^1 \int_\Delta y^T(\tau,A',B')e^{\beta^* A'(\tau-s)}\, B'\omega(dA',dB'/\tau)\ \omega(d\tau).$$

Recall that the support for ω was defined as $\Omega = \{(A',B',\tau)\in(A+\Delta_A)\times (B+\Delta_B)\times[0,1]: y^Ty = M\}$. If Lemma 7.6 holds then $\Omega = \{ (A',B',\tau)\in(A+\Delta_A)\times (B+\Delta_B)\times\{1\}: y^Ty = M_c \}$ and $z(s)$ can be simplified as follows:

$$z(s) = \int_\Delta y^T(1,A',B')e^{\beta^* A'(1-s)}\, B'\omega(dA',dB').$$

Hence for each $s\in[0,1)$ the expression for $z(s)$ is given by the following power series expansion:

$$\begin{aligned} z(s) &= \int_\Delta y^T(1,A',B')e^{\beta^* A'}\, B'\omega(dA',dB') + s\int_\Delta y^T(1,A',B')e^{\beta^* A'}\, A'B'\omega(dA',dB') + \\ &\quad \frac{s^2}{2!}\int_\Delta y^T(1,A',B')e^{\beta^* A'}\, A'^2B'\omega(dA',dB') + \ldots\ldots\ldots\ldots\ldots\ldots \\ &\approx \left\{\int_\Delta y^T(1,A',B')e^{\beta^* A'}\ \omega(dA',dB')\right\} B' + s\left\{\int_\Delta y^T(1,A',B')e^{\beta^* A'}\ \omega(dA',dB')\right\} A'B' + \\ &\quad \frac{s^2}{2!}\left\{\int_\Delta y^T(1,A',B')e^{\beta^* A'}\ \omega(dA',dB')\right\} A'^2B' + \ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

The above equality is true since the diameter of Δ_A and Δ_B are small compared to the variation in $e^{\beta^* A'}$ as A' varies over Δ_A.

Let us denote the coefficients by $(X_0, X_1, X_2, \ldots\ldots.)$.

Hence:

$$\begin{aligned} X_k &= \left\{ \int_\Delta y^T(1,A',B')e^{\beta^* A'}\ \omega(dA',dB') \right\} A'^k B' \\ &= [\, f_1\ \ f_2 \ldots.\ f_n] A'^k B' \end{aligned}$$

(Where we denote the row matrix created by the integral as $[\, f_1\ \ f_2 \ldots.\ f_n]$ and n is the order of the system.

We stack that coefficients together and expand each X_k.

$$\text{(7.8)} \qquad [X_0\ X_1 \ldots.X_k \ldots.] = [f_1\ f_2 \ldots.f_n][B'\ \ A'B' \ldots..\ A'^k B' \ldots..]$$
$$\text{or} \qquad [X_{01} \ldots.X_{0m} \mid X_{11} \ldots\mid \ldots\mid X_{k1} \ldots.] = [f_1\ f_2 \ldots.f_n][B'\ \ A'B' \ldots..\ A'^m B' \ldots..]$$

Now recall Lemma 7.2, which guaranteed that at least one component of $z^j(s)$ $(j=1,2,\ldots,m)$ is non-zero over at least a subinterval of non-zero measure in [0,1]. Without loss of generality that this is the case with the $z^1(s)$. Let us write $z^1(s)$ in terms of the coefficients defined above.

$$z^1(s) = X_{01} + X_{11}s + X_{21}s^2 + \ldots$$

Hence $z^1(s)$ is analytic and in particular a power series in s over [0,1]. It is well known that a power series is either identically zero or is non-zero almost everywhere in its domain of definition [0,1] (Conway 1978). Since by Lemma 7.2, $z^1(s)$ is non-zero at least over some interval of non-zero measure in [0,1], it follows that it is non-zero almost everywhere over [0,1]. We now write the coefficients of $z^1(s)$ in terms of Equation 7.8.

$$[X_{01}\ X_{11} \ldots\ldots\ X_{n1} \ldots\ldots] = [f_1\ f_2 \ldots.f_n][C^1 \ldots\ldots]$$

Where C^1 is the controllability matrix for the first input. Since $z^1(s)$ is non-zero, at least one of the coefficients X_{i1} $(i=0,1,\ldots.)$ is non-zero. Hence $[f_1 \ldots f_n] \neq 0$, i.e., at least one among f_i $(i=1,\ldots,n)$ is non-zero. Recall that by assumption (7.1), the controllability matrices C^j

corresponding to each of the m inputs are full rank. Hence at least one among the first n coefficients of each $z^j(s)$ $(j=1,2\ldots,m)$ is non-zero. This in turn implies that each $z^j(s)$ is non-zero almost everywhere on $(0,1)$.

This proves that $z(s)$ is a non-zero power series on $(0,1)$. It is well known that zeros of analytic functions cannot include a limit point in its domain of definition. It follows that the only possible limit points in the zeros of $z(s)$ is at $s=0,1$. Hence proved. ♦

From Theorem 7.2 and the above lemma, we conclude that $z(s)$ satisfies the following:

$z(s)v(s) \geq z(s)v^*(s)$ for a.a. $s\in[0,1]$.

Moreover each component of $z(s)\neq 0$ a.e. on $(0,1)$. Together this implies that for a.a. $s\in[0,1]$:

$$v^{*j}(s) = \begin{cases} -K & \text{when } z^j(s)>0 \\ +K & \text{when } z^j(s)<0 \end{cases}.$$

Hence, the optimal solution is purely bang-bang almost everywhere on $(0,1)$. The value of $v^*(s)$ is not known over sets of measure zero. These sets are exactly the values of s where some component of the function $z(s) = 0$. Recall that each zero crossing of any component of the switching function $z^j(s)$ $(j=1,2\ldots m)$ can potentially imply a switch between the extreme allowed values of the input $(\pm K)$. The number of such points is not finite in general and hence the solution may contain an infinite number of "switches" or sign changes between the extreme values. However, by Lemma 7.7, the zeroes of $z(s)$, or in other words the switching instants of the optimal input v^*, do not have a limit point in $(0,1)$. The only possible limit points are at $s = 0$ or 1. This implies that the optimal switching sequence may have become increasingly close together at the very beginning or the end of the optimal open loop interval. However, the open loop interval is always finite. From an engineering viewpoint, the optimal input function may be truncated near the two ends. This will result in a finite number of switches in the optimal input solution with an arbitrarily small reduction in the maximal open loop time. This will make the computation of the switching times as well as the hardware implementation feasible.

Hence the optimal v^* may be approximated for every practical purpose, by an implementable function with a finite number of switches. The engineering importance of the numerical simplification in the calculation of the optimal solution cannot be over emphasized. The optimal solution, being bang-bang, is effectively known by the computation of the switching instants. Moreover, a bang-bang solution may be implemented with the most basic in hardware and simplest of algorithms.

By a simple example, we demonstrate some of the interesting features of the problem solved above. We choose the same single dimensional system with a single input and with some uncertainty on the pole as was used in Problem 6.13. The state is the output of the system. Let us transcribe Problem 4.4 in terms of this system:

Problem 7.9: Find $\max\limits_{u \in U} t_f$ subject to the following constraints:

$$\begin{cases} \dot{x}(t) = ax(t) + u(t) & 0 \le t \le t_f \text{ and } |u(t)| \le 2 \\ x(0) = 1 & \\ x^T(t)x(t) \le M & \text{for } 0 \le t \le t_f \text{ and for all } a \in [1.2, 1.4] \end{cases}$$

We provide a solution to the problem for $M = 25$. The optimal input in this case has only one switch with the optimal switching instant at 1.3467 seconds. The maximum t_f was found to be 5.08 seconds, and in the time interval $[0, 5.08 \text{ sec}]$, every state trajectory is guaranteed to be within $[-5,5]$. For a clear understanding of the behavior of the system for different values of the uncertainty we have plotted the corresponding state trajectories for ten different uncertainty values of the pole ‘a’.

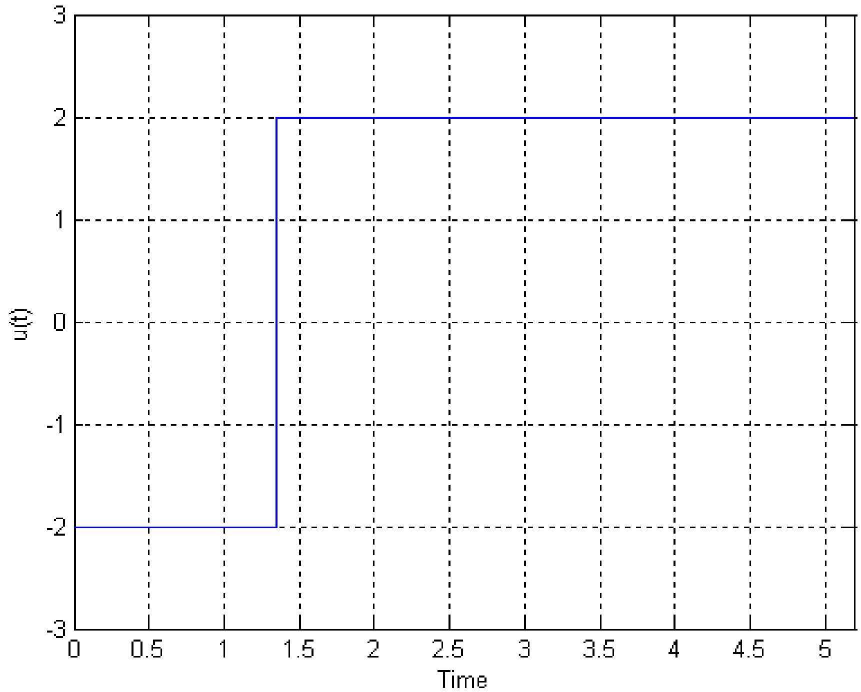

Figure 7-1: Optimal input has one switch: M = 25, t_f = 5.08

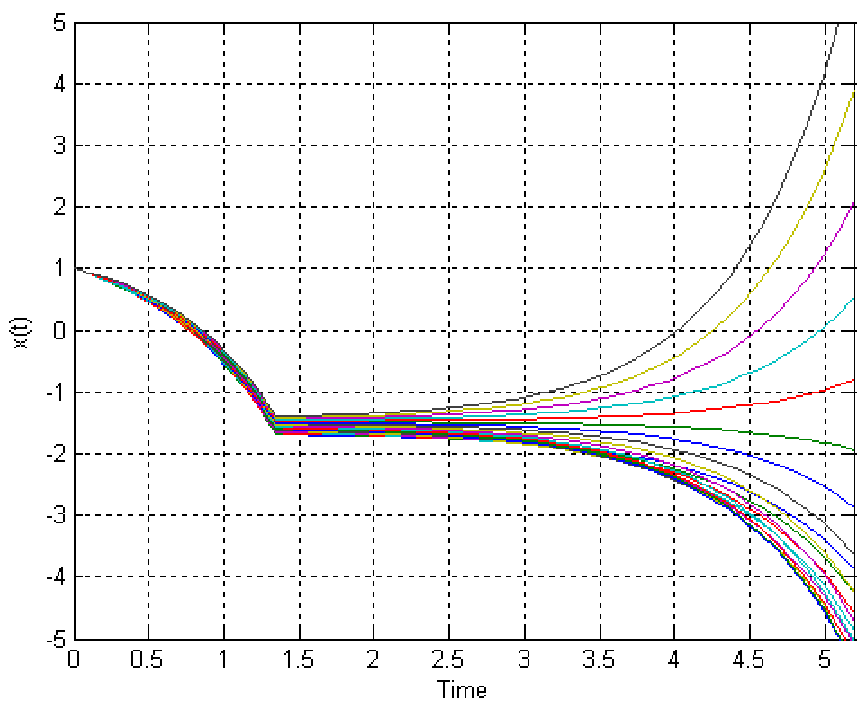

Figure 7-2: Trajectories for ten different uncertainty values of 'a'; M = 25, t_f = 5.08

CHAPTER 8
CONCLUSION AND FUTURE WORK

Conclusion

The problem of maximizing the open loop period of operation of a linear time invariant system with bounded inputs has been considered. The parameters of the controlled system are subject to bounded uncertainties. An optimal controller has been derived that maximizes the time during which the control loop can be left open, while keeping the system error within pre-specified bounds. The existence of such a control signal is first proved among the set of measurable functions. Euler-Lagrange type first order necessary conditions are then derived for calculating the optimal open loop input. It is shown that the time interval, during which the control loop can remain open, is maximized by an input, which may not be purely bang-bang over the entire maximal open loop time.

We further showed that in cases where the optimal input was not bang-bang over certain intervals, a purely bang-bang input existed which approximated the optimal input. This is of engineering importance, since the bang-bang nature of the optimal solution makes the computation of switching instants computationally feasible, as opposed to solving the entire dynamic optimization problem numerically. In the general case, the possibility of a high number of switches in the approximate input solution cannot be excluded. However, under the assumption that the input bound is small compared with the allowed error bound on the system, we have shown that the optimal input itself is purely bang-bang. Moreover, the sequence of optimal switching instants does not have a limit point in the interior of the maximal open loop interval. This facilitates the computation of the optimal open loop input. Because of the finiteness of the open loop interval of operation, the optimal input can then be approximated by a piecewise constant input with a finite number of switches.

Future Work

The results we have obtained have potential impacts on a number of application areas, including the following. Each of these applications is a potential candidate for future research that is based on the theory developed in this dissertation.

Applications

Control of Space Vehicles: Space vehicles are frequently faced with intermittent loss of signal due to obstacles in the line of vision, radiation interference, limitations in power, etc. In such situations, it is critical to know how long the system can perform within its specifications without communication with its supervisory center. This question can be directly addressed within the framework of *need-based feedback* we have developed.

Telemetry: Applications in telemetry frequently face the problem of limited power. The method we propose effectively minimizes the communication needed to control a system to a minimum and hence maximizes the longevity of the associated power sources.

Stealth Systems: Stealth and spy systems, like unmanned aerial vehicles and similar devices, prefer to reduce communications with their control center so as to reduce the chances of detection. The *need-based feedback* framework can derive the necessary control feeds to achieve this goal and reduce the probability of detection.

Agricultural Systems: Modern agricultural applications, such as soil moisture content control and fertilization control, require complex and expensive feedback processes, as most feedback data has to be collected by human experts testing in the field. Using *need-based feedback* reduces the frequency and the duration of these field tests, thus reducing costs and improving efficiency.

Biotechnology: Biological research often involves culture and preservation of cells and other organic substances under very carefully controlled environments. Regular human surveillance is necessary and extremely expensive. Our method of reducing feedback can effectively minimize requirements of human surveillance and thus reduce involved costs.

Networked Control Systems: It is increasingly common to use shared networks to control geographically separated devices. The inherent unreliability of networks may cause frequent loss of feedback signals for uncertain periods. To guarantee control objectives, it is crucial to have an estimate of how long the system can perform in open loop. Again, this can be answered in the framework of *need-based feedback* control.

Medicine: Applications are also possible in medicine and optimization of drug dosage. Typically patients are treated with drugs at regular intervals while the feedback in terms of its effectiveness is collected after large intervals. The method proposed can be potentially used to guarantee the effectiveness of the drug when measurements are not being made. Moreover the intervals after which the patient needs to be checked may be maximized.

Numerical Optimization: The actual computation of the optimal solution presents an interesting problem in numerical optimization. We conjecture that appropriate use of combinatorics may lead to a highly efficient algorithm for calculating the solution.

Theoretical Research

This problem is closely related to the viscosity solutions of the Hamilton-Jacobi-Bellman equation. It seems that the so-called "exit time problem" studied widely in relation to min-max dynamic programming is similar to the maximal open loop time studied above. But the inherent non-smoothness of the current solution makes the use of non-standard solutions necessary and hence may present an interesting future direction of research.

The theory proposed extends the results on optimal residence time of perturbed control systems. While previous results have been derived in a feedback control setting, we have shown that investigation in the framework of robust open-loop optimization framework is interesting and relevant to physical situations. A complete theory requires further research in this direction.

Questions about reachability of uncertain dynamical systems are closely intertwined with the problem investigated above. The concept of reachability of uncertain systems clearly depends on whether the system is in open or closed loop. However the correct characterization still needs to be investigated.

The problem of stabilizing an unstable system with pre-specified bounds on the inputs is related. Though this is extremely common in most controller design, a satisfactory theory is not yet available. It is thus of practical interest to calculate the minimum input bound required to stabilize an unstable uncertain system.

Further theoretical applications of the proposed theory may be found in the areas of stochastic dynamical systems, Lyapunov stability analysis of switched dynamical systems, and convex and non-convex min-max dynamical optimization.

LIST OF REFERENCES

A.V. Balakrishnan, *Introduction to Optimization Theory in Hilbert Space*, Springer-Verlag, Heidelberg, 1971.

T. Basar and G.J. Olsder, *Dynamic Noncooperative Game Theory*, London/New York: Academic Press, 1982.

P.R. Bellanger, "On the sensitivity of solutions to some optimal control problems," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, 1964.

R. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, 1957.

L.D. Berkovitz, "A differential game with no pure strategy solution", *Advances in Game Theory*, pp. 175-194, 1964, Princeton Univ. Press, Princeton, N.J.

L.D. Berkovitz, "Necessary conditions for optimal strategies in a class of differential games and control problems", *SIAM J. Control*, 5, 1967, pp. 1-24.

L.D. Berkovitz, "A variational approach to differential games", *Advances in Game Theory, Annals of Math. Study 52*, Princeton University Press, Princeton, 1964, pp. 127-174.

A.E. Bryson, Jr. and Y.C. Ho, *Applied Optimal Control*, Blaisdell, 1969.

R.M. Cannon, Jr. and E. Schmitz, "Initial experiments on the endpoint control of a flexible one-link robot", *Int. J. Robotics Res.*, Vol. 3, pp. 62-75, Fall 1984.

J.B. Conway, *Functions of One Complex Variable,* Springer, 1978.

R.J. Elliot and N. Kalton, "The existence of value in differential games ", Mem. Amer. Math. Soc. 126, 1972.

R.J. Elliot, N. Kalton and L. Markus, "Saddle points for linear differential games ", *SIAM J. Control Optimization*, Vol. 11, pp. 100-112, 1973.

B.P. Eng, "Microprocessor based servo control for gun fire control", in *Proc. IEEE IECI Annual Conf.*, Philadelphia, PA, Mar. 1979, pp. 329-333.

A.A. Feldbaum, "Dual control theory, parts I and II", *Automation and Remote Control*, Vol. 21, No. 9, April 1961, pp. 874-880 and Vol. 21, No. 11, May 1961, pp. 1033-1039. (Russian originals dated September 1960, 1240-1249 and November 1960, pp. 1453-1464)

A.A. Feldbaum, *Optimal Control Systems*, New York: Academic Press, 1965.

A.F. Filippov, "Differential equations with discontinuous right-hand side", *Amer. Math. Soc. Transl.,* Ser 2, 42, pp. 199-231, Providence, R.I., 1964.

W.H. Fleming, "The convergence problem for differential games", Journal of Mathematical Analysis and Applications, 3, pp. 102-116, 1961.

W.H. Fleming, "The convergence problem for differential games", *Advances in Game Theory*, pp. 195-210, 1964, Princeton Univ. Press, Princeton, N.J.

A. Friedman, *Differential Games*, John Wiley, 1971.

R.V. Gamkrelidze, "On some extremal problems in the theory of differential equations with applications to the theory of optimal control", J. SIAM, Ser. A Control 3, 1965, pp. 106-128.

P.R. Halmos, *A Hilbert Space Problem Book*, Springer-Verlag, New York, 1982.

N. Halyo, "Terminal area automatic navigation, guidance, and control research using the microwave landing system. Part 5-Design and development of digital automated landing system for steep final approach using modern control techniques," *NASA Contract Rep*. 3681, Apr. 1983.

M.W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, 1974.

Y.C. Ho, A.E. Bryson, Jr., and S. Baron, "Differential games and optimal pursuit-evasion strategies," IEEE Trans. Automatic Control, vol. AC-10, pp. 385-389, October 1965.

D.R. Howard and Z.V. Rekasius, "Error analysis with the maximum principle," *IEEE Trans. Automatic Control*, vol. AC-9, pp. 223-229, July 1964.

R. Isaacs, *Differential Games*, New York: Wiley, 1965.

R. Isaacs, "Differential games III", Research Memorandum RM-1411, The RAND Corporation, Santa Monica, California, 1954.

D.L. Kelendzheridze, "On the theory of optimal pursuit", *Soviet Math. Dokl.*, 2, 1961, pp. 654-656.

A.J. Koivuniemi, "Parameter optimization in systems subject to worst (bounded) disturbance", *IEEE Trans. Automatic Control*, vol. AC-11, pp. 427-433, July 1966.

A.N. Kolmogorov and S.V. Fomin, *Elements of the Theory of Functions and Functional Analysis, Vol 1 and 2*, Graylock Press, Rochester, New York, 1957 and 1961.

N.N. Krasovskii and A.I. Subbotin, *Game Theoretic Control Problems*, Springer-Verlag, Berlin, 1974.

Y.S. Ledyaev and E.F. Mishchenko, "Extrema1 problems in the theory of differential games", *Proceedings Steklov Inst. Math*., 185, pp. 165-191, 1990.

Y.S. Ledyaev and E.F. Mishchenko, "Optimization conditions for strategies in differential games of fixed duration", *Differential Equations*, 23, 1988, pp. 175-184

Y.S. Ledyaev, "Optimality conditions for feedback controls in problems of control under disturbance", *Proceedings of the 33rd Conference on Decision and Control*, Lake Buena Vista, FL, December 1994.

Y.S. Ledyaev, "The maximum principle for optimal strategies in differential games", *Soviet Math. Dokl.*, 40, pp. 545-549, 1990.

Y.S. Ledyaev, "On optimal strategies in differential games of fixed duration", *Soviet Math. Dokl.*, 33, pp. 78-80, 1986.

L.A. Liusternik and V.J. Sobolev, *Elements of Functional Analysis*, Frederick Ungar, New York, 1961.

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

S.M. Meerkov and T. Runolfsson, "Residence Time Control", *IEEE Transactions On Automatic Control*, Vol. 33, No. 4, April 1988.

E.F. Mishchenko and L. S. Pontryagin, "Linear differential games," *Soviet Math. Dokl.*, 174,. No. 1, 1967, pp. 654-656.

E.F. Mishchenko, "Pursuit and evasion problems in differential games", *Eng. Cybernetics,* 9 (1971), no. 5, 787--791

L.A. Montestruque and P.J. Antsaklis, "Stability of model-based networked control systems with time-varying transmission times," *IEEE Transactions on Automatic Control, Special Issue on Networked Control Systems*, Vol. 49, No. 9, pp.1562-1572, September 2004.

G.N. Nair, F. Fagnani, S. Zampieri and R. J. Evans, "Feedback Control under Data Rate Constraints: an Overview", to appear in *Proceedings of the IEEE (special issue on "The Emerging Technology of Networked Control Systems")*, 2007.

L.W. Neustadt, "An abstract variational theory with applications to a broad class of optimization problems I, general theory", *SIAM. J. Control*, 4, 1966, pp. 505-527.

L.W. Neustadt, "An abstract variational theory with applications to a broad class of optimization problems II, applications", *SIAM. J. Control*, 5, 1967, pp. 90-137.

J.C. Panetta and K.R. Fister, "Optimal control applied to competing chemotherapeutic cell-kill strategies", *SIAM Journal on Applied Mathematics*, Vol. 63, No. 6, 2003.

L.S. Pontryagin, V.G. Boltyansky, R.V. Gamkrelidze and E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Interscience Publishers John Wiley & Sons Inc., New York-London, 1962

L.S. Pontryagin, "On the theory of differential games", *Russ. Math. Surv.*, 1966, 21 (4),pp. 193-246.

L. Pontryagin, “Linear differential games. I, II”, *Soviet Math. Dokl*., 175, 1967 pp. 764-766

W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1966.

R.E. Skelton, “Precise pointing control of space vehicles with momentum exchange controllers”, *Sperry Rand Rep.* SP-250-754, 1973.

R.B. Vinter, “Minimax optimal control”, *SIAM Journal of Control and Optimization*, Vol. 44, No 3, pp 939-968, 2005.

R.B. Vinter, *Optimal Control*, Birkhauser, Boston, 2000.

G. Walsh, H. Ye, and L. Bushnell, “Stability analysis of networked control systems”, *Proceedings of American Control Conference*, June 1999.

J. Warga, “Minimax problems and unilateral curves in the calculus of variations”*, J. SIAM*, Ser. A, Control 3, pp. 91-105, 1965.

J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

J. Warga, “On a class of minimax control problems in the calculus of variations”, *Michigan Math. Journal*, 12, 1965, pp. 289-311.

J. Warga, “Unilateral and minimax control problems defined by integral equations”, *SIAM J. Control*, 8, 1970, pp. 372-382.

J. Warga, “On a class of pursuit and evasion problems”, *J. Differential Equations*, 9, 1971, pp. 155-167.

J. Warga, “Conflicting and minimax controls”, *J. Math. Anal. Appl*., 33, 1971, pp. 655-673.

S. Willard, *General Topology*, Addison-Wesley, Reading, MA, 1970.

H.S. Witsenhausen, “A minimax control linear problem for sampled systems”, *IEEE Transactions On Automatic Control*, Vol. AC-13, No. 1, February 1968.

L.C. Young, “Generalized curves and the existence of an attained absolute minimum in the calculus of variations”, *C. R. Sci. Lettres Varsovie*, C III 30 (1937), 212-234.

L.C. Young, *Lectures on the Calculus of Variations and Optimal Control Theory*, W. B. Saunders, Philadelphia, 1969.

E. Zeidler, *Nonlinear Functional Analysis and its Applications III* , Springer-Verlag, New York, 1985.

P.V. Zhivogyladov and R.H. Middleton, “Networked control design for linear systems”, *Automatica*, 39, pp. 743-750, 2003.

BIOGRAPHICAL SKETCH

Debraj Chakraborty was born in Calcutta, India. He earned his Bachelor of Engineering degree from Jadavpur University, Calcutta, in 2001. After completing his master's degree in electrical engineering from the Department of Electrical Engineering, Indian Institute of Technology, Kanpur, in 2003, he joined the Department of Electrical and Computer Engineering of the University of Florida, Gainesville, USA. He completed his Ph.D. in electrical and computer engineering in 2007, under the guidance of Dr. Jacob Hammer at the University of Florida. From 2003 to 2007, he was appointed an Alumni Fellow at the University of Florida.