TECHNICAL MEMORANDUM Copy______

ESL-TM-362

A PRIORI BOUNDS ON THE PERFORMANCE OF OPTIMAL SYSTEMS

by

Roberto Canales Ruiz

October, 1968

NASA Research Grant NGR-22-009-124

This technical memorandum consists of the unaltered thesis of Roberto Canales Ruiz, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology in August, 1968. This research was carried out at the M.I.T. Electronic Systems Laboratory with support extended by the National Aeronautic and Space Administration under Research Grant No. NGR-22-009-124, M I.T. DSR Project No. 76265.

Electronic Systems Laboratory
Department of Electrical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

A PRIORI BOUNDS ON THE PERFORMANCE OF OPTIMAL SYSTEMS

by

ROBERTO CANALES RUIZ

I.E. Universidad de los Andes
(1963)

B.S., Massachusetts Institute of Technology
(1963)

M.S., Massachusetts Institute of Technology
(1965)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
August, 1968

Signature of Author Roberto Canales
Department of Electrical Engineering, August, 1966

Certified by Roger W. Brockett
Thesis Supervisor

Accepted by ____________________
Chairman, Departmental Committee on Graduate Students

A PRIORI BOUNDS ON THE PERFORMANCE OF OPTIMAL SYSTEMS

by

ROBERTO CANALES RUIZ

Submitted to the Department of Electrical Engineering on August 19, 1968 in partial fulfilment of the requirements for the Degree of Doctor of Philosophy.

ABSTRACT

The problem of bounding the optimal cost of nonlinear dynamical systems is studied. Conditions are given under which it is possible to upper bound the optimal performance of a nonlinear system with a quadratic cost functional by the optimal performance of a linear system with the same cost functional. Results are stated both in the frequency and the time domain. A sharp lower bound for the optimal performance of a linear system with quadratic penalty function is given in terms of the performance of a class of suboptimal systems. The application of this result to direct evaluation of the degree of suboptimality of a given design is studied. Two numerical examples illustrate the usefulness of the results in design procedures when computational and/or structural restrictions are present.

Thesis Supervisor: Roger W. Brockett

Title: Associate Professor of Electrical Engineering

ACKNOWLEDGMENT

The author wishes to express his gratitude to Professor R. W. Brockett for his valuable suggestions, his inspiring teaching, and for the supervision of the present work. He also wants to thank Professors Michael Athans and Leonard Gould who acted as readers and made helpful suggestions during the course of this research and to Professor James R. Melcher who has been the author's Faculty Counselor during his Doctoral Program.

The author is also indebted to his colleagues at the Control Theory Group of the Electronic Systems Laboratory, Messrs. S. Greenberg, W. Levine, K. Joseph and Dr. A. Levis. The author is deeply grateful to Dr. Jan C. Willems and Mr. Michael Gruber for their fruitful discussions, many of which led to results presented in this research. The assistance of Mr. A. Rahimi in the use of the time-sharing computer is appreciated.

Special acknowledgment is extended to Mrs. Clara Conover for her excellence and promptness in typing this thesis. For the high quality of the figures in this thesis the author is thankful for the Drafting Section of ESL.

Finally, the author wants to thank his wife, Diana, not only for accepting cheerfully the tedious and difficult task of typing the draft but also for her encouragement, patience, love and understanding.

CONTENTS

CONTENTS (Contd.)

LIST OF FIGURES

CHAPTER I

INTRODUCTION

"El peor enemigo de lo bueno es lo perfecto"

1. Motivation

Some optimization problems--although they can be formulated in a precise mathematical form--are too large to solve by analytical or iterative techniques. Moreover, even if the optimal control can be found it may be very difficult to implement.

Linear programming, for instance, has been used to solve many problems of resource allocation (transportation routing, machine scheduling, product mix, oil refinery operations, etc.),[H-4†] but some problems are just too large for it--job scheduling for instance. Conceptually, linear programming could lead to an optimum assignment of start times for the thousands of jobs to be scheduled in a large shop, given some criterion, like "minimizing the idle machine time," but the number of steps necessary to reach the optimum solution--though finite--is so large as to render the method useless. In this application, linear programming is computationally inefficient since by the time the optimal solution has been reached, enough time has elapsed, even with the aid of the fastest computers, as to make the solution obsolete.

It can be said, in general, that the usual tools available for system optimization (Dynamic Programming,[B-6] Calculus of Variations,[G-3] Pontryagin's Maximum Principle[P-1] etc.), are not useful

---

† Superscripts refer to the items in the Bibliography.

on most "large problems"† because of computational infeasibility.[B-5] For this reason it is necessary to find short-cuts. The price paid, however, is that the solutions obtained are not optimal according to the minimization criterion. This type of solution, called suboptimal, is then a trade-off between a guaranteed optimal solution and a "good" one, if the computational effort to find the latter is considerably reduced.

Little attention has been paid to the problem of predicting "how far" from optimal a given suboptimal system is. When either because of computational infeasibility or because of easier or more convenient implementation (real time computation, for example) it is necessary to resort to suboptimal solutions, it becomes imperative to have at hand a way of determining the "quality" of the suboptimal system. A way of measuring the suboptimality of a system is to compare its performance with the optimal one.

In conclusion, because either

a. a computation of the optimal control is not feasible (large problems), or

b. the implementation of the control is not practical (off-line computation) or

c. initial guesses of the optimal performance are needed (iterative solutions) then, it is necessary to find suboptimal controls.

The aim of the present research is to find ways to evaluate the performance of suboptimal solutions without either computing the optimal one or using simulation.

---

† A <u>large</u> problem is defined as one in which 10^{12} or more operations are required for its solution.

When the optimal solution is impossible to compute there is, as a rule, no knowledge of the value of its performance, and, therefore, it is impossible to judge beforehand how close a suboptimal control comes to the optimal control.

Upper and lower bounds can be used to fill this gap. An a priori lower bound will put a limit on the optimal performance of the system, and therefore enables one to compute a bound on the "efficiency" of a given control. An upper bound supplies information to the effect that the optimal performance is better than certain value. The difference between the upper and lower bounds defines an interval and the search for suboptimal controls can be stopped when the performance falls within this range, if the two bounds are close enough to each other. With this criterion the "efficiency" of the controls that belong to the suboptimal interval is guaranteed to be acceptable according to the definition of optimality.

Another important area in which a priori bounds play an important role is that one of determining the effects on the performance when finite changes are made in some parameters of the system. This aspect of sensitivity analysis, from the practical point of view, is more relevant than the usual problem treated in the literature, i.e., when infinitesimal changes occur in the plant; the reason being that in actual situations the parameters used in the model of a physical system are nothing else but a rough approximation of the actual parameters. Therefore, it is important to know how the performance of an optimal system changes when a given parameter varies between some given limits. By means of upper and lower bounds we are able to guarantee that for a certain range of parameter changes the optimal performance will remain within certain values.

2. Organization

The results of the present research are divided in two farily independent parts. The first one, Chapters II and III, deals with the problem of determining a priori upper bounds on the performance of nonlinear dynamical systems for quadratic cost functionals.

Chapter II treats the problem entirely in the frequency domain and the analysis is restricted to scalar systems, that is, systems in which both the control and the output are scalar functions of time. There are two main ideas involved in the analysis presented in Chapter II. The first one is to control a nonlinear system with an optimal feedback law derived for a linear system (model) and then to express the performance of the nonlinear system associated the chosen control as the sum of two terms, one of which is the optimal performance of the model. Conditions under which the second term is negative are found; when those conditions are satisfied then it is possible to give an upper bound on the optimal performance of the nonlinear system in terms of the optimal performance of a linear system.

The second idea used in Chapter II is to find a control for a nonlinear system such that the trajectories thus generated coincide with the optimal trajectories of a linear system with respect to the same cost functional. As before the performance of the nonlinear system for that given control is expressed as the sum of two terms, one of which is the optimal performance of the linear system and conditions under which the second term is nonpositive are given.

One of the interesting features of the analysis described above is that results dealing with the problem of determining when the product of two operators is a positive operator, previously restricted to stability analysis, can be used in the determination of upper bounds.

Chapter III deals with the extensions of the results given in Chapter II to systems in which the control and the output are not necessarily scalar functions. The results in this chapter are stated in the time domain.

The second part of this thesis, Chapter IV, treats the problem of determining lower bounds on the performance of linear systems with respect to quadratic cost functionals. The main result of this chapter is a sharp lower bound on the optimal performance of a system in terms of a suboptimal performance. One of the advantages of that result is that it is possible to give a qualitative measure of the degree of suboptimality of a given (suboptimal) design entirely in terms of the design on hand. Applications of the above ideas are presented in two specific examples; one of them calling for the design of feedback control in which only part of the state vector is allowed to be fed back and the other one deals with the evaluation of a suboptimal design for a large system.

Two appendices are included. Appendix A presents some theorems about positive real functions relevant to some derivations in Chapter II. In Appendix B the proof of one lemma used in several other proofs of Chapter II is given.

Each chapter is divided into sections. Section 2.5 is the fifth section of Chapter II. Definitions, theorems, lemmas, colloraries and equations are numbered separately within each section, thus theorem 3.4.1 is the first theorem of Section 3.4. When reference is made say, to equation 9, it refers to the ninth equation of the current section, equation 3.4, is the fourth equation of Section 3 of the current chapter.

CHAPTER II

SCALAR SYSTEMS

In this chapter conditions under which it is possible to find a linear system with an optimal performance greater than the optimal performance of a given nonlinear system are derived.

1. Preliminary Concepts

In reference to dynamical systems the optimization problem is defined as follows:[A-1] Given the elements:

1. A system that satisfies a given Dynamical Equation

$$\frac{d\underline{x}(t)}{dt} = \underline{f}[\underline{x}(t), \underline{u}(t), t]$$

where

t	is the current time
$\underline{x}(t)$	is an n vector representing the state of the system at time t.
$\underline{u}(t)$	is an r vector representing the control of the system at time t.
$\underline{f}$	is a vector valued function of the state, the control and the current time.

2. A constraint set Ω in r- space to which the control $\underline{u}(t)$ should belong.

3. $\underline{x}(t_0)$ a set of initial conditions.

4. A target set T, in R^n, to which it is desired to drive the system, and

5. A cost functional

$$J = \int_{t_0}^{t_f} L[\underline{x}(t), \underline{u}(t), t]\, dt$$

Find a control u^* that satisfies the constraint $\underline{u}^* \in \Omega$, transfers the system from the initial condition to the target set, and minimizes the performance J.

Our aim is to be able to answer the following question:

How can we give bounds on $\min_{u} J$ without computing $\underline{u}^*$? A coarser classification of the elements of the optimization problem is the following:

1. A control $\underline{u}$ and a set Ω such that $\underline{u} \in \Omega$
2. The state $\underline{x}$, the initial condition $\underline{x}_0$ and the target set T
3. The cost functional J
4. The dynamic equation

One general approach to finding bounds on the optimal performance is to make simplifying assumptions on the above elements in order to reduce the computational effort and to make the changes in such a way as to be able to guarantee that the performance of the modified system fulfills the requirements of being less (or greater) than that of the original system.

In the present research only the type of modifications dealing with the dynamical equation have been treated. In this case the general statement of the problem is: Given the elements of the optimization problem then:

To find a lower bound

replace the vector valued function $\underline{f}$ by $\underline{f}_*$ such that

$$\min_{\underline{u}} J(\underline{u}) \leq \min_{\underline{u}} J(\underline{u})$$
$$\frac{dx}{dt} = \underline{f}_* \qquad\qquad \frac{dx}{dt} = \underline{f}$$

To find upper bound

replace the vector valued function $\underline{f}$ by $\underline{f}^*$ such that

$$\min_{\underline{u}} J(\underline{u}) \geq \min_{\underline{u}} J(\underline{u})$$
$$\frac{dx}{dt} = \underline{f}^* \qquad\qquad \frac{dx}{dt} = \underline{f}$$

then it can be said that in some sense a system described by the differential equation

$$\frac{dx}{dt} = \underline{f}_*[\underline{x}(t), \underline{u}(t), t] \quad (1)$$

is "better" than a system described by the differential equation

$$\frac{dx}{dt} = \underline{f}[\underline{x}(t), \underline{u}(t), t] \quad (2)$$

and in a similar fashion a system described by the differential equation (2) is in some sense "better" than the system described by the differential equation

$$\frac{dx}{dt} = \underline{f}^*[\underline{x}(t), u(t), t] \quad (3)$$

From the above considerations the problem of finding bounds on the optimal performance J is substituted by the problem of determining conditions under which it is possible to find a better system than the given one and such that the computations needed to find the optimal performance value of the modified system are simpler than those of the original system.

In the problems treated in the present research it will be assumed that Ω is R^r and that the target set is R^n.

The following definitions give a more precise meaning to the ideas presented in the paragraphs above:

Definition 1.1.1 A system Σ_2 is said to be better than a system Σ_1 † in some region S of the state space with respect to J if

† When reference is made to the two systems Σ_1 and Σ_2 it is assumed that the differential equations governing the motion of both systems is of the same order.

$$J_1^* = J_1(\underline{u}_1^*, \underline{x}_0) = \inf_{\underline{u}} {}_{\Sigma_1} J(\underline{u}, \underline{x}_0) \geq \inf_{\underline{u}} {}_{\Sigma_2} J(\underline{u}, \underline{x}_0) = J_2(\underline{u}_2^*, \underline{x}_0) = J_2^*$$

for all $\underline{x}_0 \in S$.

Definition 1.1.2 A system Σ_2 is said to be globally better than a system Σ_1 with respect to J if it is better than system Σ_1 in R^n with respect to J.

Definition 1.1.3 A system Σ_2 is said to be universally better than a system Σ_1 with respect to a class of cost functionals $\mathcal{M}$ if system Σ_2 is globally better than system Σ_1 with respect to any $J \in \mathcal{M}$.

The only problem treated in this investigation is the so called Regulator Problem. This is an optimization problem in which the cost functional is of the form†

$$J = \int_{t_0}^{t_f} [\underline{u}'(t)\underline{u}(t) + \underline{x}'(t)\underline{C}(t)\underline{C}'(t)\underline{x}(t)] \, dt$$

The only Dynamical Systems for which a solution to the regulator problem is known are those in which the dynamics are linear (Ref. A-2, K-4, K-3), and it will not be included here since there exist an extensive literature in the subject.

2. Notation and Assumptions

Denote by D the operator $\frac{d}{dt}$. Let p be a polynomial of degree n with real coefficients, q and h be polynomials of degree less than n. Let $f(t, \cdot)$ be a nonlinear time varying mapping from the real line into the real line. $\underline{x}(t)$ an element of R^n associated with the (n-1)-times differentiable function x(t) such that

† There exists more general forms of the cost functional but they can be reduced to the form presented above. (See reference K-4)

$\underline{x}(t) = \text{col}[x(t), Dx(t), \ldots, D^{n-1}x(t)]$. A function $x(t)$ is said to belong to $\mathcal{D}_o^n[0,\infty]$ if it is n times differentiable with respect to t and $\lim_{t\to\infty} \|\underline{x}(t)\|_E = 0$, where $\|\ \ \|_E$ denotes the Euclidian norm. Let S be a subset of R^n. Let Σ_1 be the system described by the equations (see Fig. 1)

$$p(D)x(t) = u(t) \quad ; \quad h(D)x(t) = y(t)$$

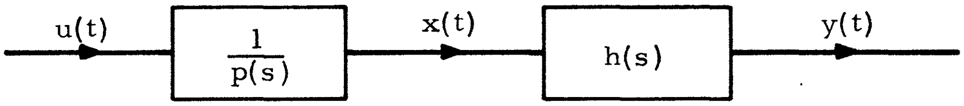

Fig. 1 The System Σ_1

and let Σ_2 be the system described by the equations (see Fig. 2)

$$p(D)x(t) + f[t, q(D)x(t)] = u(t) \quad ; \quad h(D)x(t) = y(t)$$

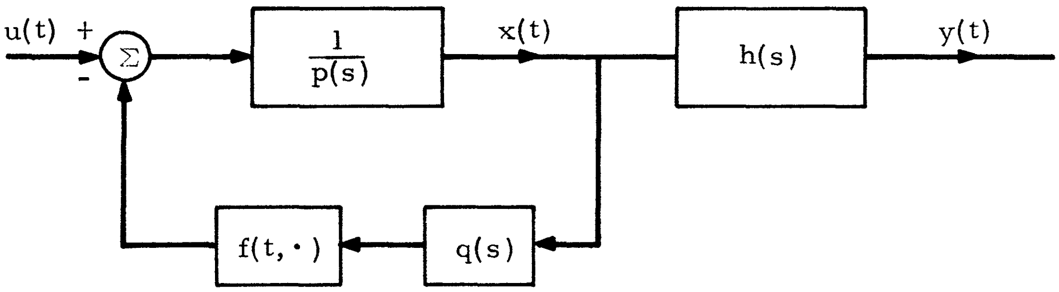

Fig. 2 The System Σ_2

Let J be the cost functional given by

$$J(u, \underline{x}_0) = \int_0^\infty \{u^2(t) + y^2(t)\}\, dt$$

It will be assumed that the polynomials p(s) and h(s) do not have common factors and that the nonlinearity $f(t, \sigma)$ satisfies the condition

$$f(t, 0) = 0$$

3. <u>Purpose</u>

The object of the present chapter is to give some sufficiency conditions on p, q, h and $f(t, \cdot)$ that will guarantee that the system Σ_2 is better, in a region S with respect to J, than the system Σ_1.

It should be recalled that the optimal feedback law for the system Σ_1 (denoted by $k_1^*(D)x(t)$) is linear and time invariant[K-1] and there are methods to compute it[B-6, L-1, L-2] while in general the optimal feedback control law for a system of the form of Σ_2 is not known when the cost functional is quadratic in u and y. If the sufficiency conditions to be derived in this chapter are satisfied then it is possible to obtain an upper bound to the optimal performance of the system Σ_2 by simply computing the optimal performance of the linear system Σ_1.

Intuitively it would be expected that if $f(t, q(D)x)$ is "close" to or in the "direction" of the optimal control law of the system Σ_1 then the system Σ_2 is better than the system Σ_1. To illustrate this point take the following example: If it is assumed that $f(t, q(D)x)$ is identically equal to the negative feedback law $k_1^*(D)x$ (the optimal feedback law for the linear system Σ_1) then, the performance of Σ_2 when no control is applied (i.e., $u(t) \equiv 0$) coincides with the optimal performance of the system Σ_1. However, as it will be shown, under less restrictive conditions, the system Σ_2 is better than the system Σ_1.

The study of the problem in the present chapter is carried out in the frequency domain, the reason being that many concepts, relevant to the solution of our problem such as passivity, are displayed more explicitly in that framework. Once the underlying ideas are understood, the extensions to multiple input-output systems becomes a less difficult task.

4. The Two Main Theorems

In this section two of the main theorems of the present chapter are derived. Theorem 2.4.1 was motivated by the observation made in Section 3. Essentially, the optimal feedback law for the linear system Σ_1 (i.e., $k_1^*(D)x$) is applied to the nonlinear system Σ_2. The performance of the system Σ_2, generated by the feedback law k_1^*, denoted by $J_2(k_1^*)$, is obtained and conditions under which

$$J_2(k_1^*) \leq J_1(k_1^*) = J_1^* \tag{1}$$

are derived. But, since by the optimality of k_2^*,

$$J_2^* = J_2(k_2^*) \leq J_2(k_1^*) \tag{2}$$

we want to guarantee that

$$J_2^* \leq J_1^*$$

The above idea is not new in the literature and has been exploited by Rissanen[R-2] and McClamrock.[M-1]

In the theorem 2.4.2, on the other hand, a different approach is used. In this case a feedback control law $(k_\beta(x(t))$ (in general time varying and nonlinear) is applied to the system Σ_2 in order to generate trajectories that are identical to the optimal trajectories of the system Σ_1. Using the same type of arguments as in theorem 2.4.1 conditions are given as to guarantee that

$$J_1^* \geq J_2(k_\beta) \geq J_2^*$$

Theorem 2.4.1: The system Σ_2 is better than the system Σ_1 in S with respect to J if

$$\text{a.} \int_0^\infty f(t, q(D)x)[p(D)x - (p\overline{p} + h\overline{h})^+(D)x]\,dt \leq 0 \;^\dagger \tag{3}$$

along trajectories satisfying

$$f(t, q(D)x) + ([pp + hh]^+(D)x) = 0 \tag{4}$$

b. All solutions of (4) with $\underline{x}_0 \epsilon$ S belong to $\mathcal{D}_o^n[0, \infty]$

Proof:

Assume that $\underline{x}_0 \epsilon$ S. Apply the optimal feedback law of system Σ_1, namely[B-1]

$$k_1^*(D)x = p(D)x - [p\overline{p} + h\overline{h}]^+(D)x \tag{5}$$

to the system Σ_2. Then the trajectories thus generated satisfy the differential equation (4).

The performance of system Σ_2 associated with the feedback law k_1 is given by

$$J_2(k_1^*) = \int_0^\infty \{[p(D)x + f(t, q(D)x)]^2 + [h(D)x]^2\}dt \tag{6}$$

with $x(t)$ satisfying the differential equation (4).

Therefore, from equations (6) and (4) it follows that

$$J_2(k_1^*) = \int_0^\infty ([p(D)x+f(t,q(D)x)]^2+[h(Dx]^2-\{f(t,q(D)x)+[p\overline{p}+h\overline{h}]^+(D)x\}^2)\,dt \tag{7}$$

† $[g]^+$ denotes the left half-plane spectral factor of the polynomial g. $\overline{p}$ is the polynomial such that $\overline{p}(s) = p(-s)$.

From the above equations it follows that

$$J_2(k_1^*) = \int_0^\infty ([p(D)x]^2 + [h(D)x]^2 - \{[p\bar{p} + h\bar{h}]^+(D)x\}^2 dt$$

$$+ 2\int_0^\infty \underline{f}(t, q(D)x)\{p(D)x - [p\bar{p} + h\bar{h}]^+(D)x\}dt \tag{8}$$

The first integral of the right hand side of (8) is independent of x(t) and is only a function of $\underline{x}_0$ and $\lim_{t\to\infty} \underline{x}(t)$.[B-1] Also if $\lim_{t\to\infty} \underline{x}(t) = \underline{0}$ that integral is identical to the optimal performance of system Σ_1[B-1] from the initial condition $\underline{x}(0) = \underline{x}_0$. But by assumption (b) if $\underline{x}(0) \in S$ and x(t) is a solution of (4) then $x(t) \in \mathcal{B}_o^n[0, \infty)$, therefore

$$J_2(k_1^*) = J_1^* + 2\int_0^\infty f(t, q(D)x)\{p(D)x - [p\bar{p} + h\bar{h}]^+(D)x\}dt \tag{9}$$

and by assumption (a) it follows that

$$J_2^* \leq J_2(k_1^*) \leq J_1^* \tag{10}$$

therefore Σ_2 is better than Σ_1 in S with respect to J.

<u>Collorary 2.4.1</u>: The system Σ_2 is better than the system Σ_1 in S with respect to J if

a. $$\int_0^\infty f[t, q(D)x][p(D)x + f(t, q(D)x)]\, dt \leq 0 \tag{11}$$

along trajectories satisfying (4), and

b. all solutions of (4) with $\underline{x}(0) \in S$ belong to $\mathcal{B}_o^n[0, \infty)$.

Collorary 2.4.2: The system Σ_2 is globally better than the system Σ_1 with respect to J if

a. Inequality (11) holds

b. The null solution of (4) is asymptotically stable in the large.

Comment:

As it was mentioned in Section 3, if $f(t, \cdot)$ is the "direction" of $-k_1(D)x(t)$ then it would be expected that the system Σ_2 is better than the system Σ_1. Equation (3) gives in some way a more precise meaning to the word "direction". Note that if $k_1(D)x(t) = -f(t, q(D)x)$ for all $x(t)$ then conditions (a) and (b) of theorem 2.4.1 are satisfied.

One of the drawbacks of the conditions given by theorem 2.4.1 and colloraries 2.4.1 and 2.4.2 is that besides testing for inequality (3) it is necessary to investigate the stability characteristics of the differential equation (4). To avoid this problem there are two alternatives: Either find conditions under which if inequality (3) is satisfied then stability condition (b) will automatically hold, or find an alternative feedback control law that will insure that the trajectories thus generated will approach zero as $t \to \infty$. We take the second alternative.

Theorem 2.4.2: System Σ_2 is better than system Σ_1 in S with respect to J if

$$\int_0^\infty f(t, q(D)x)\{2p(D)x + f(t, q(D)x)\}dt \leq 0 \tag{12}$$

along trajectories satisfying

$$[p\bar{p} + h\bar{h}]^+(D)x(t) = 0 \tag{13}$$

and $\underline{x}_0 \in S$.

Proof:

When the feedback control law

$$k_\beta(D)x = p(D)x + f(t, q(D)x) - [p\bar{p} + h\bar{h}]^+(D)x \tag{14}$$

is applied to system Σ_2 the trajectories thus generated satisfy the differential equation (13). The value of the performance of system Σ_2 associated with k_β is given by

$$J_2(k_\beta) = \int_0^\infty \{[p(D)x + f(t, q(D)x)]^2 + [h(D)x]^2\} dt \tag{15}$$

where $x(t)$ satisfies (13) and $\underline{x}(0) \in S$. Since the trajectories satisfying (13) and the initial condition are the optimal trajectories of the system Σ_1 then

$$\begin{aligned} J_2(k_\beta) &= \int_0^\infty \{[p(D)x]^2 + [h(D)x]^2\} dt + \int_0^\infty f(t, q(D)x)[2p(D) + f(t, q(D)x)] dt \\ &= J_1^* + \int_0^\infty f(t, q(D)x)[2p(D)x + f(t, q(D)x)] dt \end{aligned} \tag{16}$$

but by hypothesis the second integral of the right hand side of (16) is nonpositive then

$$J_2^* \leq J_2[k_\beta] \leq J_1^*$$

and it can be concluded that the system Σ_2 is better than the system Σ_1 in S **with** respect to J.

Collorary 2.4.3: System Σ_2 is globally better than the system Σ_1 if inequality (12) is satisfied along all trajectories satisfying (13).

5. On Universally Better Systems

In the previous section some sufficiency conditions that guarantee that a nonlinear system is better than a linear system have been derived. In this section specific conditions under which a nonlinear system is universally better than a linear system for a wide class of quadratic cost functionals are presented. For that purpose the class of nonlinearities that will be treated are classified as follows:

1. $f(t,\cdot) = f(\cdot)$ and f is a first and third quadrant nonlinearity.
2. $f(t,\cdot) = f(\cdot)$ and f is a monotone or odd monotone nonlinearity.
3. $f(t,\cdot) = k(t)(\cdot)$ and $k(t)$ is restricted.
4. $f(t,\cdot) = (\cdot)\, k$.

In this section the results derived in stability theory about positive operators will be used extensively.

As was mentioned in defintion 1.1.3, a system Σ_2 is universally better than a system Σ_1 with respect to certain classes of cost functionals $\mathcal{M}$. In the remainder of this chapter the class of cost functionals $\mathcal{M}$ will be any quadratic functional of the form

$$J = \int_0^\infty \{u^2(t) + y^2(t)\}dt$$

and $h(s)$ and $p(s)$ do not have common factors.

It is not clear if it is possible to find a system Σ_2 that is universally better than another system Σ_1 with respect to J. To gain some insight into what class of systems might be universally better than another consider the following problem:

Given a linear time invariant system Σ_1 when does there exist another system Σ_2 taken among the linear and time invariant systems of the same degree as Σ_1, such that the system Σ_2 is universally better than the system Σ_1?

The question of existence is completely solved by the following theorem:

Theorem 2.5.1: Given the system

$$\Sigma_1 : \; p(D)x = u \tag{1}$$

then there exists another system of the form

$$\Sigma_2 : p(D)x + q(D)x = u \quad ; \; q(s) \not\equiv 0 \tag{2}$$

that is universally better than Σ_1 if and only if $p(s)$ has one or more zeros in $\text{Re}[s] > 0$.

Proof: Sufficiency

Assume that $p(s)$ has one or more zeros in $\text{Re}[s] > 0$ then factor $p(s)$ as follows

$$p(s) = r(s)m(s)$$

such that $m(-s)$ is a Hurwitz polynomial. Make

$$q(s) = r(s)m(-s) - p(s) \tag{3}$$

and define

$$p^*(s) \triangleq p(s) + q(s) \tag{4}$$

The difference between the optimal performance of the system Σ_1 and Σ_2 is given by

$$J_2^* - J_1^* = \int_0^\infty \{[p^*(D)x]^2 + [h(D)x]^2 - ([p^*\overline{p}^* + h\overline{h}]^+(D)x)^2\}dt$$

$$-\int_0^\infty \{[p^*(D)x]^2 + [h(D)x]^2 - ([p\overline{p} + h\overline{h}]^+(D)x)^2\}dt \tag{5}$$

$$= \int_0^\infty \{[p^*(D)x]^2 - [p(D)x]^2 - ([p^*\overline{p}^* + h\overline{h}]^+ (D)x)^2$$

$$+ ([p\overline{p} + h\overline{h}]^+ (D)x)^2\}dt \tag{6}$$

but

$$p\overline{p}(D) = p^*\overline{p}^*(D) \tag{7}$$

by construction, then

$$J_2^* - J_1^* = \int_0^\infty \{[p^*(D)x]^2 - [p(D)x]^2\}dt \tag{8}$$

However, the integral of the right hand side of equation (8) is independent of path (i.e., depends only on the value of x and its first (n-l) derivatives at $t=0$) then

$$J_2^* - J_1^* = \int_0^\infty \{[m(-D)r(D)x]^2 - [m(D)r(D)x]^2\}dt \tag{9}$$

Denoting by

$$z(t) \triangleq r(D)x(t) \tag{10}$$

the integral in the right hand side of equation (9) depends on $z(t)$ and its first $(\beta\text{-}1)$ derivatives where β is the degree of m. Then since $m(-s)$ is a Hurwitz polynomial we can choose a trajectory such that

$$m(-D)\,z(t) = 0 \tag{11}$$

then

$$J_2^* - J_1^* \leq 0$$

with the inequality holding for a space of β degrees of freedom.

Necessity:

As in the sufficiency proof we will define

$$p^*(s) = p(s) + q(s)$$

We will show first that if $p^*(s)$ does not have the same zeros of $p(s)$ that are in $Re[s] < 0$ then there exists initial conditions and quadratic cost functional for which

$$J_2^* > J_1^* \tag{12}$$

Once the above statement has been proved we will consider systems Σ_2 where $p^*(s)$ differs from $p(s)$ only on the zeros with $Re[s] = 0$, then we will show that in the latter cases also there exist initial conditions and quadratic cost functionals for which inequality (12) holds.

Assume that there exists an s_0 with $Re[s_0] < 0$ and

$$p(s_0) = 0 \quad \text{while} \quad p^*(s_0) \neq 0$$

The difference between the cost functionals of the systems Σ_1 and Σ_2 is given by (6), and the integral of the right hand side of (6) is independent of path.

But by virtue of lemma 1B (Appendix B) it is possible to choose $h^*(s)$ such that

$$[p^*\bar{p}^* + h^*\bar{h}^*]^+(s_0) \equiv 0 \tag{13}$$

Now if $x(t)$ is chosen as

$$x(t) = \text{Re}\, e^{s_0 t} \tag{14}$$

it follows that

$$J_2^* - J_1^* = \int_0^\infty \{[p^*(D)x]^2 + ([p\bar{p} + h\bar{h}]^+(D)x)^2\}dt > 0 \tag{15}$$

The last part of the theorem is to show that if $p(j\omega_o) = 0$ and $p^*(j\omega_o) \neq 0$ then there exist initial conditions and quadratic cost functionals for which inequality (12) holds.

In a similar manner as the proof above choose

$$x(t) = e^{-\sigma t}\,\mathrm{Re}[e^{j\omega_o t}] \tag{16}$$

and h can be chosen (by virtue of lemma 1B)in such a way as to satisfy the condition

$$[p^*\bar{p} + h\bar{h}]^+(\sigma+j\omega_o) = 0 \tag{17}$$

If σ is made small enough it is possible to make the term $-\int_0^\infty [p(D)x]^2 dt$ of equation (6) as small as desired, therefore

$$\int_0^\infty \{[p^*(D)x]^2 + ([p\bar{p} + h\bar{h}]^+(D)x)^2 - [p(D)x]^2 dt > 0 \tag{18}$$

for σ sufficiently small, then

$$J_2^* - J_1^* > 0 \text{ for some } \sigma. \quad ||$$

Most of the result that follow are based largely on collorary 2.4.3.

As it was mentioned before, many of the results in stability theory dealing with the problem of the determining when the product of two operators is a positive operator will be used in this section. The reason is the following: Since the trajectories satisfying equation (4.13) approach zero as $t\to\infty$, then in order to guarantee that a system Σ_2 is universally better than a system Σ_1 it is only necessary to verify inequality (4. 12) for all $x(t)$ that approach zero as $t\to\infty$. With this idea in mind the next theorem follows trivially:

<u>Theorem 2.5.2</u>: System Σ_2 is universally better than system Σ_1 if

$$\int_0^\infty f(t, q(D)x)\{2p(D)x + f(t, q(D)x)\}dt \leq 0 \tag{19}$$

for all $x(t)\in \mathscr{D}_o^n[0,\infty)$.

It should be made clear that if inequality (19) hold for all $x(t) \in \mathcal{L}_o^n [0, \infty)$ then inequality (4.11) is automatically satisfied. However if (4.11) is satisfied for all $x(t) \in \mathcal{L}_o^n [0, \infty)$ it can not be concluded that the system Σ_2 is universally better than the system Σ_1 because of the additional stability condition on the solution of the differential equation (4.4).

Let us consider now the nonlinear systems mentioned at the beginning of this section:

The Class A_k

Definition 2.5.1: A nonlinear function -f( )- is said to belong to the class A_k if (see Fig. 3)

$$0 \leq \frac{f(\sigma)}{\sigma} \leq k \quad \text{for all } \sigma \qquad (20)$$

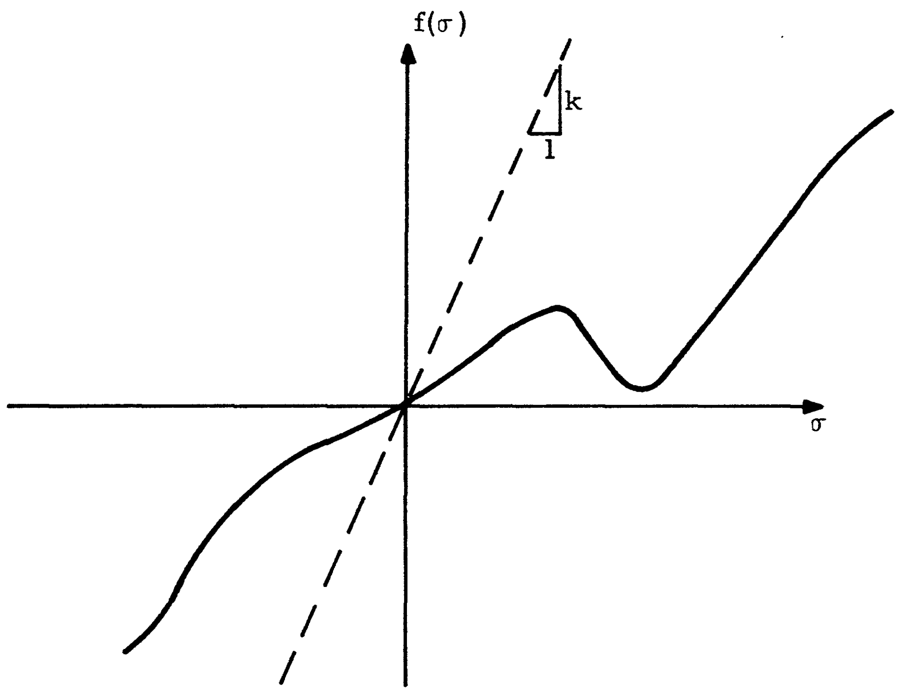

Fig. 3 An Element of the Class A_k

The following two properties of the nonlinearities of the class A_k will be used

i) $\int_0^\infty (f[z(t)] \frac{d}{dt} z(t))dt \leq 0$ provided that $z(t) \epsilon \mathscr{D}_o^1[0,\infty)$ and $f \epsilon A_\infty$ (20a)

ii) $\int_0^\infty z(t)f[z(t)][-k + \frac{f[z(t)]}{z(t)}] \leq 0$ provided that $f \epsilon A_k$ and $z(t) \epsilon \mathscr{L}_2$ (square integrable function) (20b)

Lemma 2.5.1: Inequality (19) holds for all $x(t) \epsilon \mathscr{D}_o^n[0,\infty]$ if

$$2p(s) = [cs-k]q(s) \tag{21}$$

where c is a positive constant.

Proof:

Denote by

$$a[x(t)] \stackrel{\Delta}{=} \int_0^\infty f[q(D)x]\{2p(D)x + f[q(D)x]\}\, dt$$

but by equation (21)

$$a[x(t)] = \int_0^\infty f[q(D)x]\{cDq(D)x - kq(D)x + f[q(D)x]\}dt \tag{22}$$

Denoting by

$$z(t) = q(D)x(t) \tag{23}$$

then $z(t) \epsilon \mathscr{D}_o^1[0,\infty)$ and $z(t) \epsilon \mathscr{L}_2$, therefore

$$a[x(t)] = c\int_0^\infty f[z(t)]Dz(t)dt + \int_0^\infty z(t)f[z(t)]\{-k + \frac{f[z(t)]}{z(t)}\}dt \tag{24}$$

but by properties (i) and (ii) it follows that

$$\alpha[x(t)] \leq 0$$

Theorem 2.5.3: System Σ_2 is universally better than system Σ_1 if $f \epsilon A_k$ and

$$2p(s) = [cs-k]q(s) \quad \text{with} \quad c \geq 0$$

Proof:

Follows immediately from lemma 2.5.1 and theorem 2.5.2.

Comment:

There exist a very close relationship between theorem 2.5.3 and Popov's criterion. In a similar manner as Popov's criteria theorem 2.5.3 has the same kind of limitations. From equation (21) two important restrictions of tne class of systems for which theorem 2.5.3 could give information are the following:

a. The difference in degree of the polynomials $p(s)$ and $q(s)$ should not exceed unity.

b. $p(s)$ should have at least one zero in $Re[s] > 0$. This condition, however, is a trivial consequence of theorem 2.5.1.

Example 2.5.1: Assume that system Σ_2 is given by

$$\frac{d^2}{dt^2} x(t) + f[\frac{dx(t)}{dt} + x(t)] - x(t) = u(t)$$

where $f \epsilon A_2$. Then, given any quadratic cost functional it is possible to upper bound the performance of the system Σ_2 by simply computing the optimal performance of the linear system Σ_1,

$$\frac{d^2x(t)}{dt^2} - x(t) = u(t)$$

The Classes M_k and O_k

Definition 2.5.2: A nonlinear function -f( ) - is said to belong to the class M_k if (see Fig. 4)

a. $f \epsilon A_k$

b. $(\sigma_1 - \sigma_2)[f(\sigma_1) - f(\sigma_2)] \geq 0$ for all σ_1 and σ_2.

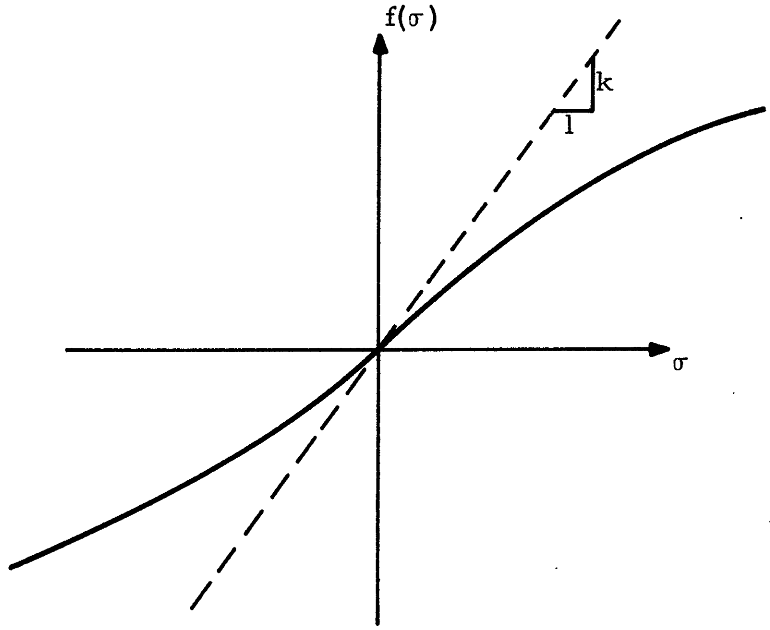

Fig. 4 An Element of the Class M_k

Properties of the class M_k Z-1

Of course the class M_k has properties (i) and (ii) of the class A_k. The following two additional properties will be needed for later developments

iii) $F(y) - F(x) + (x-y)f(x) \geq 0$ for all x and y

$$\text{where} \quad F(\sigma) = \int_0^{\sigma} f(u)du$$

Proof:

(See Fig. 5) The result is immediate from the graph. However it follows also very easily from the fact that

$$\int_x^y [f(u) - f(x)]\, du \geq 0 \qquad ||$$

$$\text{iv)} \quad \int_0^\infty x(t+\tau) f[x(t)]\, dt \leq \int_0^\infty x(t) f[x(t)]\, dt \tag{26}$$

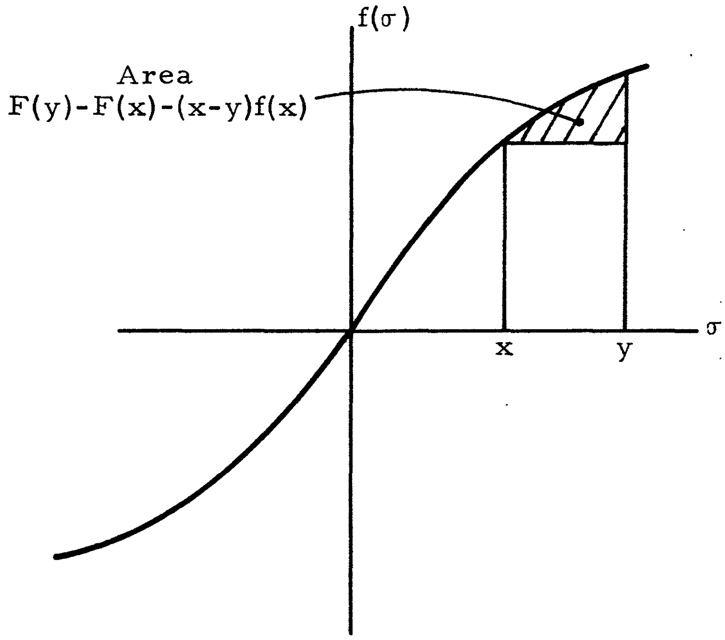

Fig. 5 If $f \in M_k$ then $F(y)-F(x) \geq (y-x)f(x)$

Proof:

From property (iii) when $y=x(t+\tau)$ and $x=x(t)$ it follows that

$$F[x(t+\tau)] - F[x(t)] + [x(t) - x(t+\tau)]\, f[x(t)] \geq 0 \tag{27}$$

Integrating with respect to t from $-\infty$ to ∞ it is obtained that

$$\int_{-\infty}^{\infty} F[x(t+\tau)]\,dt - \int_{-\infty}^{\infty} F[x(t)]\,dt + \int_{-\infty}^{\infty} x(t)f[x(t)]\,dt$$

$$- \int_{-\infty}^{\infty} x(t+\tau)f[x(t)]\,dt \geq 0 \tag{28}$$

From (28) inequality (26) follows trivially.

Lemma 2.5.2: Inequality (19) holds for any $x(t)\epsilon\mathcal{D}_o^n[0,\infty)$ if $f\epsilon M_k$ and

$$2p(s) = [-c_0 - k + c_1 s]q(s) + r(s) \tag{29}$$

where $r(s)$ is a polynomial of degree strictly less than the degree of $q(s)$ and if $[\underline{A}, \underline{b}, \underline{c}]$ is a realization^{Y-1} of $\frac{r(s)}{q(s)}$ then

$$w(t) = \underline{c}'e^{\underline{A}t}\underline{b} \geq 0 \quad \text{for all} \quad t\epsilon[0,\infty) \tag{30}$$

$$\int_0^{\infty} w(t)\,dt - \frac{r(0)}{q(0)} \leq c_0 \tag{31}$$

and c_0, c_1 are nonnegative constants

Proof:

Define

$$a[x(t)] = \int_0^{\infty} f[q(D)x]\{2p(D)x + f[q(D)x]\}dt \tag{32}$$

but since $f\epsilon M_k$ then

$$a[x(t)] \leq \int_0^{\infty} f[q(D)x]\{2p(D)x + kq(D)x\}dt \tag{33}$$

Now, substitute equation (29) into (33) to obtain

$$a[x(t)] \leq \int_0^\infty f[q(D)x][(-c_0 + c_1 D)q(D)x + r(D)x]\,dt \tag{34}$$

Define

$$z(t) \triangleq q(D)x(t) \qquad \text{then}$$
$$r(D)x(t) = \int_0^\infty w(\tau)z(t-\tau)\,d\tau \tag{35}$$

Therefore

$$a[x(t)] \leq c_1 \int_0^\infty f[z(t)]\,Dz(t)\,dt + \int_0^\infty f[z(t)]\{-c_0 z(t) + \int_0^t w(\tau)z(t-\tau)\,d\tau\}\,dt \tag{36}$$

The first integral on the right hand side of inequality (36) is negative by property (i) the fact $c_1 \geq 0$ and $z(t) \in \mathcal{L}_0^1[0,\infty)$. Define now

$$g_\sigma(t) = \begin{cases} 0 & \text{for } t < 0 \\ g(t) & \text{for } 0 \leq t \leq \sigma \\ 0 & \text{for } t > \sigma \end{cases}$$

then,

$$a[x(t)] \leq \lim_{\sigma \to \infty} \int_0^\infty f[z(t)]\{-c_0 z(t) + \int_0^t w(\tau)z(t-\tau)\,d\tau\}\,dt$$

$$= \lim_{\sigma \to \infty} \int_0^\infty f[z_\sigma(t)]\{-c_0 z_\sigma(t) + \int_0^\infty w(\tau)z_\sigma(t-\tau)\,d\tau\}\,dt$$

$$= \lim_{\sigma \to \infty} \left[\int_{-\infty}^\infty f[z_\sigma(t)][-c_0 z_\sigma(t)]\,dt + \int_0^\infty w(\tau)\int_0^\infty f[z_\sigma(t)]\,z_\sigma(t-\tau)\,d\tau\,dt\right] \tag{37}$$

by using property (iv) and condition (30) it follows that

$$a[x(t)] \leq \lim_{\sigma \to \infty} \left[\int_{-\infty}^{\infty} f[z_\sigma(t)][-c_0 z_\sigma(t)]\, dt + \int_0^{\infty} w(\tau)d\tau \int_{-\infty}^{\infty} f[z_\sigma(t)]\, z_\sigma(t)\, dt \right] \tag{38}$$

then

$$a[x(t)] \leq \int_{-\infty}^{\infty} f[z(t)]\, z(t)[-c_0 + \int_0^{\infty} w(\tau)d\tau)]\, dt$$

and finally, by inequality (31) it can be concluded that

$$a[x(t)] \leq 0 \qquad ||$$

Theorem 2.5.4: System Σ_2 is universally better than system Σ_1 if $f \in M_k$ and

i) there exist positive constants c_0 and c_1 such that

$$2p(s) = [-c_0 - k + c_1 s] q(s) + r(s)$$

ii) the degree of $r(s)$ is strictly less than the degree of $q(s)$ and if $[\underline{A}, \underline{b}, \underline{c}]$ is a realization of $\frac{r(s)}{q(s)}$ then

$$w(t) = \underline{c}' e^{\underline{A}t} \underline{b} \geq 0 \quad \text{for all } t \in [0, \infty)$$

iii) $\int_0^{\infty} w(t)dt = \frac{r(0)}{q(0)} \leq c_0$

Proof:

The above statement is an immediate consequence of lemma 2.5.2 and theorem 2.5.2.

Definition 2.5.3: A nonlinear function -f( )- is said to belong to the class O_k if

d) $f \epsilon M_k$

e) $f(\sigma) = -f(-\sigma)$ ||

The additional property of the class O_k that will be used in lemma 2.5.3 is the following

$$\text{v)} \quad \left| \int_{-\infty}^{\infty} x(t+\tau) \cdot f[x(t)]\, dt \right| \leq \int_{-\infty}^{\infty} x(t) f[x(t)]\, dt \tag{39}$$

Proof:

In inequality (28) replace $x(t+\tau)$ by $-x(t+\tau)$ then

$$\int_{-\infty}^{\infty} F[-x(t+\tau)]\, dt - \int_{-\infty}^{\infty} F[x(t)] + \int_{-\infty}^{\infty} x(t) f[x(t)]\, dt \tag{40}$$

(40)

$$+ \int_{-\infty}^{\infty} x(t+\tau) f[x(t)]\, dt \geq 0$$

but by property (e), since $F[-x(t+\tau)] = F[x(t+\tau)]$ then

$$-\int_{-\infty}^{\infty} x(t+\tau) f[x(t)]\, dt \leq \int_{-\infty}^{\infty} x(t) f[x(t)]\, dt \tag{40a}$$

Therefore inequality (40a) in conjunction with inequality (26) gives inequality (39).

With the additional property of nonlinearities of the class O_k less restrictive conditions on q(s) and p(s) can be derived in order to satisfy inequality (19). The following lemma gives those results:

Lemma 2.5.3: Inequality (19) holds for any $x(t) \epsilon \mathcal{L}_o^n[0,\infty)$ if $f \epsilon O_k$ and

$$2p(s) = [-c_0 - k + c_1 s]q(s) + r(s) \tag{41}$$

where $r(s)$ is a polynomial of degree strictly less than the degree of $q(s)$ and if $[\underline{A}, \underline{b}, \underline{c}]$ is a realization of $\frac{r(s)}{q(s)}$ then

$$\int_0^\infty |\underline{c}' e^{\underline{A}t} \underline{b}| dt \leq c_0 \tag{42}$$

where c_0 and c_1 are nonnegative constants.

Proof:

The proof of this lemma follows the same lines as the proof of lemma 2.5.2 with the difference that after inequality (37) the fact that

$$\int_0^\infty \underline{c}' e^{\underline{A}\tau} \underline{b} \int_{-\infty}^\infty f[z_\sigma(t)] z_\sigma(t-\tau) dt d\tau$$

$$\leq \int_0^\infty |c' e^{\underline{A}\tau} \underline{b}| \left| \int_{-\infty}^\infty f[z_\sigma(t)] z_\sigma(t) dt \right| d\tau \tag{43}$$

is used. The remainder of the proof is the same.

Theorem 2.5.5: A system Σ_2 is universally better than a system Σ_1 if $f \epsilon O_k$ and conditions (41) and (42) hold.

Proof:

Direct consequence of lemma 2.5.3 and theorem 2.5.2.

The Time Varying Class

In this section sufficient conditions that guarantee that a system Σ_2 in which $f(t, q(D)x) = k(t)q(D)x$ is universally better than a system Σ_1 are derived.

Lemma 2.5.4: [Gruber and Willems][G-1] The integral

$$a[x(t)] = \int_0^\infty k(t)m(D)x(t)n(D)x(t)dt$$ is nonpositive for all

a) $x(t) \in \mathscr{D}_o^n[0, \infty)$

b) $0 \leq k(t) < \infty$ for all $t > 0$

c) $\frac{dk(t)}{dt} \geq 0$ for all $t > 0$ (44)

if and only if

$$-\frac{m(-s)}{n(-s)} \text{ is positive real} \tag{45}$$

<u>Proof</u>:

Define

$$a_T[x(t)] = \int_0^T k(t)m(D)x(t)n(D)x(t)\,dt$$

then integrating by parts the above integral it follows that

$$a_T[x(t)] = -k(t)\int_t^\infty m(D)x(\tau)n(D)x(\tau)\,d\tau \Bigg|_{t=0}^{t=T} + \int_0^T \frac{dk(t)}{dt}\int_t^\infty m(D)x(\tau)n(D)x(\tau)\,d\tau\,dt \tag{46}$$

or

$$a_T[x(t)] = -k(T)\int_T^\infty m(D)x(\tau)n(D)x(\tau)\,d\tau + k(0)\int_0^\infty m(D)x(\tau)n(D)x(\tau)\,d\tau + \int_0^T \frac{dk(t)}{dt}\int_t^\infty m(D)x(\tau)n(D)x(\tau)\,d\tau\,dt \tag{47}$$

since

$$a[x(t)] = \lim_{T\to\infty} a_T[x(t)]$$

then

$$a[x(t)] = \lim_{T \to \infty} \{-k(T) \int_T^{\infty} m(D)x(\tau)n(D)x(\tau)d\tau\} +$$

$$k(0) \int_0^{\infty} m(D)x(t)n(D)x(t)\,dt + \int_0^{\infty} \frac{dk(t)}{dt} \int_t^{\infty} m(D)x(\tau)n(D)x(\tau)\,d\tau\,dt \tag{48}$$

The first term in the above expression is zero since by condition (44b) k(T) is bounded for all $T > 0$; the second term is nonpositive from (44a), (44b) and theorem 2A (Appendix). The third term is also nonpositive since its integrand is, by use of (44a), (44b) and theorem (2A) also nonpositive. The only if part follows from theorem 2A.

Lemma 2.5.5: The integral

$$a[x(t)] = \int_0^{\infty} k(t)m(D)x(t)n(D)x(t)\,dt$$

is nonpositive for all

a) $x(t) \in \mathcal{D}_o^n[0, \infty)$

b) $0 \leq k(t) < \infty$ (49)

c) $\frac{dk(t)}{dt} \geq -2\gamma k(t)$

if and only if

$$-\frac{m(-s+\gamma)}{n(-s+\gamma)} \text{ is positive real} \tag{50}$$

Proof:

Let us introduce some preliminary facts

i) $m(D)[x(t)e^{-\beta t}] = [m(D-\beta)x(t)]e^{-\beta t}$ (51)

ii) there exists a constant matrix $\underline{M}$ such that

$$(m(D+\alpha)[x(t)e^{-\gamma t}])(n(D+\alpha)[x(t)e^{-\gamma t}])=e^{-2\gamma t}\underline{x}'(t)\underline{M}\underline{x}(t) \tag{52}$$

Let

$$a[x(t)] = \int_0^\infty k(t)e^{2\gamma t}[m(D)x(t)]e^{-\gamma t}[n(D)x(t)]e^{-\gamma t}dt$$

by virtue of (51) it follows that

$$a[x(t)] = \int_0^\infty k(t)e^{2\gamma t}(m(D+\alpha)[x(t)e^{-\gamma t}](n(D+\alpha)[x(t)e^{-\gamma t}]dt \tag{53}$$

proceeding in a similar fashion as in the proof of the previous lemma, after integration by parts of (53) and by taking the limit as $T\to\infty$ it is obtained that

$$a[x(t)] = \lim_{T\to\infty}\{-k(T)e^{2\gamma t}\int_T^\infty [m(D+\gamma)(x(t)e^{-\gamma t})][n(D+\gamma)(x(t)e^{-\gamma t})]dt\}$$

$$+ k(0)\int_0^\infty [m(D)+\gamma)(x(t)e^{-\gamma t})][n(D+\gamma)(x(t)e^{-\gamma t}]dt \tag{54}$$

$$+ \int_0 \frac{dk(t)}{dt} + 2\gamma k(t)]e^{2\gamma t}\int_t^\infty [m(D+\gamma)(xe^{-\gamma\sigma})][n(D+\gamma)(xe^{-\gamma\sigma})]d\sigma\ dt$$

The last two terms in the expression above are nonpositive, since $x(t)e^{-\gamma t}\in\mathcal{B}_o^n[0,\infty)$ for $\gamma\geq 0$, (49b), (49c) and theorem 2A. It remains to be shown that the first term of the right hand side of (54) is zero. That fact can be proved as follows:

$$|k(T)e^{2\gamma t}\int_T^\infty [m(D+\gamma)(x(t)e^{-\gamma t})][n(D+\gamma)(x(t)e^{-\gamma t})]dt|$$

. . . . Continued on next page

$$\leq k(T)e^{2\gamma t}\int_T^\infty |[m(D+\gamma)(x(t)e^{-\gamma t})][n(D+\gamma)(x(t)e^{-\gamma t})]|dt$$

$$\leq k(T)e^{2\gamma t}\int_T^\infty e^{-2\gamma t}|\underline{x}'(t)\underline{M}\,\underline{x}(t)|dt \quad \text{by virtue of (ii)}$$

$$\leq k(T)e^{2\gamma t}\int_T^\infty e^{-2(\gamma+\epsilon)t}\|\underline{M}\|_E dt \quad \text{for some } \epsilon > 0 \text{ by virtue of (44a)}$$

$$= \frac{k(T)e^{-2\epsilon T}}{2(\gamma+\epsilon)} \|\underline{M}\|_E$$

and

$$\lim_{T\to\infty} \frac{k(T)\|\underline{M}\|_2}{2(\gamma+\epsilon)} e^{-2\epsilon t} = 0$$

Theorem 2.5.6: The system Σ_2 is universally better than the system Σ_1 for all $f(t, q(D)x) = k(t)q(D)x$ if

a) $0 < k(t) < k^*$

b) $\dfrac{\frac{dk(t)}{dt}}{k(t)} \geq -2\gamma$

and c) $\dfrac{-2p(-s+\gamma)}{q(-s+\gamma)} - k^*$ is positive real

Proof:

The inequality

$$\int_0^\infty k(t)q(D)x[2p(D)x+k(t)q(D)x]\,dt$$

$$\leq \int_0^\infty k(t)\, q(D)x[2p(D)x + k^*q(D)x]\,dt \leq 0$$

in conjunction with theorem 2.5.2 leads to the claim of the theorem.

The Linear Time Invariant Class

In this section it will be assumed that $f(t, q(D)x)$ is of the form $kq(D)x$.

Lemma 2.5.6: Inequality (19) holds for all $x(t) \epsilon \mathcal{D}_0^n[0,\infty]$ if and only if

$$\frac{2p(s)}{q(s)} = -Z(-s) - k \tag{54a}$$

and $Z(s)$ is positive real.

Proof:

Let us define

$$\alpha[x(t)] \triangleq \int_0^\infty kq(D)x(t)[2p(D)x(t) + kq(D)x(t)]\,dt \tag{55}$$

$$\frac{m(-s)}{q(-s)} \triangleq Z(-s) \tag{56}$$

Then $$\alpha[x(t)] = -\int_0^\infty [kq(D)x(t)][m(-D)x(t)]\,dt$$

But by virtue of theorem 2A, $\alpha[x(t)]$ is negative if and only if $-Z(-s)$ is positive real.

Theorem 2.5.8: The system Σ_2 is universally better than the system Σ_1 for $f(t, q(D)x) = kq(D)x$ if and only if condition (54a) is satisfied.

Proof:

Use lemma 2.5.6 and theorem 2.5.2.

Example 2.5.2: Consider the following specific case

$$\Sigma_2: D^2x + b_2Dx + a_2x = u$$

$$\Sigma_1: D^2x + b_1Dx + a_1x = u$$

Find conditions on a_1, a_2, b_1 and b_2 such that Σ_2 is universally better than Σ_1.

Solution:

Apply Theorem 2.5.8. In this case

$$p(s) = s^2 + b_1 s + a_1$$

and

$$q(s) = (b_2 - b_1)s + (a_2 - a_1)$$

then we want to find conditions under which

$$z(s) = \frac{2s^2 - 2b_1 s + 2a_1 - (b_2-b_1)s + a_2 - a_1}{-(b_2-b_1)s + a_1 - a_2} \quad \text{is p.r., or}$$

$$z(s) = \frac{2s^2 - (b_2+b_1)s + a_2 + a_1}{(b_2-b_1)s + a_1 - a_2}$$

These conditions are

$$-b_2 - b_1 > 0$$

$$b_2 - b_1 > 0$$

$$a_1 + a_2 > 0$$

$$a_1 - a_2 > 0$$

and they can be reduced to

$$a_1 > 0 \qquad b_1 < 0$$

$$|a_2| < a_1 \qquad |b_2| < -b_1$$

The above result is apparently new in the literature and it simply says that a second order system, with a damping less in magnitude and with a spring constant less in magnitude than a system with negative damping and positive spring constant, will have a lower quadratic criteria independent of the penalty on x and $\frac{dx}{dt}$.

6. On Globally Better Systems

The conditions under which a nonlinear system Σ_2 is universally better than a linear system derived in the previous section are easy to apply. The drawback of course is that the upper bounds obtained by computing the performance of the linear system are in general not sharp. Improvements of the upper bounds can be expected if the quadratic cost functional is specified (i.e., $-h(s)$ is fixed) since it is required that the nonlinear system performs better than the linear only with respect to a specificcost functional. Nevertheless the problem becomes more difficult because, in contrast with the results of Section 5, it is necessary to determine the negativity of an operator not in all $\mathcal{L}_o^n[0,\infty)$ but for a manifold of it. That manifold in our case is the one form by the solutions of a linear time invariant differential equation. We will present here only an extension of theorem 2.5.3.

Lemma 2.6.1: Inequality (3.12) holds for all $f \epsilon A_k$ along trajectories satisying (3.14) if there exist a nonnegative constant c_1 and c_2 arbitrary such that

$$2p(s) = [c_1 s - k]q(s) + c_2[p\bar{p} + h\bar{h}]^+(s) \tag{1}$$

Proof:

Denote by

$$\alpha[x(t)] = \int_0^\infty f(q(D)x)\{2p(D)x + f(q(D)x)\}dt \tag{2}$$

and by using equation (1), it follows that

$$\alpha[x(t)] = \int_0^\infty f(q(D)x)\{(c_1 D-k)q(D)x + c_2[p\bar{p} + h\bar{h}]^+(D)x$$

$$+ f(q(D)x)\}dt$$

but since $x(t)$ satisfies (2.13) then

$$\alpha[x(t)] = \int_0^\infty f[q(D)x]\{(c_1 D-k)q(D)x + f(q(D)x)\}dt \qquad (3)$$

Denoting by $z(t) = q(D)x(t)$ then $z(t)\in\mathcal{D}_o^1[0,\infty)$ and $z(t)\in\mathcal{L}_2[0,\infty)$ therefore

$$\alpha[x(t)] = c_1\int_0^\infty f[z(t)]Dz(t)dt + \int_0^\infty z(t)f[z(t)][-k + \frac{f[z(t)]}{z(t)}]dt \qquad (4)$$

but by virtue of properties (5.20a) and (5.20b) of the class A_k it follows that

$$\alpha[x(t)] \leq 0$$

Theorem 2.6.1: The system Σ_2 is globally better than the system Σ_1 with respect to J if condition (1) is satisfied.

Proof:

Immediate consequence of lemma 2.6.1 and collorary 2.3.3

Example 2.6.1: Consider the following second order systems

$$\Sigma_2: \frac{d^2x(t)}{dt^2} + f(\frac{dx(t)}{dt}) + b_1\frac{dx(t)}{dt} + a_1 x = u(t)$$

$$\Sigma_2: \frac{d^2x(t)}{dt^2} + b_1\frac{dx(t)}{dt} + a_1 x(t) = u(t)$$

If $f\in A_k$ find conditions on $a_1, b_1,$ and k such that the system Σ_2 is universally better than the system Σ_1.

Solution:

Apply theorem 2.6.1. In the present case

$$p(s) = s^2 + b_1 s + a_1$$

$$q(s) = s$$

then, we want to find constants c_1 and c_2 referred to in theorem 2.6.1 such that

$$2s^2 + 2b_1 s + 2a_1 = [c_1 s - k] s + c_2(s^2 + \beta_2 s + \beta_1)$$

where

$$\beta_1 > 1 \quad , \quad \beta_2 > 0 \qquad \text{(i)}$$

$$\beta_1^2 - a_1^2 \geq 0 \qquad \text{(ii)}$$

$$\beta_2^2 - b_1^2 - 2(\beta_1 - a_1) \geq 0 \qquad \text{(iii)}$$

(see Kalman[K-1])

The equations that the constants c_1, c_2 and k should satisfy are the following:

$$c_1 + c_2 = 2$$

$$-k + c_2\beta_2 = 2b_1$$

$$c_2\beta_1 = 2a_1$$

therefore $c_2 = \dfrac{2a_1}{\beta_1}$ and $c_1 = 2(1 - \dfrac{a_1}{\beta_1})$.

We can guarantee that c_1 is positive by condition (ii) above. On the other hand

$$k = c_2\beta_2 - 2b_1 = \frac{2a_1\beta_2}{\beta_1} - 2b_1$$

then, if we restrict $b_1 < 0$ and $a_1 > 0$ we have

$$k \leq |2b_1|$$

(as it can be seen the above example is an improvement over example 2.5.2 of the previous section). The conditions are then

$$f \epsilon A_{|2b_1|} \quad , \quad a_1 > 0 \quad , \quad b_1 < 0$$

Sharper results will be obtained if we specified a given cost functional. Assume for example, the following numerical values:

$$a_1 = 2 \quad , \quad b_1 = 1 \quad , \quad \beta_1 = 2 \quad \text{and} \quad \beta_2 = 4$$

then in this case

$$f \in A_2$$

CHAPTER III

GENERALIZATIONS TO THE VECTOR CASE

1. Preliminaries and Notation

In the present chapter the results presented in Chapter II will be generalized to multivariable systems. In the sequel when reference is made to system Σ_1 it will be understood to be the system described by the equations

$$\Sigma_1: \frac{d\underline{x}(t)}{dt} = \underline{A}(t)\underline{x}(t) + \underline{B}(t)u(t)$$

$$\underline{y}(t) = \underline{C}'(t)\underline{x}(t)$$

and when reference is made to system Σ_2 it will be understood to be the system described by:

$$\Sigma_2: \frac{d\underline{x}(t)}{dt} = \underline{A}(t)\underline{x}(t) + \underline{f}(t,\underline{x}) + \underline{B}(t)\underline{u}(t)$$

$$\underline{y}(t) = \underline{C}'(t)\underline{x}(t)$$

The cost functional is of the quadratic type

$$J(u,x_0) = \int_{t_0}^{T} [\underline{y}'(t)\underline{y}(t) + \underline{u}'(t)\underline{u}(t)]\, dt$$

It will be assumed that the system Σ_1 is uniformly completely controllable and uniformly completely observable.[K-4]

The nonlinearity is assumed to have the property

$$\underline{f}(t,\underline{0}) = \underline{0}$$

2. The Main Theorems

The theorems derived in the present section follow the same ideas as the theorems of Section 4 of the previous chapter.

Theorem 3.2.1: System Σ_2 is better than system Σ_1 is S with respect to J if

$$\text{i)} \quad \int_{t_0}^{T} \underline{f}'[t, \underline{x}(t)]\underline{K}(t)\underline{x}(t)dt \leq 0 \tag{1}$$

along solutions of

$$\frac{d\underline{x}(t)}{dt} = [\underline{A}(t) - \underline{B}(t)\underline{B}'(t)\underline{K}(t)]\underline{x}(t) + \underline{f}(t, \underline{x}(t)) \tag{2}$$

with $\quad \underline{x}(t_0) = \underline{x}_0 \in S$

ii) the solutions of (2) are bounded in the interval

$[t_0, T]$ for all $\underline{x}_0 \in S$

where $\underline{K}(t)$ is the unique (positive semidefinite) solution of the Riccati equation

$$\frac{d\underline{K}(t)}{dt} = -\underline{K}(t)\underline{A}(t) - \underline{A}'(t)\underline{K}(t) - \underline{C}(t)\underline{C}'(t) + \underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t) \tag{3}$$

with the boundary condition

$$\underline{K}(T) = \underline{0}$$

Proof:

Observe that the optimal feedback solution of system with respect to J is given by[K-3]

$$\underline{u}_1^*(x) = -\underline{B}'(t)\underline{K}(t)\underline{x}(t) \tag{5}$$

Apply this control law to the system Σ_2 to obtain the equation

$$\frac{d\underline{x}(t)}{dt} = [\underline{A}(t) - \underline{B}(t)\underline{B}'(t)\underline{K}(t)]\underline{x}(t) + \underline{f}[t, \underline{x}(t)] \tag{6}$$

The performance of system Σ_2 associated with the feedback law (5) is

$$J_2[\underline{u}_1^*(\underline{x})] = \int_{t_0}^{T} [\underline{x}'(t)\underline{C}(t)\underline{C}'(t)\underline{x}(t) + \underline{x}'(t)\underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{x}(t)]\,dt \tag{7}$$

along solutions of (6) or

$$J_2[\underline{u}_1^*(\underline{x})] = \int_{t_0}^{T} \{\underline{x}'(t)[\underline{C}(t)\underline{C}'(t) + \underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t)]\underline{x}(t) \tag{8}$$

$$-2[\frac{d\underline{x}(t)}{dt} - \underline{A}(t)\underline{x}(t) - \underline{f}(t, \underline{x}(t)) + \underline{B}(t)\underline{B}'(t)\underline{K}(t)\underline{x}(t)]'\underline{K}(t)\underline{x}(t)\}\,dt$$

$$= \int_{t_0}^{T} \{\underline{x}'(t)[\underline{C}(t)\underline{C}'(t) + \underline{K}(t)\underline{B}(t)\underline{B}(t)\underline{K}(t)]\underline{x}(t) - 2\frac{d\underline{x}'(t)}{dt}\underline{K}(t)\underline{x}(t) \tag{9}$$

$$+2\underline{x}'(t)\underline{A}'(t)\underline{K}(t)\underline{x}(t) + 2\underline{f}'(t, \underline{x}(t))\underline{K}(t)\underline{x}(t) - 2\underline{x}'(t)\underline{K}(t)\underline{B}(t)\underline{B}(t)\underline{K}(t)\underline{x}(t)\}\,dt$$

$$= \int_{t_0}^{T} \{\underline{x}'(t)[\underline{C}(t)\underline{C}'(t) - \underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t) + \underline{K}(t)\underline{A}(t) + \underline{A}'(t)\underline{K}(t)]\underline{x}(t)\}\,dt$$

$$- \int_{t_0}^{T} \{\underline{x}'(t)\underline{K}(t)\underline{x}(t) + \underline{x}'(t)\underline{K}(t)\underline{x}(t) - 2\underline{f}[t, \underline{x}(t)]\underline{K}(t)\underline{x}(t)\}\,dt \tag{10}$$

By using now the fact that $\underline{K}(t)$ satisfies the differential equation (3) we can write

$$J_2(\underline{u}_1^*) = \int_{t_0}^{T} \{-\underline{x}'(t)\underline{K}(t)\frac{d\underline{x}(t)}{dt} - \frac{d\underline{x}'(t)}{dt}\underline{K}(t)\underline{x}(t) - \underline{x}(t)\frac{d\underline{K}(t)}{dt}\underline{x}(t)\}dt$$

$$+ 2\int_{t_0}^{T} \{\underline{f}'(t,\underline{x}(t))\underline{K}(t)\underline{x}(t)\}dt \qquad (11)$$

The integrand of the first term of the right-hand side of (11) is a perfect differential, therefore

$$J_2(\underline{u}_1^*) = \int_{t_0}^{T} \{-\frac{d}{dt}[\underline{x}'(t)\underline{K}(t)\underline{x}(t)]\}dt + 2\int_{t_0}^{T} \{\underline{f}'(t,\underline{x}(t))\underline{K}(t)\underline{x}(t)\}dt$$

$$= \underline{x}(t_0)\underline{K}(t_0)\underline{x}(t_0) - \underline{x}'(T)\underline{K}(T)\underline{x}(t)$$

$$+ 2\int_{t_0}^{T} \{\underline{f}'[t,\underline{x}(t)]\underline{K}(t)\underline{x}(t)\}dt \qquad (12)$$

but from boundary condition (4) and from hypothesis (ii) it follows that

$$J_2(\underline{u}_1^*) = \underline{x}'(t_0)\underline{K}(t_0)\underline{x}(t_0) + 2\int_{t_0}^{T} \{\underline{f}'[t,\underline{x}(t)]\underline{K}(t)\underline{x}(t)\}dt \qquad (13)$$

but the first term of the right-hand side of (13) is exactly equal to $J_1^*(u_1) = J_1(u_1^*)^{K-3}$ then

$$J_2(u_1^*) = J_1^* + 2\int_{t_0}^{T} \{f'(t,\underline{x}(t))\underline{K}(t)\underline{x}(t)\}dt \qquad (14)$$

therefore if condition (i) holds then

$$J_2^* \leq J_1^* = J_1(u_1^*)$$

and system Σ_2 is better than system Σ_1. $||$

One of the motivations of the definitions and study of "better" systems was to find bounds on the performance of optimal systems. For that purpose some improvements over the previous theorem can be achieved by finding conditions under which

$$J_2^* \leq q\, J_1^* \quad \text{where} \quad q > 0$$

Those conditions can be given without too much effort from theorem 3.2.1 as follows:

Collorary 3.2.1: An upper bound for the optimal performance of the system Σ_2 is given by

$$J_2^*(u, \underline{x}_0) \leq q\, \underline{x}_0' \underline{K}(t_0)\underline{x}_0 \tag{15}$$

if

$$\int_{t_0}^{T} \{\underline{f}'[t, \underline{x}(t)]\underline{K}(t)\underline{x}(t)\}dt \leq \frac{q-1}{2}\, \underline{x}_0' \underline{K}(t_0)\underline{x}_0 \tag{16}$$

where $\underline{K}(t)$ is the unique solution of the differential equation (3) with the boundary condition (4) and $\underline{x}(t)$ is the solution of (2) with $\underline{x}(t_0) = \underline{x}_0$ and all solution of (2) are bounded in the interval $[t_0, T]$.

Proof:

As in the proof of the theorem after equation (13) use inequality (16) then,

$$J_2^* \leq \underline{x}'(t_0)\underline{K}(t_0)\underline{x}_0 + (q-1)\underline{x}_0'\underline{K}(t_0)\underline{x}_0 = q\underline{x}_0'\underline{K}(t_0)\underline{x}_0 \qquad ||$$

Of great interest is the case in which the system Σ_1 is time invariant and $T \to \infty$ and should be considered separately. The following theorem furnishes conditions similar to those given in theorem 3.2.1.

Theorem 3.2.2 If $\underline{A}, \underline{B}$ and $\underline{C}$ are constant matrices and $T \rightarrow \infty$ system Σ_2 is better than system Σ_1, in S with respect to J if

i) $$\int_{t_0}^{\infty} \underline{f}'[t, \underline{x}(t)]\underline{K}x(t)dt \leq 0 \qquad (17)$$

along solutions of

$$\frac{d\underline{x}(t)}{dt} = [\underline{A} - \underline{B}\underline{B}'\underline{K}]\underline{x}(t) + \underline{f}[t, \underline{x}(t)] \qquad (18)$$

with $\underline{x}(t_0) = \underline{x}_0 \epsilon S$

ii) the null solutions of (18) is asymptotically stable, $S \subset R_a$ where R_a is the region of attraction of the null solution of (18) and $\underline{K}$ is the unique positive definite solution of the algebraic equation

$$\underline{K}\,\underline{A} + \underline{A}'\underline{K} - \underline{K}\,\underline{B}\,\underline{B}'\underline{K} + \underline{C}\,\underline{C}' = \underline{0} \qquad (19)$$

Proof:

The proof of this theorem follows the same lines as the proof of theorem 3.2.1 and will be omitted. The only difference in the argument is that after equation (12) instead of using the boundary condition (4) in order to show that

$$\underline{x}'(T)\underline{K}(T)\underline{x}(T) = 0$$

the fact that $\underline{x}(t) \rightarrow \underline{0}$ as $T \rightarrow \infty$ is used based on the assumption (ii) in the statement of the theorem. ||

The counterpart of collorary 3.2.1 is the following:

Collorary 3.2.2: An upper bound for the optimal performance of the system Σ_2 is given by

$$J_2^*(\underline{u}_2, \underline{x}_0) \leq q\ \underline{x}_0 \underline{K}\,\underline{x}_0$$

if

$$\int_{t_0}^{\infty} \{\underline{f}'[t, \underline{x}(t)]\, \underline{K}\, \underline{x}(t)\} dt \leq \frac{q-1}{2}\, \underline{x}_0'\, \underline{K}\, \underline{x}_0 \qquad (20)$$

where $\underline{K}$ is the unique positive solution of (19) and $\underline{x}(t)$ is the solution of (18) with $\underline{x}(t_0) = \underline{x}_0$ and the null solution of (18) is asymptotically stable in the large. The proof of this collorary will be omitted.

Comment:

The usefulness of theorems 3.2.1 and 3.2.2 is limited by the fact that in order to apply them it is necessary to find the required stability characteristics of equations (2) and (18). That might be a difficult task. The following theorem is a simplified version of theorem 3.2.1 that overcomes the difficulty of condition (ii).

Theorem 3.2.3: If the matrices $\underline{A}, \underline{B}$ and $\underline{C}$ are constant and

$$\underline{f}'(t, \underline{x})\underline{K}\,\underline{x} < 0 \quad \text{for } \underline{x} \neq 0 \qquad (21)$$

then the system Σ_2 is globally better than the system Σ_1.

Proof:

Condition (i) of theorem 3.2.2 is trivially satisfied. To show that condition (21) implies that condition (ii) is automatically satisfied, consider the following Lyapunov function

$$v(\underline{x}) = \underline{x}'\underline{K}\,\underline{x} \qquad (22)$$

then $\frac{dv(\underline{x})}{dt}$ along solutions of (18) is given by

$$\frac{dv(\underline{x})}{dt} = -\underline{x}'\underline{C}\,\underline{C}'\underline{x} - \underline{x}'\underline{K}\,\underline{B}\,\underline{B}'\underline{K}\,\underline{x} + \underline{f}'(t, \underline{x})\underline{K}\,\underline{x}$$

which is negative definite, therefore the null solution of (18) is asymptotically stable.[H-2]

Another important class of nonlinear time varying systems that will be considered is the following:

$$\Sigma_2 : \frac{d\underline{x}(t)}{dt} = \underline{A}(t)\underline{x}(t) + \underline{B}(t)\underline{g}(t,\underline{x}(t)) + \underline{B}(t)\underline{u}(t) \tag{23}$$

$$\underline{y}(t) = \underline{C}'(t)\underline{x}(t) \tag{24}$$

Even though the above class of systems is a subclass of systems considered in theorem 3.2.1 additional conditions under which a system Σ_2 (as described by equations (23) and (24)) is better than system Σ_1 can be derived. The results are contained in the next theorem:

Theorem 3.2.4: System Σ_2 is better than system Σ_1 in S with respect to J if

$$\int_{t_0}^{T} \underline{g}'(t,\underline{x}(t))[\underline{g}(t,\underline{x}(t)) + 2\underline{B}'(t)\underline{K}(t)\underline{x}(t)]\, dt \leq 0 \tag{25}$$

along solutions of

$$\frac{d\underline{x}(t)}{dt} = [\underline{A}(t) - \underline{B}(t)\underline{B}'(t)\underline{K}(t)]\underline{x}(t) \tag{26}$$

where $\underline{K}(t)$ is the unique positive solution of (3) with boundary condition (4), and $\underline{x}_0 = \underline{x}(t_0) \epsilon S$.

Proof:

A feedback law will be chosen in such a way that the trajectories of system Σ_2 generated by that feedback law coincide with the optimal trajectories of system Σ_1. The appropriate feedback law is given by

$$\underline{u}(\underline{x}) = -\underline{B}'(t)\underline{K}(t)\underline{x}(t) - \underline{g}[t,\underline{x}(t)] \tag{27}$$

then the trajectories of systems Σ_2 generated by the feedback law (27) satisfy (26).

The performance of system Σ_2 associated with the feedback law (27) is given by

$$J_2(\underline{u}) = \int_{t_0}^{T} \{\underline{x}'(t)\underline{C}(t)\underline{C}'(t)\underline{x}(t)$$

$$+ [\underline{B}'(t)\underline{K}(t)\underline{x}(t) + \underline{g}(t,\underline{x}(t))]'[\underline{B}'(t)\underline{K}(t)\underline{x}(t) + \underline{g}(t,\underline{x}(t))]\}\,dt \tag{28}$$

$$= \int_{t_0}^{T} \{\underline{x}'(t)\underline{C}(t)\underline{C}'(t)\underline{x}(t) + \underline{x}'(t)\underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t)\underline{x}(t)\}\,dt$$

$$+ \int_{t_0}^{T} \{2\underline{x}'(t)\underline{K}(t)\underline{B}(t)\underline{g}(t,x(t)) + \underline{g}'(t,\underline{x}(t))\underline{g}(t,\underline{x}(t))\}\,dt \tag{29}$$

The first integral of the above expression is identically equal to the optimal performance of the system Σ_1 then

$$J_2^* \leq J_2(u) = J_1^* + \int_{t_0}^{T} \underline{g}'[t,\underline{x}(t)][\underline{g}(t,\underline{x}(t)) +$$

$$2\underline{B}'(t)\underline{K}(t)\underline{x}(t)]\,dt \tag{30}$$

therefore when inequality (25) is satisfied along solutions of (26) system Σ_2 is better than system Σ_1.

<u>Collorary 3.2.3</u>: An upper bound for the optimal performance of the system Σ_2 is

$$J_2^* \leq q\,\underline{x}_0'\underline{K}(t_0)\underline{x}_0$$

if

$$\int_{t_0}^{T} \underline{g}'(t, \underline{x}(t))[\underline{g}(t, x(t)) + 2\underline{B}(t)\underline{K}(t)\underline{x}(t)]\, dt$$

$$\leq (q-1)\underline{x}_0\underline{K}(t_0)\underline{x}_0 \qquad (31)$$

along solutions of (26).

3. On Universally Better Systems

It will be assumed in this section the class $\mathcal{M}$ in the definition of universally better systems is the class of cost functionals

$$J(\underline{u}, \underline{x}_0) = \int_0^{\infty} \{\underline{u}'\underline{u} + \underline{x}'\underline{C}\,\underline{C}'\underline{x}\}\, dt \qquad (1)$$

for all $\underline{C}$'s such that $[\underline{A}, \underline{C}]$ is an observable pair. The existance of universally better systems is guaranteed by the following theorem:

Theorem 3.3.1: Given any time invariant linear system Σ_1 of the form

$$\frac{d\underline{x}(t)}{dt} = \underline{A}\,\underline{x}(t) + \underline{B}\,\underline{u}(t)$$

and a cost functional (1), there always exists another linear system Σ_2 of the form $\frac{d\underline{x}(t)}{dt} = \underline{A}\,\underline{x}(t) + \underline{D}(t)\underline{x}(t) + \underline{B}\,\underline{u}$ universally better than system Σ_1

Proof:

Consider the system Σ_2 given by

$$\frac{d\underline{x}(t)}{dt} = (\underline{A} + k(t)\underline{I})\,\underline{x}(t) + \underline{B}\,\underline{u}(t), \quad \text{with } k(t) \leq \epsilon < 0$$

Since $k(t)\underline{x}'\,\underline{K}\,\underline{x} < 0$ and $\underline{K}$ is positive definite for all $J \in \mathcal{M}$ then theorem 3.3.1 follows immediately from theorem 3.2.3.

The result of the above theorem seems to contradict the result of theorem 2.5.1, however it should be kept in mind that the form of the permissible class of systems Σ_2 in theorem 2.5.1 was more restrictive than the one allowed in this chapter. In contrast with the result there it is not necessary to require that the system Σ_1 be unstable in order to find a universally better system.

4. Globally Better Systems

In this section we will focus our attention to linear systems. Let us assume that $\underline{f}(t, \underline{x}) = \underline{D}(t)\underline{x}$ then the following theorem is an immediate consequence of theorem 3.3.1.

Theorem 3.4.1: The system Σ_2 is globally better than the system Σ_1 with respect to J if $\underline{D}'(t)\underline{K}(t)$ is positive semidefinite, where $\underline{K}(t)$ is the solution of (23) with the boundary condition (24).

For the time invariant case an upper bound for the optimal cost of the system Σ_2 can be found by using the result of collorary 2.2.

Theorem 3.4.2: An upper bound for the optimal performance of the system Σ_2 is given by

$$J_2^*(\underline{u}, \underline{x}_0) \leq q\, \underline{x}_0' \underline{K} \underline{x}_0 \quad ; \quad q > 0 \tag{1}$$

if

$$(q-1)(\underline{C}\,\underline{C}' + \underline{K}\,\underline{B}\,\underline{B}'\underline{K}) - q[\underline{D}'\underline{K} + \underline{K}\,\underline{D}] = \underline{H}\,\underline{H}' \tag{2}$$

for some matrix $\underline{H}$ and if the pair $[\underline{A} + \underline{D}, \underline{C}]$ is observable.

Proof:

The proof is in two parts. First it will be shown that condition (2) implies that the null solution of

$$\frac{d\underline{x}(t)}{dt} = (\underline{A} - \underline{B}\,\underline{B}'\underline{K} + \underline{D})\underline{x}(t) \tag{3}$$

is asymptotically stable in the large.

For that purpose consider the following Lyapunov function: $v(\underline{x}) = \underline{x}'\underline{K}\,\underline{x}$ then along solutions of (3) $\frac{dv(\underline{x})}{dt}$ is given by

$$\frac{dv(\underline{x})}{dt} = \underline{x}'[\underline{A}'\underline{K} - \underline{K}\,\underline{B}\,\underline{B}'\underline{K} + \underline{D}'\underline{K} + \underline{K}\,\underline{D} + \underline{K}\,\underline{A} - \underline{K}\,\underline{B}\,\underline{B}'\underline{K}]\,\underline{x}$$

but since $\quad \underline{A}'\underline{K} + \underline{K}\,\underline{A} - \underline{K}\,\underline{B}\,\underline{B}'\underline{K} = -\underline{C}\,\underline{C}' \qquad (4)$

then

$$\frac{dv(\underline{x})}{dt} = \underline{x}'[-\underline{C}\,\underline{C}' - \underline{K}\underline{B}\underline{B}'\underline{K} + \underline{D}'\underline{K} + \underline{K}\,\underline{D}]\underline{x}$$

now, by using equation (1)

$$\frac{dv(\underline{x})}{dt} = \underline{x}'[-\frac{1}{q}\underline{H}\,\underline{H}' + (\frac{q-1}{q} - 1)(\underline{C}\,\underline{C}' + \underline{K}\,\underline{B}\,\underline{B}'\underline{K})]\,\underline{x}$$

$$= -\frac{1}{q}\underline{x}'\,[\underline{H}\,\underline{H}' + \underline{C}\,\underline{C}' + \underline{K}\,\underline{B}\,\underline{B}'\underline{K}]\,\underline{x}$$

but since q is assumed to be positive then $\frac{dv(\underline{x})}{dt} \leq 0$; for $\frac{dv(x)}{dt}$ to vanish identically it is necessary to have $\underline{B}'\underline{K}\,\underline{x}(t) \equiv 0$ which implies that (1) becomes

$$\frac{d\underline{x}}{dt} = (\underline{A} + \underline{D})\underline{x} \qquad (5)$$

but then $\underline{C}'\underline{x}(t)$ can not be identically zero, since the pair $[A+D, C]$ was assumed to be observable. Then $\frac{dv(\underline{x})}{dt} \leq 0$ and not identically zero along solutions of (1). It should be noted here that if we define

$$\underline{P}\,\underline{P}' = \underline{C}\,\underline{C}' + \underline{H}\,\underline{H}'$$

the observability assumption can be weakened by requiring only that the pair $[\underline{A} + \underline{D}, \underline{P}]$ be observable. The second part of the proof consists in showing that inequality (2.2.1) is satisfied, that is

$$\int_0^\infty e^{(\underline{A}+\underline{D}-\underline{B}\underline{B}'\underline{K})'t}(\underline{D}'\underline{K} + \underline{K}\,\underline{D})e^{(\underline{A}+\underline{D}-\underline{B}\underline{B}'\underline{K})t}dt \leq 0 \qquad (6)$$

if equation (2) holds for some $\underline{H}$.

From equations (2) and (4) it follows that

$$(q-1)[-\underline{A}'\underline{K} - \underline{K}\,\underline{A} + 2\underline{K}\,\underline{B}\,\underline{B}'\underline{K}] - q[\underline{D}'\underline{K} + \underline{K}\,\underline{D}] = \underline{H}\,\underline{H}'$$

or

$$(q-1)[-\underline{A}'\underline{K} - \underline{K}\,\underline{A} + 2\underline{K}\,\underline{B}\,\underline{B}'\underline{K} - \underline{D}'\underline{K} - \underline{K}\underline{D}] = \underline{H}\underline{H}' + \underline{D}'\underline{K} + \underline{K}\underline{D}$$

$$(q-1)[(\underline{A} - \underline{B}\underline{B}'\underline{K} + \underline{D})'\underline{K} + \underline{K}(\underline{A} - \underline{B}\underline{B}'\underline{K} + \underline{D})] = \underline{H}\,\underline{H}' + \underline{D}'\underline{K} + \underline{K}\,\underline{D}$$

Pre- and post-multiplying the above equation by $e^{(\underline{A} - \underline{B}\,\underline{B}'\underline{K} + \underline{D})'t}$ and $e^{(\underline{A} - \underline{B}\,\underline{B}'\underline{K} + \underline{D})t}$ respectively, and by integrating from 0 to ∞ it follows that

$$-(q-1)\int_0^\infty \{\frac{d}{dt}\, d^{(\underline{A} - \underline{B}\underline{B}'\underline{K} + \underline{D})'t}\, \underline{K} e^{(\underline{A} - \underline{B}\,\underline{B}'\underline{K} + D)t}\} dt$$

$$= \int_0^\infty e^{(\underline{A} - \underline{B}\underline{B}'\underline{K} + \underline{D})'t}\, \underline{H}\underline{H}'\, e^{(A - BB'K + D)t}\} dt$$

$$+ \int_0^\infty e^{(\underline{A} - \underline{B}\underline{B}'\underline{K} + \underline{D})'t} (\underline{D}'\underline{K} + \underline{K}\,\underline{D}) e^{(\underline{A} - \underline{B}\underline{B}'\underline{K} + \underline{D})t}\, dt$$

Performing the integration of the left-hand side and using the fact proved in the first part, that the eigenvalues of $(\underline{A} - \underline{B}\,\underline{B}'\underline{K} + \underline{D})$ are in $Re[s] < 0$ it is obtained that

$$(q-1)\underline{K} = \int_0^\infty e^{(\underline{A} - \underline{B}\underline{B}'\underline{K} + D)'t}\, \underline{H}\underline{H}'\, e^{(\underline{A} - \underline{B}\underline{B}'\underline{K} + D)t} dt$$

$$+ \int_0^\infty e^{(\underline{A} + \underline{D} - \underline{B}\underline{B}'\underline{K})'t} [\underline{D}'\underline{K} + \underline{K}\,\underline{D}]\, e^{(\underline{A} + \underline{D} - \underline{B}\,B'\underline{K})t}\, dt$$

which implies that inequality (6) is satisfied, and by collorary 3.2.2 inequality (1) holds. ||

A very similar result to the above theorem was presented in Ref. M-1.

CHAPTER IV

ON LOWER BOUNDS ON THE PERFORMANCE OF THE REGULATOR PROBLEM

In this chapter the problem of determining a priori lower bounds on the optimal performance of linear systems is analyzed. Then the problem of evaluating the degree of suboptimality of a given design is studied. Two examples illustrating the advantages of the obtained results are given.

1. Notation and Preliminaries

Consider the linear time varying system described by the equation

$$\Sigma_1 : \frac{d\underline{x}(t)}{dt} = \underline{A}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) \qquad (1)$$

$$\underline{y}(t) = \underline{C}'(t)\underline{x}(t) \quad ; \quad \underline{x}(t_0) = \underline{x}_0 \qquad (2)$$

and the cost functional

$$J(\underline{u}, \underline{x}_0) = \int_{t_0}^{T} [\underline{y}'(t)\underline{y}(t) + \underline{u}'(t)\underline{u}(t)]\,dt \quad ; \quad \infty > T > t_0 \qquad (3)$$

also denote by

$$\|\underline{u}\|^2 \triangleq \int_{t_0}^{T} \underline{u}'(t)\underline{u}(t)dt \qquad (4)$$

$$\|\underline{y}\|^2 \triangleq \int_{0}^{T} \underline{y}'(t)\underline{y}(t)dt \qquad (5)$$

and the operator G mapping $R^r x[t_0, T]$ into $R^m x[t_0, T]$ as

$$G\underline{u} \triangleq \int_{t_0}^{t} \underline{C}'(t)\underline{\Phi}_A(t,\sigma)\underline{B}(\sigma)\underline{u}(\sigma)d\sigma \quad ; \quad t_0 \leq t \leq T \tag{6}$$

where $\underline{\Phi}_A(t,t_0)$ is the transition matrix associated with the matrix $\underline{A}(t)$ that is $\underline{\Phi}_A$ should satisfy the following two conditions

$$\text{i)} \quad \underline{\Phi}_A(t_0,t_0) = \underline{I}$$

$$\text{ii)} \quad \frac{d\underline{\Phi}_A}{dt}(t,t_0) = \underline{A}(t)\underline{\Phi}_A(t,t_0) \tag{8}$$

The gain of the operator G is defined as

$$g = ||G|| = \sup_{\underline{u}\in\mathcal{L}_2[t_0,T]} \frac{||G\underline{u}||}{||\underline{u}||} \tag{9}$$

The inner product of two vectors functions of time, $\underline{\Psi}(t)$ and $\underline{\Gamma}(t)$ is defined as

$$< \underline{\Psi}(t), \underline{\Gamma}(t) > = \int_{t_0}^{T} \underline{\Psi}'(t)\underline{\Gamma}(t)dt \tag{10}$$

$\underline{y}_0$ will denote the homogeneous output of system Σ_1, that is

$$\underline{y}_0(t) = \underline{C}'(t)\underline{\Phi}_A(t,t_0)\underline{x}_0 \tag{11}$$

The performance associated with the control $\underline{u}(t) \equiv \underline{0}$ will be called the homogeneous performance and is denoted by

$$J_0 \triangleq ||\underline{y}_0||^2 \tag{12}$$

Note that for any $\underline{u}(t)$ the performance associated with that control, by virtue of linearity, can be expressed alternatively as

$$J(u,\underline{x}_0) = ||\underline{u}||^2 + ||G\underline{u} + \underline{y}_0||^2 \tag{13}$$

2. A Fundamental Inequality

In this section we will derive an inequality that is essential for the development of latter sections. First we will derive a theorem giving some characterization of optimal feedback controls.

Lemma 4.2.1: (A characterization of optimal controls). If $\underline{u}^*$ is the optimal control for the system Σ_1 and $\underline{y}^*$ is the corresponding output then

$$J^* = \langle \underline{y}^*, y_0 \rangle = \int_{t_0}^{T} \underline{y}^{*\prime}(t)\underline{y}_0(t)dt \tag{1}$$

Proof:

It has been determined[K-3] that the optimal trajectories generated by the optimal control satisfy the equation

$$\frac{d\underline{x}(t)}{dt} = [\underline{A}(t) - \underline{B}(t)\underline{B}'(t)\underline{K}(t)]\underline{x}(t) \tag{2}$$

where $\underline{K}(t)$ is the unique (positive semidefinite) solution of the Matrix Riccati equation

$$\underline{A}'(t)\underline{K}(t) + \underline{K}(t)\underline{A}(t) - \underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t) + \frac{d\underline{K}(t)}{dt} = -\underline{C}(t)\underline{C}'(t) \tag{3}$$

and $\underline{K}(T) \equiv \underline{0}$. In addition, the optimal cost is given by

$$J^* = \underline{x}_0' \underline{K}(t_0)\underline{x}_0 \tag{4}$$

Denote by $\underline{\Phi}_K(t, t_0)$ the transition matrix associated with the matrix $\underline{A}(t) - \underline{B}(t)\underline{B}'(t)\underline{K}(t)$ then, after pre- and post-multiplying (3) by $\underline{x}_0' \underline{\Phi}_A'(t, t_0)$ and $\underline{\Phi}_K(t, t_0)\underline{x}_0$ respectively, one obtains that

$$\underline{x}'_0 \underline{\Phi}'_A(t,t_0)\underline{A}'(t)\underline{K}(t)\underline{\Phi}_K(t,t_0)\underline{x}_0$$

$$+ \underline{x}'_0\underline{\Phi}'_A(t,t_0)\underline{K}(t)[\underline{A}(t) - \underline{B}(t)\underline{B}'(t)\underline{K}(t)]\underline{\Phi}_K(t,t_0)\underline{x}_0 \tag{5}$$

$$+ \underline{x}'_0 \underline{\Phi}'_A(t,t_0)\frac{dK(t)}{dt}\underline{\Phi}_K(t,t_0)\underline{x}_0 = -\underline{x}'_0 \underline{\Phi}_A(t,t_0)\underline{C}(t)\underline{C}'(t)\underline{\Phi}_K(t,t_0)\underline{x}_0$$

However, the left-hand side of equation (5) is a perfect differential, then equation (5) can be written as

$$\frac{d}{dt}\{\underline{\phi}'_A(t,\underline{x}_0,t_0)\underline{K}(t)\underline{\phi}_K(t,\underline{x}_0,t_0)\} = -\underline{\phi}'_A(t,\underline{x}_0,t_0)\underline{C}(t)\underline{C}'(t)\underline{\phi}_K(t,\underline{x}_0,t_0) \tag{6}$$

where we have used the following definitions:

$$\underline{\phi}_A(t,\underline{x}_0,t_0) \triangleq \underline{\Phi}_A(t,t_0)\underline{x}_0 \tag{7}$$

$$\underline{\phi}_K(t,\underline{x}_0,t_0) \triangleq \underline{\Phi}_K(t,t_0)\underline{x}_0 \tag{8}$$

Integrating both sides of (6) from t_0 to T it is obtained that

$$-\underline{x}'_0\underline{K}(t_0)\underline{x}_0 + \underline{\phi}_A(T,\underline{x}_0,t_0)\underline{K}(T)\underline{\phi}_K(T,\underline{x}_0,t_0) = -\langle y_0, y^* \rangle \tag{9}$$

since $\qquad \underline{y}_0(t) = \underline{C}'(t)\,\underline{\phi}_A(t,\underline{x}_0,t_0)$

and $\qquad \underline{y}^*(t) = \underline{C}'(t)\,\underline{\phi}_K(t,\underline{x}_0,t_0)$

From (4), (6) and the boundary condition on $\underline{K}(t)$ it follows that

$$J^* = \langle y_0, y^* \rangle \qquad ||$$

<u>Collorary 4.2.1</u>: If J_0 and J^* are the optimal performance and the homogeneous performance of the system Σ_1 respectively, and if $\underline{u}^*$ is the optimal control then

a) $$J^* = J_0 - ||Gu^*||^2 - ||u^*||^2 \tag{10}$$

b) $$J^* = J_0 + \langle Gu^*, y_0 \rangle \tag{11}$$

Proof:

Because of linearity of the system, by relation (1.13) and equation (1) it is obtained that

$$\|\underline{u}^*\|^2 + \|G\underline{u}^*\|^2 + 2\langle G\underline{u}^*, \underline{y}_0\rangle + \|\underline{y}_0\|^2 = \|\underline{y}_0\|^2 + \langle G\underline{u}^*, \underline{y}_0\rangle \tag{12}$$

or
$$\|\underline{u}^*\|^2 + \|G\underline{u}^*\|^2 + \langle G\underline{u}^*, \underline{y}_0\rangle = 0 \tag{13}$$

then, by using equation (13) and equation (1.13) both equations (10) and (11) follow immediately. | |

Another interesting, nontrivial by-product of lemma 4.2.1 is the following

Collorary 4.2.2: An upper bound for the optimal energy is

$$\|\underline{u}^*\|^2 \leq \frac{J_0}{4} \tag{14}$$

and equality holds if and only if $\underline{y}^* = 2\underline{y}_0$.

Proof:

From equation (1) it follows that

$$\|\underline{u}^*\|^2 + \|\underline{y}^*\|^2 = \langle \underline{y}^*, \underline{y}_0\rangle \tag{15}$$

then, by completing the square, it is obtained that

$$\|\underline{u}^*\|^2 + \|\underline{y}^*\|^2 - \langle \underline{y}^*, y_0\rangle + \|\underline{y}_0\|^2/4 = \|\underline{y}_0\|^2/4 \tag{16}$$

therefore
$$\|\underline{u}^*\|^2 + \|\underline{y}_0/2 - \underline{y}^*\|^2 = J_0/4 \tag{17}$$

from which the statement of the collorary follows immediately.

Theorem 4.2.1:† The ratio of the optimal performance to the homogeneous performance of system Σ_1 obeys the inequality

† The original statement and proof of this particular theorem is due to R.W. Brockett. His proof however is different to the one presented here.

$$\frac{J^*}{J_0} \geq \frac{1}{1+g^2} \tag{18}$$

where $g = \| G \|$.[B-4]

Proof:

By using equation (1) and the fact that $\| G\underline{u} \| \leq \| G \| \; \| u \|$ it follows that

$$J^* \leq J_0 - \| G\underline{u}^* \|^2 (1 + \frac{1}{g^2}) \tag{19}$$

Similarly by using equation (11) and the Schwarz inequality, (i.e.,

$$< G\underline{u}^*, y_0 > \geq - \| G\underline{u}^* \| \; \| y_0 \| \quad) \tag{20}$$

then $$J^* \geq J_0 - \| G\underline{u}^* \| \; \| \underline{y}_0 \| \tag{21}$$

therefore $$\| G\underline{u}^* \| \geq \frac{J_0 - J^*}{\| \underline{y}_0 \|}$$

but since both sides of the inequality are nonnegative, then

$$\| G\underline{u}^* \|^2 \geq \frac{(J_0 - J^*)^2}{J_0} \tag{22}$$

By making the definition

$$R \triangleq \frac{J^*}{J_0}$$

it follows from (22) that

$$\frac{\| G\underline{u}^* \|^2}{J_0} \geq (1-R)^2 \tag{23}$$

by dividing both sides of inequality (19) by J_0 it is obtained that

$$R \leq 1 - \frac{\| G\underline{u}^* \|^2}{J_0} (1 + \frac{1}{g^2}) \tag{24}$$

By substituting inequality (23) into inequality (24) it follows that

$$R \leq 1 - (1-R)^2 (1 + \frac{1}{g^2})$$

or

$$(1-R) \geq (1-R)^2\left(1 + \frac{1}{g^2}\right) \tag{25}$$

Inequality (25) is satisfied (see Fig. 6) if and only if R satisfies

$$1 \geq R \geq \frac{1}{1+g^2}$$

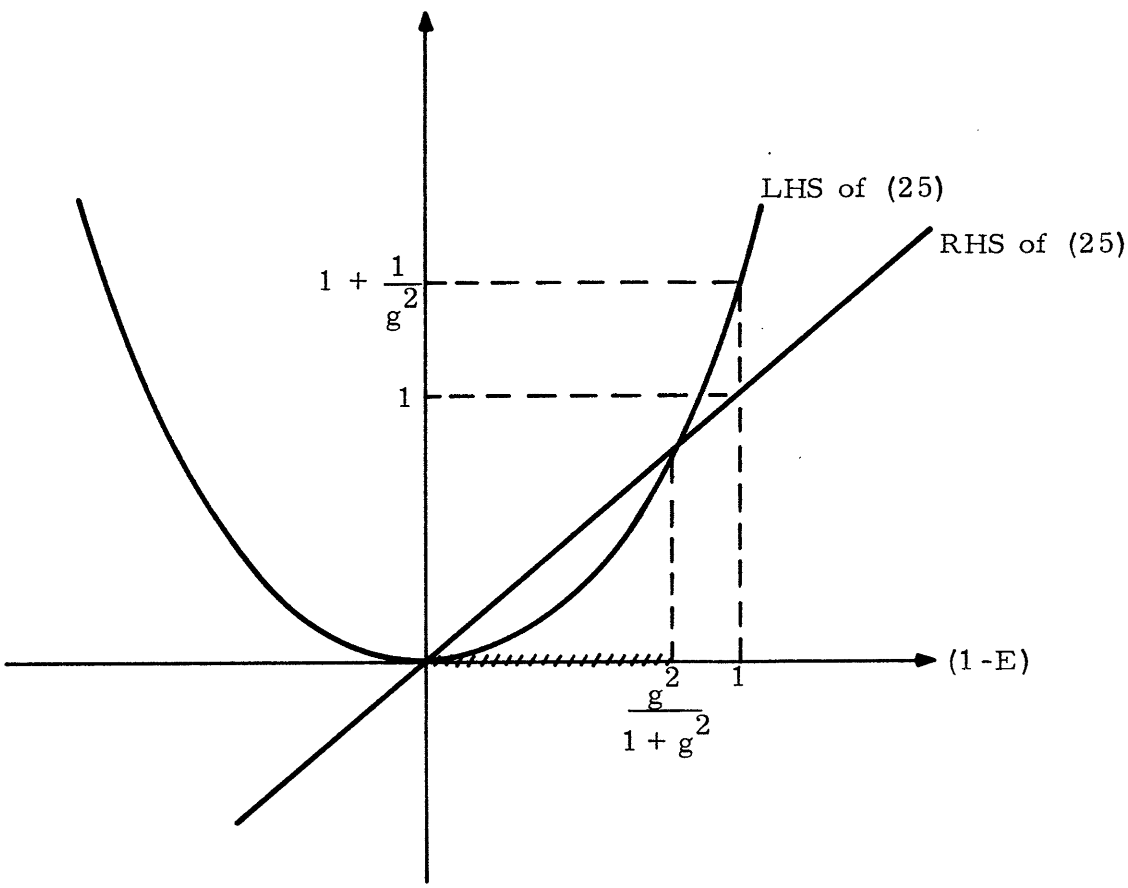

Fig. 6 Region for which inequality (25) is satisfied.

therefore

$$1 \geq \frac{J^*}{J_0} \geq \frac{1}{1+g^2} \qquad ||$$

<u>Comment:</u>

The proof of the theorem above displays explicitly the two approximations made. For completeness we repeat them here:

a) $\| G\underline{u}^* \|^2 \leq \| G \| \;\; \| \underline{u}^* \|$ (26)

b) $< G\underline{u}^*, \underline{y}_0 > \geq - \| G\underline{u}^* \| \;\; \| \underline{y}_0 \|$ (27)

The difference between the right- and left-hand sides of inequality (26) becomes arbitrarily small for some u*. This fact follows immediately from the definition of the gain of the operator G. Inequality (27), on the other hand, becomes equality whenever $G\underline{u}^* = +c\underline{y}_0$ where c is a negative constant. Therefore it can be said that inequality (18) is the best of its kind, that is, given that only known parameter of the system Σ_1 is the gain of the operator G then there does not exist a function $f(\| G \|)$ such that

$$\frac{J^*}{J_0} \geq f(\| G \|) > \frac{1}{1 + \| G \|^2} \tag{28}$$

One of the natural questions that arise at this point is the following: Given that the only known parameter of the system is the gain of the operator G it is possible to find an upper bound strictly less than unity for the ratio $\frac{J^*}{J_0}$? Posing the problem in a more precise manner, does there exist a function $f(\| G \|)$ such that

$$1 > f(\| G \|) \geq \frac{J^*}{J_0} \; ?$$

The answer to this question is negative as shown in the following theorem.

Theorem 4.2.2: Given any positive real number g then there exist a linear time invariant system with a transfer function G^{B-3} such that

$$g = \| G \|$$

and a set of initial conditions such that

$$J^* = J_0 \tag{29}$$

Proof:

Assume that

$$G(s) = \frac{\alpha q(s)}{p(s)}$$

and that $p(s)$ is a Hurwitz polynomial. Without loss of generality it can be assumed that $p(s)$ is chosen with all its roots in the real axis and α will be adjusted accordingly in order to make $\| G \| = g$.

As it was shown earlier[B-1] the optimal performance of the system

$$p(D)x(t) = u(t)$$

$$\alpha q(D)x(t) = y(t)$$

for the cost functional

$$J(u) = \int_0^\infty \{u^2(t) + y^2(t)\}\,dt$$

is given by

$$J^* = \int_0^\infty \{[p(D)x(t)]^2 + [\alpha q(D)x(t)]^2 - [(p\bar{p} + \alpha^2 q\bar{q})^+(D)x(t)]^2\}\,dt \tag{30}$$

where $x(t)$ satisfies some prescribed initial conditions.

Let us assume that s_0 is a root of $p(s)$. $q(s)$ is chosen such that

$$q(-s_0) = 0 \quad \text{but} \quad q(s_0) \neq 0 \tag{31}$$

therefore

$$(p\bar{p} + \alpha^2 q\bar{q})(s_0) = 0 \tag{32}$$

Since the integral of the right-hand side of (30) is independent of path, it can be evaluated along solutions of the equation

$$p(D)x(t) = 0 \tag{33}$$

then

$$J^* = \int_0^\infty \{[\alpha q(D)x(t)]^2 - [(p\overline{p} + \alpha^2 q\overline{q})^+(D)x(t)]^2\}\,dt \tag{34}$$

whenever $x(t)$ satisfies (33).

If the initial conditions are chosen in such a way that

$$x(t) = e^{s_0 t} \tag{36}$$

then by using (32) it follows that

$$J^* = \int_0^\infty [\alpha q(s_0)]^2 e^{2s_0 t}\,dt$$

but on the other hand

$$J_0 = \int_0^\infty [\alpha q(s_0)]^2 e^{2s_0 t}\,dt$$

therefore $J^* = J_0$. $||$

Now based on theorem 4.2.1 the fundamental inequality of the section will be derived.

Consider the system Σ_2 described by the equations

$$\dot{\underline{x}}(t) = \underline{F}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) \tag{36}$$

$$\underline{y}(t) = \underline{D}'(t)\underline{x}(t) \tag{37}$$

and in a similar manner as in Section 1, define the operator H by the relation

$$H\underline{u} = \int_{t_0}^{t} \underline{\Phi}_F(t,\sigma)\underline{B}(\sigma)\underline{u}(\sigma)\,d\sigma \qquad t_0 \le t \le T \tag{38}$$

where $\underline{\Phi}_F$ is the transition matrix associated with the matrix $\underline{F}(t)$.

<u>Theorem 4.2.3:</u> The inequality

$$\underline{x}_0' \int_{t_0}^{T} \underline{\Phi}_F'(t,t_0)\underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t)\underline{\Phi}_F(t,t_0)dt\,\underline{x}_0$$

$$\leq \left(\frac{\|H\|^2}{1+\|H\|^2}\right) \underline{x}_0' \int_{t_0}^{T} \underline{\Phi}_F'(t,t_0)\underline{D}(t)\underline{D}'(t)\underline{\Phi}_F(t,t_0)dt\,\underline{x}_0 \tag{39}$$

holds for all $\underline{x}_0$ where $\underline{K}(t)$ is the unique (positive semi-definite) solution of the differential equation

$$\frac{d\underline{K}(t)}{dt} + \underline{F}'(t)\underline{K}(t) + \underline{K}(t)\underline{F}(t) - \underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t) = -\underline{D}(t)\underline{D}'(t) \tag{40}$$

with $$\underline{K}(T) = \underline{0} \tag{41}$$

<u>Proof:</u>

Pre- and post-multiplying equation (40)by $\underline{x}_0'\underline{\Phi}_F'(t,t_0)$ and $\underline{\Phi}_F(t,t_0)\underline{x}_0$ respectively, and by noticing that the left-hand side becomes a perfect differential, it follows that

$$\frac{d}{dt}\{\underline{x}_0'\underline{\Phi}_F'(t,t_0)\underline{K}(t)\underline{\Phi}_F(t,t_0)\underline{x}_0\}$$

$$+\underline{x}_0'\underline{\Phi}_F'(t,t_0)\underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t)\underline{\Phi}_F(t,t_0)\underline{x}_0 = -\underline{x}_0'\underline{\Phi}_F'(t,t_0)\underline{D}(t)\underline{D}'(t)\underline{\Phi}_F(t,t_0)\underline{x}_0 \tag{42}$$

by integrating equation (42) from t_0 to T and use of the boundary condition (41) it follows that

$$\underline{x}_0'\underline{K}(t_0)\underline{x}_0 + \int_{t_0}^{T} \underline{x}_0'\underline{\Phi}_F'(t,t_0)\underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t)\underline{\Phi}_F(t,t_0)\underline{x}_0\,dt$$

$$= \int_{t_0}^{T} \underline{x}_0'\underline{\Phi}_F'(t,t_0)\underline{D}(t)\underline{D}'(t)\underline{\Phi}_F(t,t_0)\underline{x}_0\,dt \tag{43}$$

However, $\underline{x}_0'\underline{K}(t_0)\underline{x}_0$ is the optimal cost of the system Σ_2 with respect to the cost functional

$$J(\underline{u};\underline{x}_0) = \int_{t_0}^{T} \{\underline{u}'(t)\underline{u}(t) + \underline{y}'(t)\underline{y}(t)\}\,dt$$

and $\int_{t_0}^{T} \underline{x}_0'\,\underline{\Phi}_F'(t,t_0)\underline{D}(t)\underline{D}'(t)\underline{\Phi}_F(t,t_0)\underline{x}_0\,dt$ is the homogeneous performance of the system Σ_2, then

$$J^* + \int_{t_0}^{T} \underline{x}_0'\,\underline{\Phi}_F'(t,t_0)\underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t)\underline{\Phi}_F(t,t_0)\underline{x}_0\,dt = J_0 \qquad (44)$$

but from theorem 4.2.1

$$J^* \geq \frac{1}{1+\|H\|^2}\,J_0 = J_0 - \frac{\|H\|^2}{1+\|H\|^2}\,J_0 \qquad (45)$$

then from equation (44) and inequality (45) the statement of the theorem follows immediately. | |

3. A Measure of the Degree of Sub-Optimality

In theorem 4.2.1 a lower bound for the ratio of the optimal to the homogeneous performance of a system has been given as a function of the gain the operator G. In this section a lower bound for the ratio of the optimal to some sub-optimal design will be given. We will restrict our study to those sub-optimal feedback laws given by

$$\underline{u}(t) = -\underline{B}(t)\underline{L}(t)\underline{x}(t) \qquad (1)$$

where $\underline{L}(t)$ is a positive semidefinite matrix that satisfies a differential equation of the form

$$\frac{d\underline{L}(t)}{dt} + \underline{A}'(t)\underline{L}(t) + \underline{L}(t)\underline{A}(t) - \underline{L}(t)\underline{B}(t)\underline{B}'(t)\underline{L}(t) + \underline{C}'(t)C(t) = \underline{D}'(t)\underline{D}(t) \qquad (2)$$

and $\underline{L}(T) = 0$.

Theorem 4.3.1: Given a suboptimal feedback law (1) where $\underline{L}(t)$ satisfies equation (2) then

$$\frac{J^*}{J_L} \geq \frac{1}{1+g^2} + \frac{1}{1+g^2} \frac{\underline{x}_0 \underline{L}(t_0)\underline{x}_0}{J_L} \tag{3}$$

where J_L is the performance associated with the feedback law (1) and g is the gain of the operator defined by

$$G\underline{w} = \int_{t_0}^{t} \underline{D}'(t)\, \underline{\Phi}_L(t,\sigma)\underline{B}(\sigma)\underline{w}(\sigma)d\sigma \tag{4}$$

and $\underline{\Phi}_L$ is the transition matrix associated with the matrix

$$\underline{A}(t) - \underline{B}(t)\underline{B}'(t)\underline{L}(t)$$

Proof:

By adding to both sides of equation (2.3) the matrix

$$-\underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{L}(t) - \underline{L}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t)$$

it is obtained that

$$\frac{d\underline{K}(t)}{dt} + \underline{K}(t)[\underline{A}(t)-\underline{B}(t)\underline{B}'(t)\underline{L}(t)] + [\underline{A}(t)-\underline{B}(t)\underline{B}'(t)\underline{L}(t)]'\underline{K}(t)$$

$$- \underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t)$$

$$= -\underline{C}(t)\underline{C}'(t) - \underline{K}(t)\underline{B}(t)\underline{B}'(t)\underline{L}(t) - \underline{L}(t)\underline{B}(t)\underline{B}'(t)\underline{K}(t) \tag{5}$$

Subtracting from both sides of equation (5) the matrix $\underline{L}(t)\underline{B}(t)\underline{B}'(t)\underline{L}(t)$ and rearranging terms it follows that

$$\frac{d\underline{K}(t)}{dt} + \underline{K}(t)[\underline{A}(t)-\underline{B}(t)\underline{B}'(t)\underline{L}(t)] + [\underline{A}(t)-\underline{B}(t)\underline{B}'(t)L(t)]'K(t)$$

$$-[\underline{K}(t)-\underline{L}(t)]\underline{B}(t)\underline{B}'(t)[\underline{K}(t)-\underline{L}(t)] = -\underline{C}(t)\underline{C}'(t) - \underline{L}(t)\underline{B}(t)\underline{B}'(t)\underline{L}(t) \tag{6}$$

Now by pre- and post-multiplying the above equation by $\underline{x}_0'\underline{\Phi}_L'(t,t_0)$ and $\underline{\Phi}_L(t,t_0)\underline{x}_0$ respectively and integrating from t_0 to T, the following relation is obtained:

$$\underline{x}_0'\underline{K}(t_0)\underline{x}_0 + \underline{x}_0'\int_{t_0}^{T} \underline{\Phi}_L'(t,t_0)[\underline{K}(t)-\underline{L}(t)]\underline{B}(t)\underline{B}'(t)[\underline{K}(t)-\underline{L}(t)]\underline{\Phi}_L(t,t_0)\underline{x}_0 dt = J_L \tag{7}$$

therefore

$$J^* + \underline{x}_0'\int_{t_0}^{T} \underline{\Phi}_L'(t,t_0)[\underline{K}(t)-\underline{L}(t)]\underline{B}(t)\underline{B}'(t)[\underline{K}(t)-\underline{L}(t)]\underline{\Phi}_L(t,t_0)\underline{x}_0 dt \tag{8}$$

$$= J_L$$

Since the matrix $\underline{M}(t) = [\underline{K}(t)-\underline{L}(t)]$ satisfies the differential equation†

$$\frac{d\underline{M}(t)}{dt} + \underline{F}'(t)\underline{M}(t) + \underline{M}(t)\underline{F}(t) - \underline{M}(t)\underline{B}(t)\underline{B}'(t)\underline{M}(t) = -\underline{D}(t)\underline{D}'(t)$$

where $\quad \underline{F}(t) = [\underline{A}(t) - \underline{B}(t)\underline{B}'(t)\underline{L}(t)]$ (9)

then by using theorem 4.2.3 and equation (8) it follows that

$$J_L - J^* \le \frac{\|G\|^2}{1+\|G\|^2}\left(\int_{t_0}^{T} \underline{x}_0'\,\underline{\Phi}_L'(t,t_0)\underline{D}'(t)\underline{D}(t)\underline{\Phi}_L(t,t_0)\underline{x}_0 dt\right) \tag{10}$$

Now the left-hand side of inequality (10) can again be expressed in terms of J_L and $\underline{x}_0'\underline{L}(t_0)\underline{x}_0$. In order to do this consider equation (2), by adding $\underline{L}(t)\underline{B}(t)\underline{B}'(t)\underline{L}(t)$ to both sides of the equation it is found that

$$\frac{d\underline{L}(t)}{dt} + [\underline{A}(t)-\underline{B}(t)\underline{B}'(t)\underline{L}(t)]'\underline{L}(t) + \underline{L}(t)[\underline{A}(t)-\underline{B}(t)\underline{B}'(t)\underline{L}(t)]$$

......Continued on next page

† To verify the above statement simply substitute $\underline{M}(t)$ by $[\underline{K}(t)-\underline{L}(t)]$ and use the differential equations that $\underline{K}(t)$ and $\underline{L}(t)$ satisfy.

$$= -\underline{C}(t)\underline{C}'(t) + \underline{D}(t)\underline{D}'(t) - \underline{L}(t)\underline{B}(t)\underline{B}'(t)\underline{L}(t) \qquad (11)$$

After pre- and post-multiplying the terms of equation (11) by $\underline{x}_0'\underline{\Phi}'_L(t,t_0)$ and $\underline{\Phi}_L(t,t_0)\underline{x}_0$ respectively and by integrating every term, it follows that

$$\underline{x}_0\underline{L}\,\underline{x}_0 = J_L - \underline{x}_0'\int_{t_0}^{T} \underline{\Phi}'_L(t,t_0)\underline{D}(t)\underline{D}'(t)\underline{\Phi}_L(t,t_0)\underline{x}_0\,dt \qquad (12)$$

By using equation (12) in conjunction with inequality (10) the following relation is obtained

$$J_L - J^* \leq \frac{\| G \|^2}{1+\| G \|^2}\,(J_L - \underline{x}_0'\underline{L}(t_0)\underline{x}_0) \qquad (13)$$

by dividing both sides of (13) by J_L then

$$1 - \frac{J^*}{J_L} \leq \frac{\| G \|^2}{1+\| G \|^2}\left(1 - \frac{\underline{x}_0'\underline{L}(t_0)\underline{x}_0}{J_L}\right)$$

then

$$\frac{J^*}{J_L} \geq \frac{1}{1+\| G \|^2} + \frac{\| G \|^2}{1+\| G \|^2}\left(\frac{\underline{x}_0'\underline{L}(t_0)\underline{x}_0}{J_L}\right) \qquad ||$$

Collorary 4.3.1

$$\frac{J^*}{J_L} \geq \frac{1}{1+\| G \|^2}$$

The importance of both collorary 4.3.1 and theorem 4.3.1 are obvious for the following reasons:

a) The lower bound of the ratio J^*/J_L is given entirely as a function of L,

b) The bound given in theorem 4.3.1 is sharp, that is, there exist initial conditions for which equality is attained.

c) The bound given in collorary 4.3.1 is independent of the initial conditions, and last

d) It is possible to give a quantitative measure of the degree of suboptimality of a large class of designs by evaluation of the lower bound on $\frac{J^*}{J_L}$.

It has been assumed that the optimization interval (i.e., $[t_0, T]$) is finite, and results given so far are based very strongly on the finiteness of T.

As has been mentioned before, time invariant systems are of great importance in conjunction with the optimization interval $[0, \infty]$. For this case the optimal feedback control is linear and time invariant and therefore very attractive from the implementation point of view.

It would be expected that the results given in theorem 4.3.1 generalizes for the case in which the system Σ_1 is time invariant and $T \to \infty$. However, some problems of existence and uniqueness are introduced that should be taken into account, for example, existence of integrals of the form $\underline{x}_0' \int_0^\infty e^{\underline{A}'t} \underline{R}\,\underline{R}' e^{\underline{A}t} x_0 \, dt$ is not guaranteed, unless specific assumptions on the location of the eigenvalues of the matrix $\underline{A}$ are made. Also uniques of positive semidefinite solutions of the algebraic Riccati equation is not guaranteed. In order to overcome these difficulties several assumptions will be made and some introductory results will be proved before stating the version of theorem 4.3.1 for the case in which $T \to \infty$.

The system Σ_1 under consideration is the system described by the equations

$$\Sigma_1 : \frac{d\underline{x}(t)}{dt} = \underline{A}\ \underline{x}(t) + \underline{B}\, u(t) \tag{14}$$

$$\underline{y}(t) = \underline{C}'\ \underline{x}(t) \quad ; \quad \underline{x}(0) = \underline{x}_0 \tag{15}$$

with the qualification that the system Σ_1 is both completely controllable and completely observable. The cost functional is given by

$$J(\underline{u}, x_0) = \int_0^\infty \{\underline{u}'(t)\underline{u}(t) + \underline{y}'(t)\underline{y}(t)\}\, dt \tag{16}$$

The matrix $\underline{L}$ is assumed to be <u>positive semidefinite</u> and there exists a matrix $\underline{D}$ such that

$$\underline{A}'\underline{L} + \underline{L}\underline{A} - \underline{L}\underline{B}\underline{B}'\underline{L} + \underline{C}'\underline{C} = \underline{D}'\underline{D} \tag{17}$$

It will be assumed that the suboptimal feedback law $\underline{u} = -\underline{B}'\underline{L}\underline{x}$ is such that the eigenvalues of $(\underline{A} - \underline{B}\underline{B}'\underline{L})$ lie entirely in $\mathrm{Re}[s] < 0$.

The following theorem establishes the positive semidefiniteness of $\underline{K} - \underline{L}$.

<u>Theorem 4.3.2:</u> If the matrix $\underline{L}$ satisfies equation (17) then $\underline{K}-\underline{L}$ is positive semidefinite where $\underline{K}$ is the unique positive definite solution of the algebraic Riccati equation

$$\underline{A}'\underline{K} + \underline{K}\underline{A} - \underline{K}\underline{B}\underline{B}'\underline{K} = -\underline{C}\underline{C}' \tag{18}$$

<u>Proof:</u>

After some algebraic manipulation from equations (17) and (18) the following relation is obtained:

$$[\underline{A} - \underline{B}\underline{B}'\underline{K}]'[\underline{K} - \underline{L}] + [\underline{K} - \underline{L}][\underline{A} - \underline{B}\underline{B}'\underline{K}]$$

$$= -\underline{D}'\underline{D} - [\underline{K} - \underline{L}]\underline{B}\underline{B}'[\underline{K} - \underline{L}] \tag{19}$$

and since the eigenvalues of $[\underline{A} - \underline{B}\underline{B}'\underline{K}]$ are in $\mathrm{Re}[s] < 0$ it follows that $\underline{K} - \underline{L}$ is positive semidefinite.

The following theorem guarantees the existence of a unique positive semidefinite solution of the algebraic Riccati equation.

Theorem 4.3.3: If the matrix $\underline{A}$ is stable (all the eigenvalues are in $Re(s) < 0$) and the pair $[\underline{A}, \underline{B}]$ is controllable, then there exists a unique positive semidefinite solution to the algebraic equation

$$\underline{A}\,\underline{K} + \underline{K}\,\underline{A} - \underline{K}\,\underline{B}\,\underline{B}'\,\underline{K} = -\underline{D}'\underline{D} \tag{20}$$

(Note here that the usual observability assumption on the pair $[\underline{A}, \underline{D}]$ is absent.)

Proof:

First we will show that if two solutions of (20), say $\underline{K}_1$ and $\underline{K}_2$ are such that the matrices are $[\underline{A} - \underline{B}\,\underline{B}'\underline{K}_1]$ and $[\underline{A} - \underline{B}\,\underline{B}'\underline{K}_2]$ are stable then $\underline{K}_1$ should be identical to $\underline{K}_2$. So if $\underline{K}_1$ and $\underline{K}_2$ are solutions of (20) then

$$\underline{A}\,\underline{K}_1 + \underline{K}_1\underline{A} - \underline{K}_1\underline{B}\,\underline{B}'\underline{K}_1 = -\underline{D}\,\underline{D}'$$

$$\underline{A}'\underline{K}_2 + \underline{K}_2\underline{A} - \underline{K}_2\underline{B}\,\underline{B}'\underline{K}_2 = -\underline{D}\,\underline{D}'$$

by subtracting the two equations and rearranging terms it is obtained that

$$[\underline{A} - \underline{B}\,\underline{B}'\underline{K}_2][\underline{K}_1 - \underline{K}_2]' + [\underline{K}_1 - \underline{K}_2][\underline{A} - \underline{B}\,\underline{B}'K_1] = \underline{0}$$

However the equation $\underline{A}\,\underline{X} + \underline{X}'\underline{B} = \underline{0}$ has a unique solution $\underline{X} = \underline{0}$ if $\underline{A}$ and $-\underline{B}'$ do not have common eigenvalues,[G-2] then it can be concluded that

$$\underline{K}_1 = \underline{K}_2$$

By using the result obtained by Wonham (Theorem 4.1 of Ref. W-2) that if any positive semidefinite solution of (20), say $\underline{P}$, with $\underline{A}$ stable and $[\underline{A}, \underline{B}]$ controllable, has the property that the matrix $\underline{A}$-$\underline{B}\underline{B}'\underline{P}$ is stable. Then by the first part of the theorem it follows that equation (20) has a unique positive semidefinite solution. | |

For the system under consideration with the assumption that the matrix $\underline{A}$ is stable, both lemma 4.2.2 and theorem 4.2.2 still hold. The proofs for the time invariant case and $T \to \infty$ are identical to the ones presented in the previous section, and will be omitted. By assuming stability of $\underline{A}$ the existence of the terms in equations (2.9) and (2.22) are assured. Theorem 2.3 also holds for the time invariant case under the assumption that the matrix $\underline{F}$ is stable and the pair $[\underline{F}, \underline{B}]$ is controllable. With the above observation we can state the following theorem:

Theorem 4.3.4: Given a suboptimal control law

$$\underline{u}(t) = -\underline{B}'\underline{L}\underline{x}(t) \tag{21}$$

with $\underline{L}$ positive semidefinite, satisfying the algebraic equation (17) and $(\underline{A} - \underline{B}\,\underline{B}'\underline{L})$ stable, then

$$\frac{J^*}{J_L} \geq \frac{1}{1+g^2} + \frac{g^2}{1+g^2}\;\frac{\underline{x}_0'\underline{L}\,\underline{x}_0}{J_L} \tag{22}$$

where g is the gain of the operator†

$$\underline{\mathbf{G}}\,\underline{w} = \int_0^t \underline{D}'\, e^{(\underline{A}-\underline{B}\underline{B}'\underline{L})(t-\sigma)}\underline{B}\,\underline{w}(\sigma)d\sigma \tag{23}$$

† It is not difficult to verify that the gain of the operator G is given by

$$g^2 = \|G\|^2 = \sup_{\omega} \sup_{\underline{v}} \frac{\tilde{\underline{v}}'\underline{D}'(-Ij\omega-\underline{A}+BB'\underline{L})^{-1}\underline{B}\underline{B}'(+Ij\omega-\underline{A}+\underline{B}\underline{B}'\underline{L})^{-1}\underline{D}\underline{v}}{\tilde{\underline{v}}'\underline{v}}$$

where $\tilde{\underline{v}}$ is the complex conjugate of $\underline{v}$.

Proof:

Proceed as in theorem 4.3.1. The equation (27) becomes

$$\underline{x}_0'\underline{K}\underline{x}_0 + \underline{x}_0'\int_0^\infty e^{(\underline{A}-\underline{B}\underline{B}'\underline{L})'t}[\underline{K}-\underline{L}]\underline{B}\underline{B}'[K-L]e^{(\underline{A}-\underline{B}\underline{B}'\underline{L})t}\underline{x}_0 dt = J_L \tag{24}$$

By virtue of the fact that feedback does not destroy controllability, theorem 4.3.2 and theorem 4.3.3, guarantee that the matrix $\underline{M}$ is the unique positive semidefinite solution of the algebraic equation

$$[\underline{A}-\underline{B}\underline{B}'\underline{L}]'\underline{M} + \underline{M}[\underline{A}-\underline{B}\underline{B}'\underline{L}] - \underline{M}\underline{B}\underline{B}'\underline{M} = -\underline{D}\underline{D}' \tag{25}$$

then by theorem 4.2.3

$$J_L - J^* \leq \frac{\|G\|^2}{1+\|G\|^2}\left(\int_0^\infty \underline{x}_0' e^{(\underline{A}-\underline{B}\underline{B}'\underline{L})'t}\underline{D}\underline{D}' e^{(\underline{A}-\underline{B}\underline{B}'L)t}\underline{x}_0 dt\right) \tag{26}$$

By using the fact that the eigenvalues of $(\underline{A}-\underline{B}\underline{B}'L)$ are in the left-half plane then

$$\underline{x}_0'L\underline{x}_0 = J_L - \int_0^\infty \underline{x}_0' e^{(\underline{A}-\underline{B}\underline{B}'L)'t}\underline{D}\underline{D}' e^{(\underline{A}-\underline{B}\underline{B}'L)t}\underline{x}_0 dt \tag{27}$$

therefore from equations (26) and (27) the claim of the theorem follows.| |

4. Frequency Domain Interpretation

One of the advantages of treating some problems in the frequency domain is that a great deal of insight is gained and new interpretation of results can be found that otherwise, in the time domain, would be very difficult to give. Above all, the most important advantages of the frequency domain analysis is that fundamental properties, such as the input output relations, are coordinate free (i.e., they do not depend upon the particular choice of the state variables).

Consider the algebraic equation that the matrix $\underline{L}$ satisfies (i.e., 4.3.17), then, by adding and subtracting $\underline{L}$'s to the left-hand side and rearranging terms it follows that

$$(-s\underline{I} - \underline{A} + \underline{B}\underline{B}'\underline{L})'\underline{L} + \underline{L}(\underline{I}s - \underline{A} + \underline{B}\underline{B}'\underline{L}) = \underline{C}\underline{C}' - \underline{D}\underline{D}' + \underline{L}\underline{B}\underline{B}'\underline{L} \tag{1}$$

Defining now

$$\underline{\Phi}_L = (\underline{I}s - \underline{A} + \underline{B}\underline{B}'\underline{L})^{-1} \tag{2}$$

then, by pre- and post-multiplying the terms of equation (1) by $\underline{B}'\underline{\Phi}'_L(-s)$ and $\underline{\Phi}_L(s)\underline{B}$ respectively, it follows that

$$\begin{aligned}&\underline{B}'\underline{L}\underline{\Phi}_L(s)\underline{B} + \underline{B}'\underline{\Phi}'_L(-s)\underline{L}\underline{B}\\ &= \underline{B}'\underline{\Phi}'_L(-s)\underline{C}\underline{C}'\underline{\Phi}_L(s)\underline{B} - \underline{B}'\underline{\Phi}'_L(-s)\underline{D}\underline{D}'\underline{\Phi}_L\underline{B} + \underline{B}'\underline{\Phi}'_L(-s)\underline{L}\underline{B}\underline{B}'\underline{L}\Phi_L(-s)\underline{B}\end{aligned} \tag{3}$$

By calling $\underline{K}_L = \underline{L}\underline{B}$ (the suboptimal feedback gain matrix) then

$$\begin{aligned}&\underline{B}'\underline{\Phi}'_L(-s)\underline{D}\underline{D}'\underline{\Phi}_L(+s)\underline{B}\\ &= \underline{B}'\underline{\Phi}'_L(-s)\underline{C}\underline{C}'\underline{\Phi}_L(s)\underline{B} + [\underline{I} - \underline{B}'\underline{\Phi}'_L(-s)\underline{K}_L][\underline{I} - \underline{K}_L\underline{\Phi}_L(s)\underline{B}]\end{aligned} \tag{4}$$

It should be noted that $\underline{G}(s) = \underline{D}'\underline{\Phi}_L(s)\underline{B}$ is the Laplace transform associated with the convolution operator G of equation (3.23). Similarly $\underline{H}_K(s) = \underline{C}'\underline{\Phi}_L(s)\underline{B}$ is the transfer function of the close loop system with the feedback law $\underline{u} = -\underline{K}'_L\underline{x}$ then equation (4) takes the simple form

$$\underline{G}'(-s)\underline{G}(s) = \underline{H}'_K(-s)\underline{H}_K(s) + [\underline{I} - \underline{B}'\underline{\Phi}'_L(-s)\underline{K}_L][\underline{I} - \underline{K}'_L\underline{\Phi}_L(s)\underline{B}] - I \tag{5}$$

by replacing $s = j\omega$ then

$$\underline{G}'(-j\omega)\underline{G}(-j\omega) = \underline{H}'_K(-j\omega)\underline{H}_K(j\omega) + [\underline{I} - \underline{B}'\underline{\Phi}_L(-j\omega)\underline{K}_L][\underline{I} - \underline{K}_L\underline{\Phi}_L(j\omega)\underline{B}] - \underline{I} \tag{6}$$

From equation (5) it follows that if the matrix $\underline{L}$ satisfies an equation of the form of (3.17) implies (and is implied by) the condition†

$$\|\underline{H}_K(j\omega)\| + \|\underline{I} - \underline{K}_i \underline{\Phi}_L(j\omega)\underline{B}\| - \underline{I} \geq 0 \tag{7}$$

Note that only under optimality of the feedback law the equality holds, and in that case inequality (7) becomes the main result of Kalman.[K-1]

In order to interpret inequality (3.3) in its simplest form assume that the system Σ_1 is a system with a scalar transfer function, then $\underline{B}$, $\underline{K}_L$ and $\underline{C}$ become vectors and will be denoted by $\underline{b}$, $\underline{k}$ and $\underline{c}$ respectively. Under these circumstances equation (5) becomes

$$\|\underline{G}(j\omega)\| = |h_k(j\omega)|^2 + |1 - \underline{k}_L' \underline{\Phi}_L(j\omega)\underline{b}|^2 - 1 \tag{8}$$

Now, since $\underline{G}(j\omega)$ is exactly the operator G of theorem 4.3.4 in the frequency domain,‡ then in order to find a bound for the suboptimal control law it is necessary to have

$$|h_k(j\omega)|^2 + |1 - \underline{k}_L \underline{\Phi}_L(j\omega)\underline{b}\Phi^2 \geq 1 \tag{9}$$

and if

$$\Delta = \sup_\omega \{ |h_k(j\omega)|^2 + |1 - \underline{k}_L \underline{\Phi}_L(j\omega) b|^2 \} \tag{10}$$

then

$$\frac{J^*}{J_L} \geq \frac{1}{\Delta}$$

Notice that $1 - \underline{k}_i \underline{\Phi}(j\omega)\underline{b}$ in "classical control" language is nothing else but the inverse of the "return difference" of the feedback law $-\underline{k}_L \underline{x}$.[H-2]

---

† $\|\underline{A}\|$ denotes $\tilde{\underline{A}}\ \underline{A}$ where $\tilde{A}$ is the adjoint (complex conjugate transpose) of $\underline{A}$.

‡ $\underline{G}(j\omega)$ is the Fourier transform of the L_1 function $\underline{D}' e^{(\underline{A} - \underline{b}\underline{k}_L')t} \underline{b}$

If the return difference is denoted by $T_k(j\omega)$ the conditions for optimality becomes simply

$$|h_k(j\omega)|^2 + \frac{1}{|T_k(j\omega)|^2} = 1 \tag{11}$$

which says that the square of the magnitude of transfer function plus the inverse square of the return difference should be unity for all frequencies. The theorem of Kalman[K-1] which states that a feedback law is optimal with respect to some quadratic cost functional if and only if $|T_k(j\omega)| \geq 1$, follows immediately from the relation (11). In addition, we obtain the condition that the gain of the closed loop transfer function should be less than unity. Going back to our problem of suboptimality of given feedback law, in order to be able to apply our criterion of suboptimality we require that

$$|h_k(j\omega)|^2 + \frac{1}{|T_k(j\omega)|^2} \geq 1 \tag{12}$$

then

$$\frac{J^*}{J_L} \geq \frac{1}{\sup_{\omega} \{ |h_k(j\omega)|^2 + \frac{1}{|T_k(j\omega|^2} \}} \tag{13}$$

The importance of the above result stems from its simplicity since once a feedback law is given then condition (12) is easily checked and the left-hand side of (13) is not difficult to compute, it is possible to give a quick measure of the degree of suboptimality of a given feedback law.

5. Applications

Suboptimal design with restricted feedback structure. One realistic assumption from the practical point of view is that not all the state variables of a system to be optimized can be measured directly.

Another restriction from the implementation point of view is that the feedback structure is restricted by the available instrumentation. With these constraints in mind the following example illustrates an application of some of the inequalities presented in this chapter.

Problem:

Given a system described by the equations (see Fig. 7)

$$\frac{d^2}{dt^2} x(t) - \frac{dx(t)}{dt} + 3x(t) = u(t)$$

$$y(t) = \frac{dx(t)}{dt} + x(t)$$

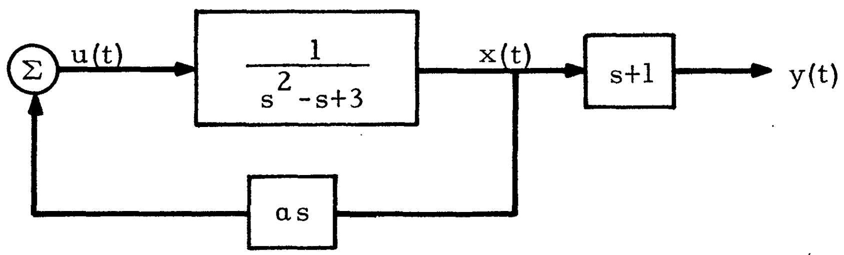

Fig. 7 System to be Optimized

the restriction is that $u(t) = -a \frac{dx}{dt}$ and the cost functional

$$J = \int_0^\infty \{u^2(t) + y^2(t)\}dt$$

The problem is to determine a value of a such that the performance of the system is "close" to the optimal performance (if all the states were measurable and the structure of the feedback were not restricted).

Solution:

The closed loop transfer function of the system with a feedback of the form $u(t) = -a \frac{dx(t)}{dt}$ is given by

$$h(s) = \frac{s+1}{s^2+(a-1)s+3}$$

and the return difference is given by

$$T_\alpha = \frac{s^2 + (\alpha-1)s+1}{s^2 - s + 3}$$

In order to have stability of the feedback system it is necessary to restrict α to be larger than unity. On the other hand, to satisfy condition (4.12) it is necessary to have

$$\frac{(-\omega^2+3)^2 + 2\omega^2 + 1}{(-\omega^2+3)^2 + (\alpha-1)^2\omega^2} \geq 1$$

which implies that $\alpha \leq \sqrt{2} + 1$, therefore

$$1 < \alpha \leq \sqrt{2} + 1$$

If α is chosen to be $\sqrt{2} + 1$, then Δ (as given by equation 4.10) is

$$\Delta = \frac{5}{4}$$

therefore $$\frac{J^*}{J_s} \geq \frac{4}{5}$$

In Fig. 8 a lower bound for the ratio of $\frac{J^*}{J_s}$ is displayed as a function of α, for the allowable range of α.

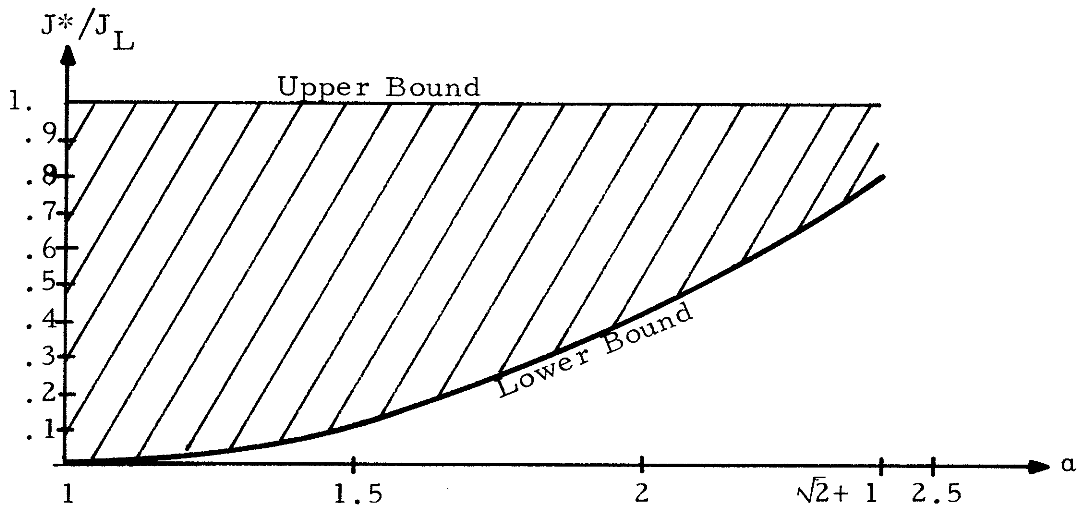

Fig. 8 Region where the Performance Ratio should lie as a function of the parameter α

However, the above lower bound is quite conservative since it is based on collorary 4.1. A sharper result is obtained if full advantage of the lower bound derived in theorem 4.4 is taken, that is, when the term $\frac{g^2}{1+g^2} \frac{\underline{x}_0' \underline{L} \underline{x}_0}{J_L}$ is included. Of course more computations are needed but still they are quite simple.

Let us assume, based on the graph of Fig. 8, that it has been decided to make

$$u(t) = -(\sqrt{2} + 1) \frac{dx}{dt}$$

We will compute the corresponding $\underline{L}$ matrix and the cost functional J_L. The matrix $\underline{L}$ should be positive definite and satisfy the equation

$$\underline{A}'\underline{L} + \underline{L}\,\underline{A} - \underline{L}\,\underline{B}\,\underline{B}'\underline{L} + \underline{C}\,\underline{C}' = \underline{D}'\underline{D}$$

In our case

$$\underline{A} = \begin{bmatrix} 0 & 1 \\ -3 & 1 \end{bmatrix} \quad ; \quad \underline{B}'\underline{L} = \begin{bmatrix} 0 \\ \sqrt{2}+1 \end{bmatrix} \quad ; \quad \underline{C}\,\underline{C}' = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

then the value obtained for $\underline{L}$ is given by

$$\underline{L} = \begin{bmatrix} 6.23 & 0 \\ 0 & 2.41 \end{bmatrix}$$

In order to obtain $J_L (= \underline{x}_0' \underline{Q} \underline{x}_0)$ it is necessary to solve the linear equation

$$(\underline{A} - \underline{B}\,\underline{B}'\underline{L})\underline{Q} + \underline{Q}(\underline{A} - \underline{B}\,\underline{B}'\underline{L}) = -\underline{C}\,\underline{C}' - \underline{L}\,\underline{B}\,\underline{B}'\underline{L}$$

the corresponding value of the matrix $\underline{Q}$ is

$$\underline{Q} = \begin{bmatrix} 6.83 & .166 \\ .166 & 2.53 \end{bmatrix}$$

therefore

$$\frac{\underline{x}_0' \, \underline{L} \, \underline{x}_0}{J_L} = \frac{6.23\beta^2 + 2.41}{6.83\beta^2 + 0.33\beta + 2.53}$$

where $$\beta = \frac{x_1(0)}{x_2(0)}$$

then a lower bound for the performance of the feedback $u(t) = -(\sqrt{2} + 1)\frac{dx(t)}{dt}$ is given by

$$\frac{J^*}{J_L} \geq 0.8 + 0.2 \frac{6.23\beta^2 + 2.41}{6.83\beta^2 + 0.33\beta + 2.53}$$

In Fig. 9, a graph of the lower bound on the performance ratio versus the ratio of the initial position to the initial velocity is given. The result is quite impressive since the minimum of that lower bound occurs in the neighborhood of $\beta=1$ and its value is larger than .977, therefore

$$\frac{J^*}{J_s} \geq .977$$

then, no matter what type of feedback is used (even by allowing the feedback of $x(t)$) the performance cannot be improved by more than two percent! The optimal feedback solution, as computed by digital computer was obtained to be

$$u(t) = -.158\, x(t) - 2.521 \frac{dx(t)}{dt}$$

However, when simulation was made for the worst initial condition, that is, $x(0) = 1$ and $\frac{dx(0)}{dt} = 1$ it was difficult to distinguish between the optimal and suboptimal trajectories (see Figs. 10a and 10b).

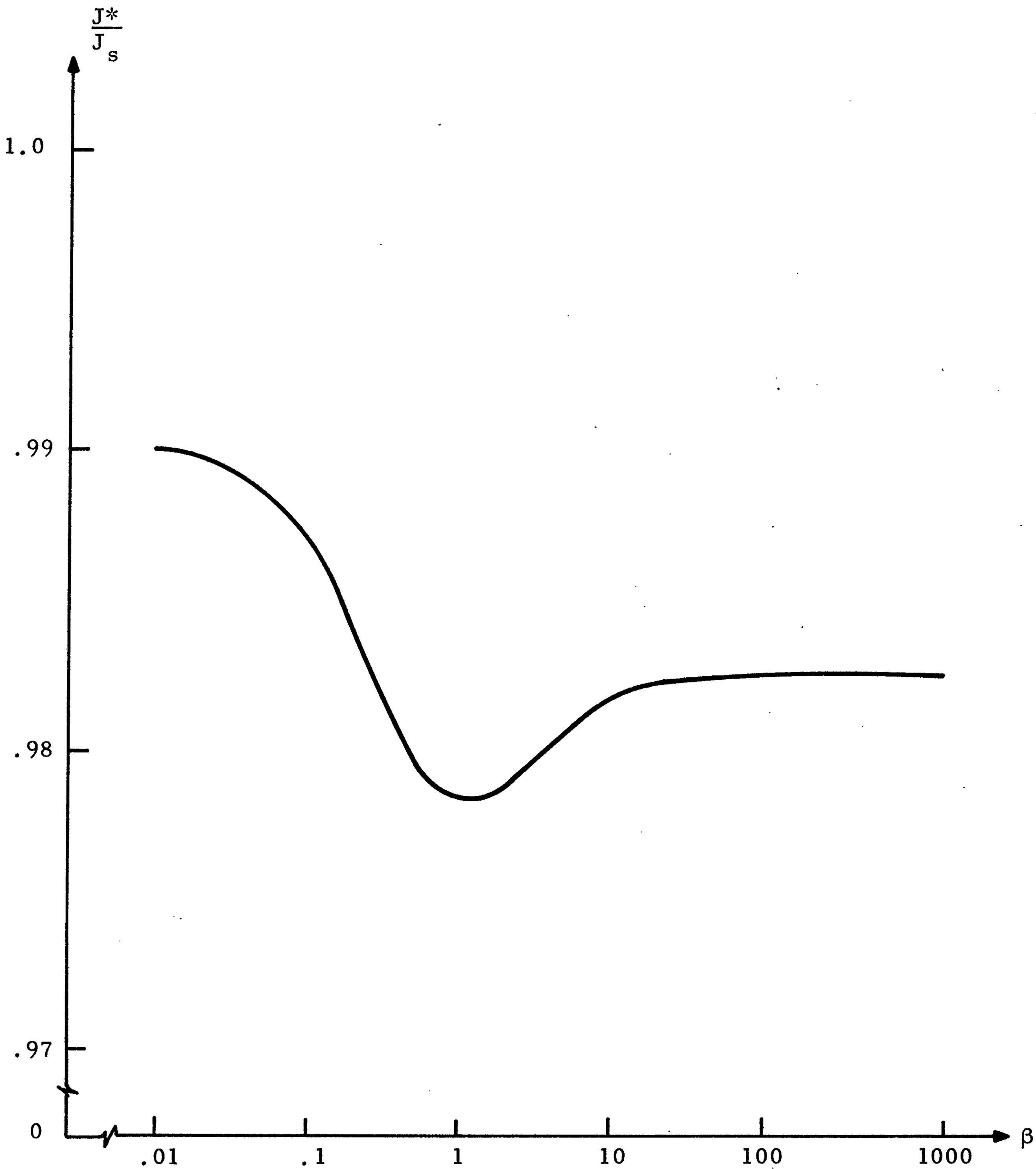

Fig. 9 Lower bound on the suboptimal performance as a function of the ratio of the initial position to the initial velocity.

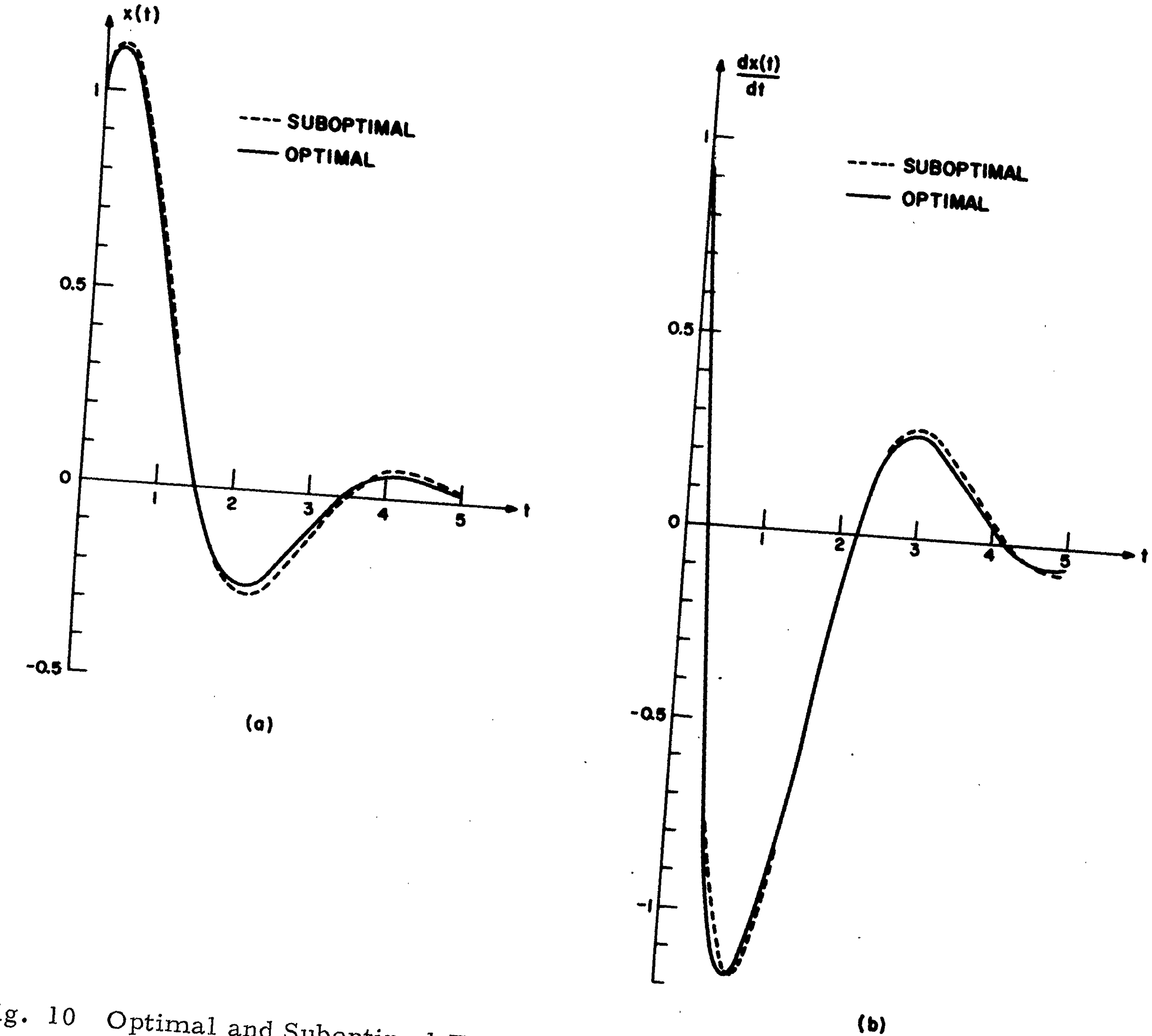

Fig. 10 Optimal and Suboptimal Trajectories for the "Worst" Initial Condition

Suboptimal Design of Weakly-Coupled Systems

The objective of this application is to demonstrate that by the use of the bounds derived in the present chapter it is possible to find sub-optimal controls for large systems, composed of two (or more) weakly coupled systems, with a guaranteed performance ratio, by finding the optimal feedback controls of the individual systems that make up the large system. In order to present the above idea in a more precise context consider, for simplicity, that the large system is composed of only two weakly coupled systems.

Assume that a system, denoted by S_1 is given by

$$\dot{\underline{x}}_1 = \underline{A}_1 \underline{x}_1 + \underline{B}_1 \underline{u}_1$$

$$\underline{y}_1 = \underline{C}_1' \underline{x}_1$$

and a system S_2 given by

$$\dot{\underline{x}}_2 = \underline{A}_2 \underline{x}_2 + \underline{B}_2 \underline{u}_2$$

$$\underline{y}_2 = \underline{C}_2' \underline{x}_2$$

and the large scale system S_L given by

$$\dot{\underline{x}} = \begin{bmatrix} \dot{\underline{x}}_1 \\ \dot{\underline{x}}_2 \end{bmatrix} = \begin{bmatrix} \underline{A}_1 & \underline{A}_{12} \\ \underline{A}_{21} & \underline{A}_2 \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} + \begin{bmatrix} \underline{B}_1 & \underline{0} \\ \underline{0} & \underline{B}_2 \end{bmatrix} \begin{bmatrix} \underline{u}_1 \\ \underline{u}_2 \end{bmatrix}$$

$$\underline{y} = \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix} = \begin{bmatrix} \underline{C}_1 & \underline{0} \\ \underline{0} & \underline{C}_2 \end{bmatrix} \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix}$$

The S_L is said to be weakly coupled if the matrices $\underline{A}_{12}$ and $\underline{A}_{21}$ are in some sense "small" relative to the matrices $\underline{A}_1$ and $\underline{A}_2$. Intuitively it would be expected that the optimal feedback law that minimizes the cost functional

$$J = \int_0^\infty \{\underline{u}_1'\underline{u}_1 + \underline{u}_2'\underline{u}_2 + \underline{y}_1'\underline{y}_1 + \underline{y}_2'\underline{y}_2\}dt$$

will be "close" to the decoupled feedback law generated when the systems S_1 and S_2 are assumed to be decoupled (i.e., when $\underline{A}_{12}$ and $\underline{A}_{21}$ are identically zero).

With the aid of the results presented in this chapter it is possible to give a qualitative measure of the degree of suboptimality when the above approximation is made.

In order to be able to apply the results derived in Section 4 certain preliminary considerations are in order.

Assume that $\underline{K}_1$ and $\underline{K}_2$ are respectively the unique positive definite solutions of the equations

$$\underline{A}_1'\underline{K}_1 + \underline{K}_1\underline{A}_1 - \underline{K}_1\underline{B}_1\underline{B}_1'\underline{K}_1 = -\underline{C}_1\underline{C}_1' \tag{1}$$

$$\underline{A}_2'\underline{K}_2 + \underline{K}_2\underline{A}_2 - \underline{K}_2\underline{B}_2\underline{B}_2'\underline{K}_2 = -\underline{C}_2\underline{C}_2' \tag{2}$$

the matrix

$$\underline{L} \triangleq \begin{bmatrix} \underline{K}_1 & \underline{0} \\ \underline{0} & \underline{K}_2 \end{bmatrix}$$

does not necessarily satisfy the condition imposed by equation 4.3.17, that is the matrix

$$\underline{M} = \begin{bmatrix} \underline{A}_1' & \underline{A}_{21}' \\ \underline{A}_{12}' & \underline{A}_2' \end{bmatrix} \underline{L} + \underline{L} \begin{bmatrix} \underline{A}_1 & \underline{A}_{12} \\ \underline{A}_{21} & \underline{A}_2 \end{bmatrix} - \underline{L} \begin{bmatrix} \underline{B}_1 & \underline{0} \\ \underline{0} & \underline{B}_2 \end{bmatrix} \begin{bmatrix} \underline{B}_1' & \underline{0} \\ \underline{0} & \underline{B}_2' \end{bmatrix} \underline{L} + \begin{bmatrix} \underline{C}_1\underline{C}_1' & 0 \\ \underline{0} & \underline{C}_2\underline{C}_2' \end{bmatrix}$$

$$= \begin{bmatrix} \underline{0} & \underline{A}_{21}'\underline{K}_2 + \underline{K}_1\underline{A}_{12} \\ \underline{K}_2\underline{A}_{21} + \underline{A}_{12}'\underline{K}_1 & \underline{0} \end{bmatrix}$$

is not always positive semidefinite. However if we consider instead

$$\underline{L}(\epsilon,\delta) = \begin{bmatrix} (1-\epsilon)\underline{K}_1 & 0 \\ 0 & (1-\delta)\underline{K}_1 \end{bmatrix}$$

then

$$\underline{M}(\epsilon,\delta) = \begin{bmatrix} \epsilon\, \underline{C}_1\underline{C}_1' + \epsilon(1-\epsilon)\underline{K}_1\underline{B}_1\underline{B}_1'\underline{K}_1 & (1-\delta)\underline{A}_{21}'\underline{K}_2 + (1-\epsilon)\underline{K}_1\underline{A}_{12} \\ (1-\delta)\underline{K}_2\underline{A}_{21} + (1-\epsilon)\underline{A}_{12}'\underline{K}_1 & \delta\underline{C}_2\underline{C}_2' + \delta(1-\delta)\underline{K}_2\underline{B}_2\underline{B}_2'\underline{K}_2 \end{bmatrix}$$

can be made positive semidefinite if δ and ϵ are made sufficiently large $(0 < \epsilon < 1\ ,\ 0 < \delta < 1)$. It should be kept in mind that for our criterion of suboptimality to be applicable it is required to have the suboptimal closed loop system stable.

If $\underline{M}(\epsilon,\delta)$ is positive semidefinite then

$$\frac{J^*}{J_L} \geq \frac{1}{1+g^2}$$

where g is the appropriate gain defined by equation (3.23).

In order to demonstrate the procedure in detail consider the following specific system:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & .05 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & -2 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -2 & 1 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\underline{y} = \underline{x}$$

and the cost functional

$$J = \int_0^\infty \{ \sum_{i=1}^{2} u_i^2 + \sum_{i=1}^{6} y_i^2 \} dt$$

If we consider the two systems

$$S_1: \quad \frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u_1 \; ; \; \underline{y}_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$S_2: \quad \frac{d}{dt}\begin{bmatrix} x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_4 \\ x_5 \\ x_6 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u_2 \; ; \; \underline{y}_2 = \begin{bmatrix} x_4 \\ x_5 \\ x_6 \end{bmatrix}$$

then the positive definite solutions of equations (1) and (2) are respectively

$$\underline{K}_1 = \begin{bmatrix} 2.19 & 1.83 & .41 \\ 1.83 & 3.81 & .96 \\ .41 & .96 & .45 \end{bmatrix}$$

$$\underline{K}_2 = \begin{bmatrix} .921 & 4.2 & .24 \\ 4.20 & 1.36 & 4.22 \\ .24 & 4.22 & 1.67 \end{bmatrix}$$

At this point it is necessary to pick the values of ϵ and δ in order to guarantee that the matrix $\underline{M}$ is positive semidefinite. One way of doing so is to chose ϵ and δ large enough as to make the matrix $\underline{M}$ diagonal dominant, that is that for every row the diagonal element is larger in magnitude than the sum of the absolute value of the off diagonal elements of that row. By Gersagorin theorem[G-2] all the eigenvalues of $\underline{M}$ will be positive. By this method the tedious task of determining the nonnegativity (Sylvester test) of $\underline{M}$ will be avoided. However on the other hand the degree of suboptimality will be increased. For our particular example the values of the parameters δ and ϵ

chosen are .01 and .1 respectively. For these values of ϵ and δ the corresponding $\underline{M}$ matrix is

$$\underline{M} = \begin{bmatrix} .115 & .035 & .016 & 0 & .060 & 0 \\ .035 & .185 & .038 & 0 & -0.04 & 0 \\ .016 & .038 & .118 & 0 & -.021 & 0 \\ 0 & 0 & 0 & .011 & .010 & .040 \\ .060 & -.040 & -.021 & .010 & .188 & .070 \\ 0 & 0 & 0 & .04 & .070 & .028 \end{bmatrix}$$

which is indeed positive definite. The closed loop system with the suboptimal feedback corresponding to the values of ϵ and δ becomes

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & .05 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -1.38 & -2.86 & -3.40 & 0 & -.1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -2.24 & -3.22 & -3.67 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}$$

and the suboptimal feedback law is

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = -\begin{bmatrix} .38 & .86 & -.40 & 0 & 0 & 0 \\ 0 & 0 & 0 & .24 & 4.22 & 1.67 \end{bmatrix}\underline{x}$$

In order to have a measure of the degree of suboptimality of the given control law it is necessary to compute the gain of the operator G defined by (4.23). After some algebraic manipulations the following relation is obtained:

$$\underline{G}(s)\underline{G}(-s) = \begin{bmatrix} \dfrac{.115-.153s^2+.118s^4}{p_1(s)p_1(s)} & \dfrac{.06s-.04s^2+.021s^3}{p_1(-s)p_2(+s)} \\[2ex] \dfrac{-.06s+.04s^2-.021s^3}{p_1(s)p_2(-s)} & - \dfrac{.011-.108s^2+.028s^4}{p_2(s)p_2(-s)} \end{bmatrix}$$

where $p_1(s) = (s^3 + 3.40s^2 + 2.86s + 1.38)$

and $p_2(s) = (s^3 + 3.67s^2 + 3.22s + 2.24)$

the value of $\|G\|^2$ is .092

therefore $\frac{J^*}{J_s} \geq \frac{1}{1.092} = .91$

then, the suboptimal feedback law has a guaranteed performance ratio of .91. However as it was demonstrated vividly in the previous example this lower bound is quite conservative since the term due to the initial conditions has been ignored.

CHAPTER V

CONCLUSIONS

The research presented in the first part of this thesis was motivated by the lack of systematic methods in finding optimal feedback control laws for systems with nonlinear dynamics and quadratic cost functionals. The contribution of the present work to the field of optimal control has been the study of the possibility of being able to give upper bounds to the optimal performance of a nonlinear system by computing the optimal performance of a linear system. One of the methods used was to derive the optimal feedback law of a linearized version of the nonlinear system and apply that control law to the nonlinear system. However this method had the disadvantage that in order to apply it, it is necessary to determine the closed loop stability characteristic of the nonlinear system with linear feedback. In order to overcome this difficulty a second method was investigated--namely, use of a nonlinear feedback control law such that the trajectories thus generated coincide with the optimal trajectories of a linear system, known a priori to be stable. In the development of this second method extensive use was made of the known results of the problem of determining when the product of two operators is a positive operator, thus bringing results already widely used in stability theory to the area of optimal control. Specific conditions were given under which the second method can be applied.

The second part of the research dealt with the problem of obtaining lower bounds on the performance of linear systems with quadratic cost functionals. An attainable lower bound was derived in terms of a suboptimal performance, thus obtaining, a qualitative measure of the degree of suboptimality of a given design. The lower bound was given in terms of a measure of the deviation of a given control law from satisfying a certain optimality condition (the Riccati equation). Two numerical examples were presented in order to illustrate the usefulness of the derived lower bound in the design of suboptimal systems, when, either it is desired to avoid a large number of computations or, due to practical limitations, there exists a structural constraint on the class of allowable feedback controls. The main contribution of the second part of this research has been the derivation of a simple measure of the degree in which a given suboptimal design can be improved. This measure is specified entirely in terms of the given design hence avoiding the need for computation of the optimal performance.

APPENDIX A

POSITIVE REAL FUNCTIONS

Definition 1A[K-2] A function f of a complex variable s - f(s) - is said to be positive real (p. r. in short) if

$$\text{Re}[f(s)] \geq 0$$

throughout the region

$$\text{Re}[s] \geq 0$$

of the complex s-plane where f(s) is defined and†

$$f(\tilde{s}) = \tilde{f}(s) \qquad \text{for all complex } s.$$

Theorem 1A [Weinberg and Slepian][W-1]

If f(s) is a rational function of s - $f(s) = \frac{q(s)}{p(s)}$ - where q(s) and p(s) do not have common factors, then a necessary and sufficient condition for f(s) to be positive real is that

a) $\text{Re}\left[\frac{q(j\omega)}{p(j\omega)}\right] \geq 0$ for all real ω

b) $q(s) + p(s)$ has all its zeros in $\text{Re}[s] < 0$

Proof: Necessity

Assume that f(s) is positive real, then (a) follows immediately from the definition of positive realness

If $q(s_0) + p(s_0) = 0$ then

$$\frac{q(s_0)}{p(s_0)} = -1$$

†$\tilde{a}$ indicates the complex conjugate of a.

therefore

$$\operatorname{Re} \frac{q(s_0)}{p(s_0)} = -1 < 0$$

but since f(s) is positive real it follows that $\operatorname{Re}[s_0] < 0$

Sufficiency Define

$$w(s) = \frac{1 - \frac{q(s)}{p(s)}}{1 + \frac{q(s)}{p(s)}} = \frac{p(s) - q(s)}{p(s) + q(s)}$$

then

$$\operatorname{Re}\left[\frac{q(s)}{p(s)}\right] \geq 0 \quad \text{if and only if } |w(s)| \leq 0$$

Apply now the maximum modulus theorem[R-1] to the restriction w(s) to the closure of the right-half plane and use the fact that w(s) is analytic in $\operatorname{Re}[s] \geq 0$ and $\operatorname{Re}[f(j\omega)] \geq 0$ in order to conclude that

$$|w(s)| \leq 1 \quad \text{for} \quad \operatorname{Re}[s] \geq 0, \qquad \text{then} \qquad \operatorname{Re}\left[\frac{q(s)}{p(s)}\right] \geq 0 \quad \text{for} \quad \operatorname{Re}[s] \geq 0$$

Theorem 2A (on integral properties of p. r. functions)

$$\int_0^\infty p(D)x(t)q(D)x(t)dt \leq 0 \quad \text{for all} \quad x(t) \in \mathscr{D}_0^n[0, \infty)$$

$$\text{if and only if } -\frac{p(-s)}{q(-s)} \text{ is p. r.}$$

Proof: Sufficiency

$$\int_0^\infty p(D)x(t)q(D)x(t)dt = \int_0^\infty \{p(D)x(t)q(D)x(t) + [(\mathrm{Ev}[\bar{p}q]^{-}(D)x(t)]^2\}dt$$

$$- \int_0^\infty \{(\mathrm{Ev}[\bar{p}q])^{-}(D)x(t)\}^2 dt$$

The second integral is negative. The first integral is independent of path [B-1] and (depends only on $\underline{x}(0)$ and $\lim_{t\to\infty} \underline{x}(t)$) but since $x(t) \in \mathcal{D}_0^n[0,\infty)$ then

$$\lim_{t\to\infty} \underline{x}(t) = \underline{0}$$

therefore we can write

$$F[\underline{x}(0)] = \int_0^{\infty} \{p(D)x(t)q(D)x(t) + [(Ev[\bar{p}q])^-(D)x(t)]^2\}\, dt$$

$$= \int_{t[\underline{x}(0)]}^{t[\underline{0}]} \{p(D)z(t)q(D)z(t) + [(Ev[\bar{p}q])^-(D)z(t)]^2\}dt$$

where

$$\underline{z}[t(\underline{x}(0))] = \underline{x}(0) \quad \text{and} \quad \underline{z}[t(\underline{0})] = \underline{0}.$$

Changing now t by $-t$, it is obtained that

$$F[\underline{x}(0)] = \int_{-t[\underline{x}(0)]}^{-t(\underline{0})} \{p(-D)z(-t)q(-D)z(-t)] - [(Ev[\bar{p}q])^+(+D)z(-t)]^2\}dt$$

Choose a $z(t)$ such that

$$p(-D)z(-t) - q(-D)z(-t) = 0$$

with $\underline{z}\{t[\underline{x}(0)]\} = \underline{x}(0)$. But since $-\frac{p(-s)}{q(-s)}$ is p. r. $\underline{z}[t(\underline{0})] = \underline{0}$ then

$$F[\underline{x}(0)] = \int_{-t[\underline{x}(0)]}^{-t(\underline{0})} \{-[p(-D)z(-t)]^2 - [(Ev[\bar{p}q])^+(D)z(-t)]^2\}dt$$

which is certainly nonpositive.

Necessity: This is more complicated. Here we want to show that if $-\frac{p(-s)}{q(-s)}$ is not p. r. then

$$\int_0^\infty p(D)x(t)q(D)x(t)dt > 0 \quad \text{for some} \quad x(t) \in \mathcal{D}_0^n[0,\infty)$$

If $-\frac{q(-s)}{p(-s)}$ is not p. r. then, either

a) there exists an s_0 with positive real part such that

$$q(-s_0) - p(-s_0) = 0 \tag{1}$$

and

$$\text{Re}\left[-\frac{q(-j\omega)}{p(-j\omega)}\right] \geq 0 \quad \text{for all real } \omega, \text{ or} \tag{2}$$

b) $$\text{Re}\left[-\frac{q(-j\omega_0)}{p(-j\omega_0)}\right] < 0 \quad \text{for some real } \omega_0 \tag{3}$$

Consider the first case. Define

$$x_0(t) = a_0 e^{-s_0 t} + \tilde{a}_0 e^{-\tilde{s}_0 t} \in \mathcal{D}_0^n[0,\infty)$$

then

$$\alpha[x_0(t)] = \int_0^\infty p(D)x_0(t)q(D)x_0(t)dt$$

$$= \int_0^\infty [p(-s_0)a_0 e^{-s_0 t} + p(-\tilde{s}_0)\tilde{a}_0 e^{-\tilde{s}_0 t}][q(-s_0)a_0 e^{-s_0 t}$$

$$+ q(-\tilde{s}_0)\tilde{a}_0 e^{-\tilde{s}_0 t}]dt$$

but by virtue of (1)

$$\alpha[x_0(t)] = \int_0^\infty [p(-s_0)a_0e^{-s_0t} + p(-\tilde{s}_0)\tilde{a}_0e^{-\tilde{s}_0t}]^2 dt \geq 0$$

Consider now the second case. Inequality (3) can be written as

$$q(-j\omega_0)p(j\omega_0) + p(-j\omega_0)q(j\omega_0) > 0$$

observe that by continuity in ω it follows that there exists an $\epsilon > 0$ such that

$$q(-j\omega_0-\sigma_0)p(+j\omega_0-\sigma_0) + p(-j\omega_0-\sigma_0)q(j\omega_0-\sigma_0) \geq 0$$

for all $|\sigma_0| < \epsilon$

take now $x(t) = e^{-\sigma_0 t}[e^{-j\omega_0 t} + e^{j\omega_0 t}] \in \mathcal{D}_0^n[0,\infty)$ if $\sigma_0 > 0$

then

$$\sigma[x(t)] = \int_0^\infty p(D)x(t)q(D)x(t)dt$$

$$= \int_0^\infty e^{-2\sigma_0 t}[p(-j\omega_0-\sigma_0)e^{-j\omega_0} + p(j\omega_0-\sigma_0)e^{j\omega_0 t}][q(-j\omega_0-\sigma_0)e^{-j\omega_0 t}$$

$$+ q(j\omega_0-\sigma_0)e^{j\omega_0 t}]dt$$

$$= \int_0^\infty \{e^{-2\sigma_0 t}[p(-j\omega_0-\sigma_0)q(-j\omega_0-\sigma_0)e^{-2j\omega_0 t} + p(j\omega_0-\sigma_0)q(j\omega_0-\sigma_0)e^{2j\omega_0}$$

$$+ [p(-j\omega_0-\sigma_0)q(j\omega_0-\sigma_0) + q(-j\omega_0-\sigma_0)p(j\omega_0-\sigma_0)]e^{-2\sigma_0 t}\}dt$$

After performing the integration it is obtained that

$$a[x(t)] = \mathrm{Re}\left[\frac{p(j\omega_0-\sigma_0)q(j\omega_0-\sigma_0)}{2[-j\omega_0-\sigma_0]}\right] + \mathrm{Re}\left[\frac{p(j\omega_0-\sigma_0)q(-j\omega_0-\sigma_0)}{2\sigma_0}\right]$$

the term on the right is positive. Choose σ_0 small enough such that

$$\left|\mathrm{Re}\left[\frac{p(j\omega_0-\sigma_0)q(j\omega_0-\sigma_0)}{2(-j\omega_0-\sigma_0)}\right]\right| \leq \mathrm{Re}\left[\frac{p(j\omega_0-\sigma_0)q(-j\omega_0-\sigma_0)}{2\sigma_0}\right]$$

therefore $a[x(t)] > 0$.

APPENDIX B

Lemma 1B Given any monic real polynomial p(s) of order n (n> 2) and any complex number s_0 with $Re[s_0] \neq 0$ then there exists another real polynomial h(s) of order less than n such that

$$p(s_0)p(-s_0) + h(s_0)h(-s_0) = 0$$

Furthermore there exist a polynomial h of second order that satisfies the above equality.

Proof: i) Assume that s_0 is real. Then we will show that for any real number x there exist real numbers a and b such that:

$$x = (as_0+b)(-as_0+b) = -a^2s_0^2+b^2 \tag{1}$$

Simply choose a such that

$$a^2 > \frac{|x|}{s_0^2} \tag{2}$$

For that choice of a, b is guaranteed to be real.

ii) The proof when s_0 is complex $[Im(s_0) \neq 0]$ is more involved. We will show that for any complex number $x + jy$† then exists real constants a, b and c such that

$$(as_0^2+bs_0+c)(as_0^2-bs_0+c) = x + iy \tag{3}$$

or

$$(as_0^2+c)^2-b^2s_0^2 = x + iy \tag{4}$$

---

† $j = \sqrt{-1}$

then in order to satisfy equation (4) for any s_0, given that $s_0^2 = \sigma_0 + j\omega_0$ it is necessary to have

$$a^2(\sigma_0^2 - \omega_0^2) + (2ac - b^2)\sigma_0 + c^2 = x \tag{5}$$

$$2a^2\omega_0\sigma_0 + (2ac - b^2)\omega_0 = y \tag{6}$$

Substituting (6) into (5) it follows that

$$a^2 = \frac{-x + y/\omega_0 + c^2}{\sigma_0^2 + \omega_0^2} \tag{7}$$

therefore if c is chosen such that

$$c^2 > -x + y/\omega_0 \tag{8}$$

a will be real.

The problem now is to determine if it is possible to choose b real and satisfy (6). The expression for b^2 as a function c is given by

$$b^2 = \left(\frac{-x + y/\omega_0 + c^2}{\sigma_0^2 + \omega_0^2}\right)\sigma_0 + 2\left(\frac{-x + y/\omega_0 + c^2}{\sigma_0^2 + \omega_0^2}\right)^{1/2} c - y/\omega_0$$

If $c \rightarrow \infty$ then b approaches asymptotically to

$$b^2 \rightarrow 2c^2\left(\frac{\sigma_0}{\sigma_0^2 + \omega_0^2} + \frac{1}{\sqrt{\sigma_0^2 + \omega_0^2}}\right) - \frac{y}{\omega_0}$$

however

$$\frac{\sigma_0^2}{\sigma_0^2 + \omega_0^2} + \frac{1}{\sqrt{\sigma_0^2 + \omega_0^2}} > 0 \quad \text{for all } \sigma_0, \text{ then}$$

$b^2 \rightarrow$ positive number, therefore for c large enough both a and b will be real.

REFERENCES

A-1. Athans, M., "The Status of Optimal Control Theory and Applications for Deterministic Systems," IEEE Transactions on Automatic Control, Vol. 11, No. 3, July, 1966, pp. 580-596.

A-2. Athans, M. and Falb, P. F., Optimal Control, McGraw Hill Book Co., New York, 1966.

B-1. Brockett, R. W., "Path Integrals, Liapunov Functions and Quadratic Minimization," Proceedings of 4th Annual Allerton Conference on Circuits and Systems Theory, Urbana, Illinois, University of Illinois, 1966.

B-2. Brockett, R. W., "On the Stability of Non-Linear Feedback Systems," IEEE Transactions on Applications and Industry, Vol. 83, November, 1963, pp. 443-449.

B-3. Brockett, R. W., Qualitative Behavior of Time Varying and Non-Linear Feedback Systems, to be published by John Wiley and Sons, New York.

B-4. Brockett, R. W. "A Priori Bounds on the Solution of Regulator Problems," Control Theory Group, Electronic Systems Laboratory, M. I. T. (unpublished, undated internal memorandum).

B-5. Brockett, R. W. and Canales, R., "Organization of System Control," Electronic Systems Laboratory Report ESL-R-296, November, 1966.

B-6. Bellman, R., Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

B-7. Bellman, R. and Kalaba, R., "Quasilinearization and Non-Linear Boundary-Valued Problems," RAND Report No. R-438, 1965.

G-1. Gruber, M. and Willems, J. L., "A Generalization of the Circle Criterion," Proceedings of 4th Annual Allerton Conference on Circuits and Systems Theory, University of Illinois, Urbana, Illinois, 1966.

G-2. Grantmacher, F. R., The Theory of Matrices, Chelsea, 1959.

G-3. Gelfand, I. M. and Fomin, S. V., Calculus of Variations, Prentice-Hall Inc., Engelwood Cliffs, New Jersey, 1963.

H-1. Hanh, W., Theory and Applications of Lyapunov's Direct Method, Prentice Hall, 1963.

REFERENCES (Contd.)

H-2. Horowitz, I. M., Synthesis of Feedback Systems, Academic Press, New York, 1963.

H-3. Hadley, G., Linear Programming, Addison-Wesley Publishing Co., Reading, Mass., 1963.

K-1. Kalman, R. E., "When is a Linear Control System Optimal " Trans. ASME (Journal of Basic Engineering), Part D, March, 1968, pp. 51-60.

K-2. Kuh, E. S., and Rohrer, R. A., Theory of Linear Active Networks, Holden-Day Ind., 1967, p. 139.

K-3. Kalman, R. E., "Contributions to the Theory of Optimal Control," Boletin de la Sociedad Matematica Mexicana, 1960.

K-4. Kleiman, D. L., "On the Linear Regulator Problem and the Matrix Riccati Equation," Electronic Systems Laboratory Report 271, M.I.T., 1966.

L-1. Levis, A., "On the Optimal Sample-Date Control of Linear Processes," Sc.D. Thesis, Department of Mechanical Engineering, M.I.T., June, 1968.

L-2. Levine, W., and Athans, W. S., "On the Numerical Solution of the Matrix Riccati Differential Equation Using a Runge-Kutta Scheme," Electronic Systems Laboratory Report ESL-R-276, M.I.T., July, 1966.

M-1. McClamrock, "A Result on the Performance Deterioration of Optimum Systems," IEEE Trans. on Automatic Control, Vol. AC-12, April 1967, pp. 209-211. (correspondence)

P-1. Pontryagin, et al., Mathematical Theory of Optimal Processes, John Wiley and Sons, Inc., (Interscience), 1962.

R-1. Rudin, W., Real and Complex Analysis, McGraw-Hill Book Co., 1966. p. 213.

R-2. Rissanen, J. J., "Performance Deterioration of Optimum Systems," IEEE Trans. on Automatic Control, Vol. AC-11, July, 1966, pp. 530-532.

W-1 Weinberg, L., and Slepian, P., "Positive Real Matrices," Journal of Mathematics and Mechanics, Vol. 9, No. 1, 1960, pp. 71-83.

W-2 Wonham, W. M., "On the Matrix Quadratic Equations and Matrix Riccati Equations," Technical Report 67-5, Center for Dynamical Systems, Brown University, August, 1967.

REFERENCES (Contd.)

Y-1. Youla, D. C., "The Synthesis of Linear Dynamical Systems from Prescribed Weighting Patterns," J. SIAM Appl. Math., Vol. 14, No. 3, 1966.

Z-1. Zames, G., and Falb, P. L., "Stability Conditions for Systems with Monotone and Slope-Restricted Nonlinearities," SIAM Journal of Control, Vol. 6, No. 1, February, 1968, pp. 89-108.