OPTIMAL COMPENSATOR STRUCTURE FOR LINEAR TIME-INVARIANT PLANT WITH INACCESSIBLE STATES

Patrick Jean-Pierre Blanvillain

Electronic Systems Laboratory

Decision and Control Sciences Group

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139

Department of Electrical Engineering

OPTIMAL COMPENSATOR STRUCTURE FOR LINEAR TIME-INVARIANT
PLANT WITH INACCESSIBLE STATES

by

Patrick Jean-Pierre Blanvillain

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

OPTIMAL COMPENSATOR STRUCTURE FOR LINEAR TIME-INVARIANT

PLANT WITH INACCESSIBLE STATES

by

Patrick Jean-Pierre BLANVILLAIN

Ingenieur, Ecole Centrale de Lyon, France
(1973)
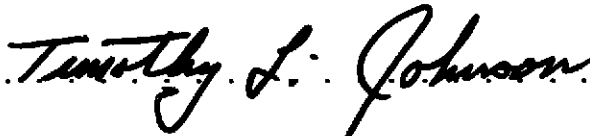
SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June, 1974

Signature of Author . . . .
                    Department of Electrical Engineering, May 10, 1974

Certified by . . . .
                                                  Thesis Supervisor

Accepted by . . . .
              Chairman, Departmental Committee on Graduate Students

OPTIMAL COMPENSATOR STRUCTURE FOR LINEAR TIME-INVARIANT

PLANT WITH INACCESSIBLE STATES

by

PATRICK JEAN-PIERRE BLANVILLAIN

Submitted to the Department of Electrical Engineering on May 10, 1974, in partial fulfillment of the requirements for the Degree of Master of Science.

ABSTRACT

This thesis considers the problem of designing an optimal linear time-invariant dynamic compensator for the regulation of an n-th order linear time-invariant plant with m independent outputs. The initial plant state is characterized by its first and second moments, and the cost is the usual quadratic infinite-time penalty on the state and control, averaged over the initial plant and compensator states.

The compensator is based on a minimal-order Luenberger observer and consequently has fixed dimension n-m. Necessary and sufficient conditions are derived for optimality of the compensator gains. The optimal compensator is shown to be unique if the plant has a particular canonical form, and, in general, for any arbitrary plant, the class of all optimal compensators is precisely determined.

The main contributions of this thesis are to show that for this problem an equivalent of the "separation property" holds, and that the compensator gains are uniquely determined in terms of a unique design matrix, which, in particular, completely specifies the dynamics of the minimal-order Luenberger observer. In addition, several examples are worked out, and generalizations which treat any sort of initial plant and compensator states are provided.

Thesis Supervisor: Timothy L. Johnson
Title: Assistant Professor of Electrical Engineering

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In the past two decades, a number of time domain methods have been
proposed to regulate linear plants. The so-called "linear regulator prob-
lem" has been extensively studied and the work of Kalman, [14], [15], [16],
[17], must be mentioned here. Recall, however, that for this linear regula-
tor problem, the optimal control requires the <u>entire state</u> to be generated.
On the contrary, in most applications, one does not have a direct access to
the complete state, but rather to part of it, through the output vector.
In [12], [18], Levine solves the problem of finding the optimal <u>output</u>
feedback controller, assuming that the plant is output stabilizable. Now,
there exist cases, and we will see an example in Section 414, where <u>constant</u>
output feedback gains cannot stabilize the plant (i.e., the plant is not
<u>output stabilizable</u>). In these cases, a <u>dynamic compensation</u> is required.
Several methods have then been proposed for designing dynamic compensators
to regulate n-th order linear plants with only m independent outputs.

In a stochastic context, a Kalman-Bucy filter [19] of dimension n is
used to give an estimate of the entire state from noisy measurements, and
then the control is generated with this estimate via the optimal gains we
would have for the noise-free linear regulator problem [20]; this is the well
known separation theorem [21] for the linear quadratic Gaussian problem.
In case the measurements are noise-free, but the plant is still driven by
a noisy process, a Bryson-Johansen filter [22] of dimensions n-m can be
designed; the separation property holds in this case also, and Bucy [23]

proved the optimality of this design.

Strangely enough, the state of affairs in a completely deterministic case is not so clearly defined. Mimicking what has been done in the stochastic case, a minimal-order Luenberger observer of dimension n-m, [2], [24], [25], is very often used to generate an estimate of the state from the available outputs, and then a control is generated with this estimate via the optimal gains for the linear regulator problem. Recall that if we know perfectly the initial state of the plant, a Luenberger observer can be designed, having arbitrary dynamics, that reconstructs perfectly the complete state; recall also that if we don't know the initial plant state, then the error between the true state and the estimate of the state can be never-theless be made to decay arbitrarily fast. The above design philosophy is then certainly a reasonable one. It can also be shown that in this case, the total cost is composed of two parts: the optimal cost for the linear regulator problem, (i.e., the optimal cost we would have if we had complete state feedback), and in addition, another term due to the fact that we have to reconstruct part of the state to implement the control. The increase in the cost incurred by the use of a dynamic compensator can then be precisely computed.

The above design philosophy does have an important practical drawback: there is no way to "optimally" determine the dynamics of the observer part of the compensator. It was then suggested that an arbitrarily fast observer would make the incurred increase in the cost arbitrarily small. However, it was soon demonstrated by counter-examples [8], [9], [10], that this was obviously wrong. The designer, then, is left with his "engineering judgment"

and very often, a common design will be one with an observer slightly faster than the plant, to insure good tracking properties. The conclusion is that if we retain this completely deterministic formulation, there is no general optimal solution for the compensator dynamics.

Notice now that very often we don't know the exact initial state of the plant. The idea (introduced in [12]) is then to design a compensator which will be optimal over a set of possible initial states for the plant. This leads to a design which is optimal in an "average sense." The fundamental results using this approach are given in [26]. In this paper, a new quadratic cost is introduced which weighs quadratically both the input signals to the plant and the input signals to the compensator. The optimal compensator gains are then found to be the solutions of coupled Riccati and Lyapunov-type equations. The main contribution of [26] comes from the fact that the approach taken does not require the compensator to have the dimension of the minimal order Luenberger observer (i.e., n-m). In particular, an optimal reduced order compensator of fixed dimension s (s < n-m) can be found with this approach (if one exists). A slight drawback of this method is that it provides no insight on the internal structure of the optimal compensator.

Following this "average-sense" optimal philosophy, what we will do in this thesis is to reconsider the problem of using a minimal-order Luenberger observer to estimate the state of the plant, and then apply a control generated with this estimate via the optimal gains for the linear regulator problem. However, we will look for a design which is optimal over some set of initial plant states. More precisely, we will consider the initial plant state as a random vector specified by its first and second moments.

As we will mainly consider time-invariant plants, the cost used will be the usual quadratic infinite time penalty on the state and control averaged over initial plant and compensator states.

We note that this approach has already been taken by some authors. Newmann in [5], [6] took it, but was only able to derive a set of necessary conditions for optimality of the compensator gains. However, these conditions consisted of a large set of non-linear coupled equations of the Riccati-type, and as a consequence were not readily solved, so that no explicit structure for the optimal compensator was given. More recently, Rom and Sarachik in [7] have given a partially satisfying answer to the problem. They actually identified the problem we are trying to solve with a steady-state stochastic control problem whose solution is known and requires the use of a Bryson-Johansen filter. Imposing some particular conditions on the plant and compensator initial states for their comparison to be valid, they actually showed by this indirect method one of the main results of this thesis, namely that if the plant has a particular canonical form, there exists a unique optimal compensator, and that an equivalent of the separation property holds for this problem.

We will completely solve the problem in this thesis using a direct approach. Namely, we will find the optimal compensator leading to the smallest increase in the cost. Our results are more general than those of [7] because the proposed method can treat any sort of plant and compensator initial states, because a final answer is also given in the case where the initial output covariance matrix of the plant is singular, and most of all because the class of all optimal compensators for an arbitrary plant is

precisely determined.

We will give now a brief survey of the different chapters.

In Chapter II, the linear regulator problem is revisited, and the effect of a similarity transformation on the plant states is considered. Since the main part of this thesis will consider a particular canonical form for the plant, it is important to know how to go from an arbitrary plant to this canonical plant, and to know what the optimal design for this new plant becomes. Also a brief review of observer theory is given.

In Chapter III, the design philosophy we have taken is exposed. One of the most important original parts of this chapter is the derivation of the initial error covariance matrix for the observer. This is truly a very important step towards a concise formulation of the problem. Another important part (although not original) is the computation of the total cost, which is shown to be composed of the two parts we previously mentioned. Particularly, the increase in the cost due to the need to reconstruct part of the state is clearly expressed. The problem is simply now to minimize this increase in the cost.

In Chapter IV, the main result is stated and proved. The technique for the proof involves gradient matrices. The optimal compensator is shown to be unique, due to the canonical form we have taken for the plant. Different cases are considered, depending on the non-singularity or singularity of the initial output covariance matrix. Numerous examples are provided as well as generalizations to treat different possible plant and compensator initial states.

In Chapter V, we return to the original arbitrary plant. Namely, we precisely determine the class of all compensators which are optimal for

the given plant   (all of them leading to the same increase in the cost, and the same dynamics for the overall system plant + compensator).

In Chapter VI, a possible extension of our results towards a compensator of lower order s (s < n − m) is presented.

CHAPTER II

PRELIMINARY BACKGROUND

2.1 Optimal Control and Similarity Transformations

In this paragraph we shall derive some results about similarity

transformations of the linear time-invariant, completely controllable and

completely observable system:

$$\dot{\bar{x}}(t) = \bar{A}\bar{x}(t) + \bar{B}u(t); \quad \bar{y}(t) = \bar{C}\bar{x}(t); \quad \bar{x}(0) = \bar{x}_o \tag{2.1.1}$$

where $\bar{x}$ (an n-vector) is the state of the plant,

u (an r-vector) is the control vector,

$\bar{y}$ (an m-vector) is the output of the plant.

The matrices $\bar{A}$, $\bar{B}$, $\bar{C}$ have respective dimensions (nxn), (nxr), (mxn),

and we will assume that $\bar{C}$ has full rank, (i.e., Rank $\bar{C}$ = m).

Assume the initial state of the plant is distributed according to:

$$E\{\bar{x}_o\} = 0; \quad E\{\bar{x}_o \bar{x}_o'\} = \bar{\Sigma}_o \tag{2.1.2}$$

where $\bar{\Sigma}_o$ is a positive semi-definite matrix.

Due to the stochastic initial state, consider the following cost:

$$\hat{\bar{J}} = E\left\{\int_0^\infty \{\bar{x}'(t)\ \bar{Q}\ \bar{x}(t) + u'(t)\ \bar{R}\ u(t)\}\ dt\right\} \tag{2.1.3}$$

where $\bar{Q}$ is an nxn positive semi-definite matrix, and $\bar{R}$ is an rxr positive

definite matrix. Moreover, assume $(\bar{A}, \bar{Q}^{1/2})$ is observable.

With these conditions, it is well known [1], that the optimal complete

state control for the cost (2.1.3) and the plant (2.1.1), is given by:

$$u*(t) = -\bar{R}^{-1}\bar{B}' \; \bar{\Pi} \; \bar{x}(t) = \bar{K}*\bar{x}(t) \qquad (2.1.4)$$

where $\bar{\Pi}$ is solution of the Riccati equation:

$$\bar{\Pi}\bar{A} + \bar{A}'\bar{\Pi} + \bar{Q} - \bar{\Pi}\bar{B}\bar{R}^{-1}\bar{B}'\bar{\Pi} = 0 \qquad (2.1.5)$$

and where we have defined:

$$\bar{K}* = -\bar{R}^{-1}\bar{B}'\bar{\Pi} \qquad (2.1.6)$$

The expected optimal cost is in this case:

$$\hat{\bar{J}}* = \text{tr}(\bar{\Pi}\bar{\Sigma}_o) \qquad (2.1.7)$$

whereas the closed-loop dynamics are governed by:

$$\dot{\bar{x}}(t) = (\bar{A} + \bar{B}\bar{K}*) \; \bar{x}(t); \; \bar{x}(0) = \bar{x}_o \qquad (2.1.8)$$

$\bar{x}_o$ being distributed as in (2.1.2).

Note that even if $E\{\bar{x}_o\} = \hat{\bar{x}}_o$ and $\text{cov}(\bar{x}_o, \bar{x}_o) = \bar{\Sigma}_o$, then the optimal control is still given by (2.1.4). However the optimal cost becomes in this case:

$$\hat{\bar{J}}* = \text{tr} \; \bar{\Pi}(\hat{\bar{x}}_o \; \hat{\bar{x}}_o' + \bar{\Sigma}_o) \qquad (2.1.9)$$

Now for the main part in this thesis, we will be interested in using a particular canonical form for the plant (2.1.1), in which the output matrix has the following form:

$$C = [\; I_{m \times m} \quad O_{m \times (n-m)}] \qquad (2.1.10)$$

This is absolutely no loss of generality because having assumed that Rank $\bar{C} = m$, one can always find an nxn non-singular matrix P such that $C = \bar{C}P^{-1}$. However, the original system is not very often given in this particular canonical form, and it is interesting to be able to transform an optimization problem given $\bar{Q}$, $\bar{R}$, $\bar{A}$, $\bar{B}$, $\bar{C}$ into an optimization problem given Q, R, A, B, C such that the dynamics of the two closed-loop optimal systems are identical. The problem is then to find P such that $C = \bar{C}P^{-1}$, where C is given by (2.1.10), and to relate Q and R to $\bar{Q}$ and $\bar{R}$.

A - Construction of P

Write $\quad \bar{C} = \begin{bmatrix} \bar{c}_1 \\ \cdots \\ \bar{c}_i \\ \cdots \\ \bar{c}_m \end{bmatrix}$ From the condition: Rank $\bar{C} = m$,

the m n-dimensional row vectors $\bar{c}_i$, $i = 1, \ldots m$, are linearly independent.

Define now the new state-variables:

$$x_i(t) = \bar{c}_i\,\bar{x}(t), \ i = 1, \ldots m; \ x_j(t) = t_j\,\bar{x}(t), \ j = m + 1, \ldots n \tag{2.1.11}$$

where the t_j's are n-m n-dimensional row vectors, such that the $\bar{c}_i$'s and t_j's are linearly independent and form a basis of R^n.

Write $\quad T = \begin{bmatrix} t_{m+1} \\ \cdots \\ t_j \\ \cdots \\ t_n \end{bmatrix}$; then (2.1.11) can be written:

$$x(t) = \begin{bmatrix} \bar{C} \\ T \end{bmatrix} \bar{x}(t) = P\,\bar{x}(t) \tag{2.1.12}$$

By construction, the output relation for the plant (2.1.1) is now:

$$\bar{y}(t) = y(t) = C \ x(t) = \bar{C}P^{-1} \ x(t) \qquad (2.1.13)$$

where C is given by (2.1.10) and P by (2.1.12).

Note that the matrix P is not unique, as T can be chosen arbitrarily, as long as the $\bar{c}_i$'s and t_j's are linearly independent and form a basis of R^n.

Under the similarity transformation P, the system (2.1.1) now becomes:

$$\dot{x}(t) = A \ x(t) + B \ u(t); \ y(t) = C \ x(t); \ x(0) = x_o \qquad (2.1.14)$$

$$E \ \{ \ x_o \ \} = 0 \ ; \ E \ \{ \ x_o \ x_o' \} = \Sigma_o \qquad (2.1.15)$$

where we have the following relations:

$$A = P\bar{A}P^{-1}; \ B = P\bar{B}; \ C = \bar{C}P^{-1}; \ x_o = P\bar{x}_o; \ \Sigma_o = P\bar{\Sigma}_oP' \qquad (2.1.16)$$

B – Relations between Q, R, $\bar{Q}$, $\bar{R}$.

For the new system (2.1.14), we would like to reformulate the optimization problem we solved for the plant (2.1.1) associated with the cost (2.1.3). In other words, we would like to find the matrices Q and R (respectively positive semi-definite and positive definite), such that the cost:

$$\hat{J} = E \ \left\{ \ \int_o^\infty \ \{x'(t) \ Q \ x(t) + u'(t) \ R \ u(t)\} \ dt \ \right\} \qquad (2.1.17)$$

associated with the plant (2.1.14) leads to the same closed loop dynamics as those given by (2.1.8).

We will give now a sufficient condition for this to be true.

<u>Theorem</u>. Consider the two following plants related by the similarity

transformation P:

$$\dot{\bar{x}}(t) = \bar{A}\,\bar{x}(t) + \bar{B}\,u(t); \quad \bar{y}(t) = \bar{C}\,\bar{x}(t); \quad \bar{x}(0) = \bar{x}_o \qquad (2.1.1)$$

$$E\{\bar{x}_o\} = 0; \quad E\{\bar{x}_o\bar{x}_o'\} = \bar{\Sigma}_o \qquad (2.1.2)$$

associated with the cost functional:

$$\hat{\bar{J}} = E\left\{ \int_0^\infty \{\bar{x}'(t)\,\bar{Q}\,\bar{x}(t) + u'(t)\,\bar{R}\,u(t)\}\,dt \right\} \qquad (2.1.3)$$

and

$$\dot{x}(t) = A\,x(t) + B\,u(t); \quad y(t) = C\,x(t); \quad x(0) = x_o \qquad (2.1.14)$$

$$E\{x_o\} = 0 \; ; \quad E\{x_o x_o'\} = \Sigma_o \qquad (2.1.15)$$

where we have the relations:

$$A = P\bar{A}P^{-1}; B = P\bar{B}; \; C = \bar{C}P^{-1}; \; x_o = P\bar{x}_o; \; \Sigma_o = P\bar{\Sigma}_o P' \qquad (2.1.16)$$

associated with the cost functional:

$$\hat{J} = E\left\{ \int_0^\infty \{x'(t)\,Q\,x(t) + u'(t)\,R\,u(t)\}\,dt \right\} \qquad (2.1.17)$$

Then a sufficient condition for the plant (2.1.14, 15) associated

with the cost (2.1.17) to have the same optimal closed loop dynamics as

the plant (2.1.1, 2) associated with the cost (2.1.3), is that:

$$Q = (P')^{-1}\,\bar{Q}\,P^{-1} \text{and } R = \bar{R} \qquad (2.1.18)$$

Moreover, in this case:

$$\Pi = (P')^{-1} \bar{\Pi} P^{-1} \text{ and } K* = \bar{K}*P^{-1} \qquad (2.1.19)$$

where we have defined

$$K* = - R^{-1} B' \Pi \qquad (2.1.20)$$

and finally, the optimal expected costs are equal:

$$\hat{\bar{J}}* = \hat{J}* \qquad (2.1.21)$$

Proof. The proof is very simple and relies on the fact that a suffi-cient condition for $A - BR^{-1}B'\Pi$ and $\bar{A} - \bar{B}\bar{R}^{-1}\bar{B}'\bar{\Pi}$ to have the same eigenvalues is that:

$$A - BR^{-1}B'\Pi = P(\bar{A} - \bar{B}\bar{R}^{-1}\bar{B}'\bar{\Pi})P^{-1}$$

Now using (2.1.16), we can write:

$$A - BR^{-1}B'\Pi = P(\bar{A} - \bar{B}\bar{R}^{-1}\bar{B}'P'\Pi P)P^{-1}$$

so that an obvious choice is to take:

$$R = \bar{R} \quad \text{and} \quad \Pi = (P')^{-1} \bar{\Pi} P^{-1} \qquad (2.1.22)$$

Now $\bar{\Pi}$ is given by:

$$\bar{\Pi}\bar{A} + \bar{A}'\bar{\Pi} + \bar{Q} - \bar{\Pi}\bar{B}\bar{R}^{-1}\bar{B}'\bar{\Pi} = 0 \qquad (2.1.23)$$

so that plugging (2.1.22) in (2.1.23) and making use of (2.1.16) yields:

$$(P')^{-1}\bar{\Pi}\bar{A}P^{-1} + (P')^{-1} \bar{A}'\bar{\Pi}P^{-1} + Q - (P')^{-1} \bar{\Pi}\bar{B}\bar{R}^{-1}\bar{B}'\bar{\Pi}P^{-1} = 0 \qquad (2.1.24)$$

A left multiplication of (2.1.24) by P' and a right multiplication by P yields:

$$\bar{\bar{\Pi}}\bar{\bar{A}} + \bar{A}'\bar{\bar{\Pi}} + P'QP - \bar{\bar{\Pi}}\bar{B}\bar{R}^{-1}\bar{B}'\bar{\bar{\Pi}} = 0 \qquad (2.1.25)$$

Now a comparison of (2.1.5) and (2.1.25) leads one to choose:

$$Q = (P')^{-1}\bar{Q} \, P^{-1}$$

Now obviously $K^* = - R^{-1}B'\Pi$ can be rewritten as follows:
$K^* = - \bar{R}^{-1}\bar{B}'P'(P')^{-1}\bar{\bar{\Pi}}P^{-1}$, that is $K^* = \bar{K}^*P^{-1}$.

The equality of the two optimal costs is easily proved by expressing Π and Σ_o in terms of $\bar{\bar{\Pi}}$ and $\bar{\Sigma}_o$, making use of the trace identity tr (AB) = tr(BA).

$$\hat{J}^* = \text{tr}\,(\Pi\Sigma_o) = \text{tr}\,\{(P')^{-1}\bar{\bar{\Pi}}P^{-1} \, P\bar{\Sigma}_oP'\}$$

$$= \text{tr}\,\{(P')^{-1}\bar{\bar{\Pi}}\bar{\Sigma}_oP'\} = \text{tr}\,\{(P')^{-1} \, P'\bar{\bar{\Pi}}\bar{\Sigma}_o\}$$

$$= \text{tr}\,(\bar{\bar{\Pi}}\bar{\Sigma}_o) = \hat{\bar{J}}^*$$

<div align="right">Q. E. D.</div>

Remark. Note that we have only here a sufficient condition. This comes from the fact that a sufficient condition for A and B to have the same eigenvalues is that A and B be similar, (i.e., $A = TBT^{-1}$ where T is non-singular). The condition is obviously not necessary as can be easily seen by choosing $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$.

We have now a complete picture of the effect of a similarity transformation on an optimal control design using complete state feedback.

2.2 Observer Theory: Review of Results

We will now briefly state some results about observer theory to introduce our notations. For a detailed proof of these results, the reader is

referred to the tutorial paper by Luenberger, [2].

A - Complete State Observer

<u>Theorem.</u> Corresponding to an n-th order completely observable linear
time invariant plant, an identity observer can be constructed having arbi-
trary eigenvalues.

Consider the linear time-invariant dynamical system:

$$\dot{x}(t) = A\, x(t) + B\, u(t); \quad y(t) = C\, x(t) \qquad\qquad (2.2.1)$$

where (A, C) is a completely observable pair.

We can define $\hat{x}(t)$, the estimate of $x(t)$, such that the error:

$$e(t) = x(t) - \hat{x}(t) \qquad\qquad (2.2.2)$$

has arbitrary dynamics. We will start from the dynamical system:

$$\dot{\hat{x}}(t) = F\, \hat{x}(t) + L\, u(t) + H\, y(t) \qquad\qquad (2.2.3)$$

and choose by comparison with (2.2.1)

$$F = A - HC \quad \text{and} \quad L = B \qquad\qquad (2.2.4)$$

where H is an nxm design matrix.

Using (2.2.4) and subtracting (2.2.3) from (2.2.1), we obtain:

$$\dot{e}(t) = (A - HC)\, e(t) \qquad\qquad (2.2.5)$$

and $\hat{x}(t)$ is governed by:

$$\dot{\hat{x}}(t) = A\, \hat{x}(t) + B\, u(t) + H[y(t) - C\, \hat{x}(t)] \qquad\qquad (2.2.6)$$

Now if (A, C) is a completely controllable pair, we see immediately,
[3], that H can be chosen such that A - HC has arbitrary eigenvalues.

A block diagram representation of the identity observer is shown in
Figure 1, page 23.

B - Minimal Order Luenberger Observer

There is a redundancy in the previous design, because if Rank C = m,
then we already know exactly m linear combinations of the state, and we
don't need to reconstruct all the n states.

Theorem. Corresponding to an n-th order, completely observable, linear
time invariant plant, having m independent outputs, a state observer of
order n-m can be constructed having arbitrary eigenvalues.

We have already seen in paragraph 2.1, that if Rank $\bar{C}$ = m, then we can
always find a non-singular matrix P, such that after similarity transforma-
tion the system (2.1.1) becomes:

$$\dot{x}(t) = A\ x(t) + B\ u(t);\ y(t) = C\ x(t) \tag{2.2.7}$$

where C has the following structure:

$$C = [\ I_{mxm} \qquad O_{mx(n-m)}]$$

Define now $x \overset{\Delta}{=} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ where x_1 is an m-vector and x_2 is an (n-m)
vector. With this definition, $y(t) = C\ x(t)$ becomes $y(t) = x_1(t)$.

Now partition accordingly A and B to obtain:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad ; \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

where A_{11} is $m \times m$, A_{12} $m \times (n-m)$, A_{21} $(n-m) \times m$, A_{22} $(n-m) \times (n-m)$, B_1 $m \times r$ and B_2 $(n-m) \times r$.

(2.2.7) can then be written:

$$\dot{x}_1(t) = A_{11} x_1(t) + A_{12} x_2(t) + B_1 u(t)$$

$$\dot{x}_2(t) = A_{21} x_1(t) + A_{22} x_2(t) + B_2 u(t) \qquad (2.2.8)$$

$$y(t) = x_1(t)$$

Note that we perfectly know $x_1(t)$ so that all we need is to estimate $x_2(t)$. Rewrite (2.2.8) in the following form:

$$\dot{x}_2(t) = A_{22} x_2(t) + v(t); \; v(t) = A_{21} x_1(t) + B_2 u(t) \qquad (2.2.9)$$

$$y_2(t) = A_{12} x_2(t) \; ; \; y_2(t) = \dot{x}_1(t) - A_{11} x_1(t) - B_1 u(t) \qquad (2.2.10)$$

and consider $v(t)$ and $y_2(t)$ as new control and observation vectors for $x_2(t)$. Then we can apply the theory developed in A - to estimate $x_2(t)$.

Write $\dot{\hat{x}}_2(t) = F \hat{x}_2(t) + Lv(t) + Hy_2(t)$, and choose $F = A_{22} - HA_{12}$, $L = I$ where H is $(n-m) \times m$.

We obtain:

$$\dot{\hat{x}}_2(t) = A_{22} \hat{x}_2(t) + v(t) + HA_{12}[x_2(t) - \hat{x}_2(t)] \qquad (2.2.11)$$

so that by subtracting (2.2.11) from (2.2.9) we have:

$$\dot{e}(t) = (A_{22} - HA_{12}) e(t) \qquad (2.2.12)$$

where

$$e(t) = x_2(t) - \hat{x}_2(t) \tag{2.2.13}$$

It can be easily shown that if (A, C) is completely observable, then so is (A_{22}, A_{12}). Then by appropriate choice of H, $e(t)$ can be assigned arbitrary dynamics.

Now from (2.2.10) we see that $y_2(t) = A_{12} \, x_2(t)$ contains a differentiation of the data $x_1(t)$. To avoid this, define:

$$z(t) = \hat{x}_2(t) - H \, x_1(t) \tag{2.2.14}$$

It is then easy to see that $z(t)$ satisfies:

$$\dot{z}(t) = F \, z(t) + G \, x_1(t) + D \, u(t) \tag{2.2.15}$$

where we have defined:

$$F = A_{22} - HA_{12}$$

$$G = FH + (A_{21} - HA_{11}) \tag{2.2.16}$$

$$D = B_2 - HB_1$$

A block diagram representation for this minimal order Luenberger observer is shown in Figure 2, page 23.

Important Remark. Note that this observer depends only on a single design matrix, namely H. The main goal of this thesis will be to optimally determine this matrix, in some precise sense.
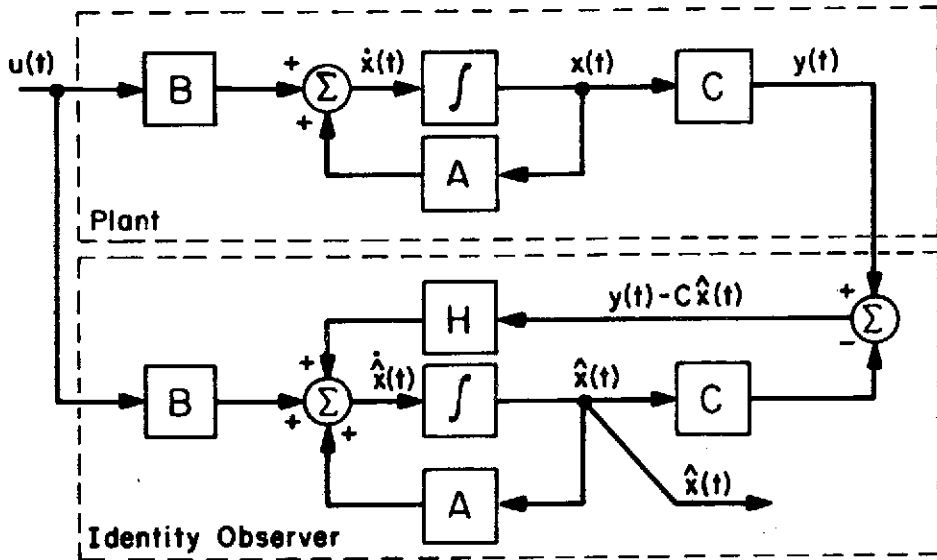
Fig. 1
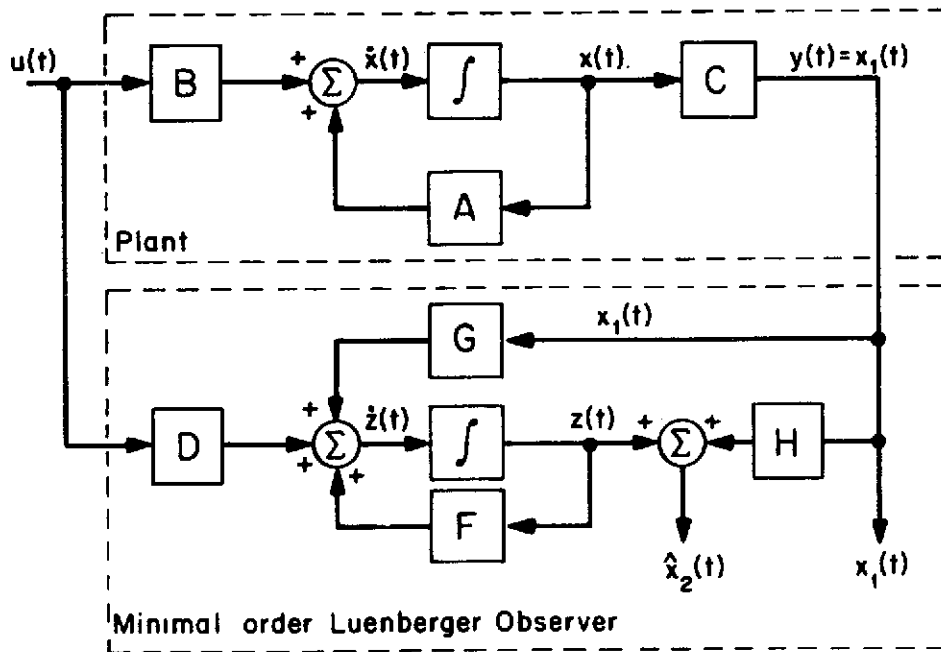
Fig. 2

CHAPTER III

DESIGN PHILOSOPHY: PRESENTATION OF THE APPROACH

3.1 Introduction

We have seen in Chapter II, paragraph 2.1, that one can easily
translate a complete state optimization problem, given a realization
$(\bar{A}, \bar{B}, \bar{C})$ of a plant and a cost involving $\bar{Q}$ and $\bar{R}$, into another equivalent
complete state optimization problem involving another realization (A, B, C)
of the same plant, associated with a cost involving new weighting matrices
Q and R.

If we use the similarity transformation $x(t) = P \bar{x}(t)$ between the two
realizations, then we showed that one should choose: $Q = (P')^{-1} \bar{Q} P^{-1}$ and
$\bar{R} = R$ in order to achieve the same design.

With this result in mind, we will concentrate for the remainder of this
chapter and Chapter IV on the following canonical plant:

$$\dot{x}(t) = A \, x(t) + B \, u(t); \quad y(t) = C \, x(t); \quad x(0) = x_o \qquad (3.1.1)$$

where x (an n-vector) is the state of the plant,

   y (an m-vector) is the output of the plant,

   u (an r-vector) is the control vector.

We will assume furthermore that x_o is a random variable specified by
its first and second moments:

$$E \{x_o\} = 0 \qquad E \{x_o x_o'\} = \Sigma_o \qquad (3.1.2)$$

In (3.1.1), A is an nxn matrix, B an nxr matrix, C an mxn matrix having the following canonical form:

$$C = [I_{mxm} \quad O_{mx(n-m)}] \qquad (3.1.3)$$

Assume also that (A, B) is a completely controllable pair, that (A, C) is a completely observable pair, but that the plant is not output stabilizable, so that a dynamic compensator is required if we have only access to y(t).

Assume now that we are given the cost functional:

$$\hat{J}_1 = E \left\{ \int_0^\infty \{x'(t) \ Q \ x(t) + u'(t) \ R \ u(t)\} \ dt \right\} \qquad (3.1.4)$$

where Q is an nxn positive semi-definite matrix, $(A, Q^{1/2})$ a completely observable pair, and R is an rxr positive definite matrix.

Supposing we have access to the complete state, the optimal control is then given by:

$$u^*(t) = K^* \ x(t) = - R^{-1} \ B'\Pi \ x(t) \qquad (3.1.5)$$

where Π satisfies the Riccati equation:

$$\Pi A + A'\Pi + Q - \Pi \ BR^{-1}B'\Pi = 0 \qquad (3.1.6)$$

The optimal closed-loop system is then:

$$\dot{x}(t) = (A + BK^*) \ x(t); \ y(t) = C \ x(t); \ x(0) = x_0 \qquad (3.1.7)$$

$$E \{x_0\} = 0 \qquad E \{x_0 x_0'\} = \Sigma_0 \qquad (3.1.2)$$

and the optimal cost for this design is:

$$\hat{J}_1^* = \text{tr}(\Pi \Sigma_o) \tag{3.1.8}$$

Note that from (3.1.7) we have:

$$x(t) = e^{(A + BK^*)t} x_o \tag{3.1.9}$$

and that, as a consequence, $\hat{J}_1^*$ can be rewritten:

$$\hat{J}_1^* = \text{tr} \left\{ \int_0^\infty e^{(A + BK^*)'t} (Q + K^{*'}RK^*) e^{(A + BK^*)t} dt\, \Sigma_o \right\}$$

so that Π satisfies also:

$$\Pi(A + BK^*) + (A + BK^*)'\Pi + Q + K^{*'} RK^* = 0 \tag{3.1.10}$$

Now in practice we don't have access to $x(t)$, but rather to $y(t)$. What we intend to do is:

- Use an observer to form $\hat{x}(t)$, an estimate of $x(t)$.

- Then apply the control $u(t) = K^* \hat{x}(t)$, where K^* is as in (3.1.5).

This approach has been taken by many authors [4], [5], [6], and assumes a sort of "separation property". However, up to now, no precise answer has been given to the question: what are the optimal dynamics to be chosen for the observer part of the compensator? We will solve completely the problem in this thesis, and show that a separation property effectively holds.

Important Remark: In [7] was recently given a partial answer to this question, assuming restricted possibilities of initial conditions for the plant and the compensator. The answer is derived, however, by identifying

the deterministic problem we are trying to solve with a stochastic steady-state control problem, and as such does not constitute a direct approach to the question. We will return later on the solution proposed by Newman, [5], [6].

3.2   Formulation of the Problem

Define as we did in Chapter II, paragraph 2.2,

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad ; \quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad ; \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \qquad (3.2.1)$$

(The reader is referred to the previous chapter for the dimensions of vectors and matrices and the derivation of the minimal order Luenberger observer.)

In view of (3.1.3) and (3.2.1), we have then $y(t) = C\,x(t) = x_1(t)$, so that a minimal order Luenberger observer can be implemented to estimate $x_2(t)$. We obtain:

$$\hat{x}_2(t) = z(t) + H\,x_1(t) \qquad (3.2.2)$$

$$\dot{z}(t) = F\,z(t) + G\,x_1(t) + D\,u(t) \qquad (3.2.3)$$

$$F = A_{22} - HA_{12}$$

$$G = FH + (A_{21} - HA_{11}) \qquad (3.2.4)$$

$$D = B_2 - HB_1$$

The error $e(t)$, defined as:

$$e(t) = x_2(t) - \hat{x}_2(t) \qquad (3.2.5)$$

satisfies the following differential equation:

$$\dot{e}(t) = F \ e(t) \qquad\qquad (3.2.6)$$

Define now:

$$K* \overset{\Delta}{=} [K_1^* \qquad K_2^*] \qquad\qquad (3.2.7)$$

where $K*$ is as in (3.1.5), K_1^* is rxm, and K_2^* rx(n-m). Noting that $\hat{x}(t)$, the estimate of x(t) is:

$$\hat{x}(t) = \begin{bmatrix} x_1(t) \\ \hat{x}_2(t) \end{bmatrix} \qquad\qquad (3.2.8)$$

we then want to implement the control $u(t) = K* \ \hat{x}(t)$, which becomes:

$$u(t) = K_1^* \ x_1(t) + K_2^* \ \hat{x}_2(t) \qquad\qquad (3.2.9)$$

that is using (3.2.2):

$$u(t) = (K_1^* + K_2^* \ H) \ x_1(t) + K_2^* \ z(t) \qquad\qquad (3.2.10)$$

We can now use this value for u(t) in the equation (3.2.3) to obtain:

$$\dot{z}(t) = \hat{F} \ z(t) + \hat{G} \ x_1(t) \qquad\qquad (3.2.11)$$

where we have defined

$$\hat{F} = F + DK_2^*$$
$$\hat{G} = G + D(K_1^* + K_2^* \ H) \qquad\qquad (3.2.12)$$

Assume for the present time that we start the compensator from $z(0) = 0$. We shall see in the sequel that this condition can be removed.

A block diagram of the overall design is shown in Figure 3, page 31, indicating clearly the two parts of the compensator. In Figure 4, page 31, a compact representation is shown, corresponding to the compensator we will practically implement.

Define now the augmented state:

$$\zeta(t) = \begin{bmatrix} x(t) \\ z(t) \end{bmatrix} \qquad (3.2.13)$$

then using (3.1.1, 2) and (3.2.10, 11) the overall closed loop system is given by:

$$\dot{\zeta}(t) = \Phi \, \zeta(t) \, ; \; \zeta(0) = \zeta_0 \qquad (3.2.14)$$

where we have defined:

$$\Phi = \begin{bmatrix} A + B(K_1^* + K_2^* H) & BK_2^* \\ \hat{G}C & \hat{F} \end{bmatrix} \qquad (3.2.15)$$

Now it is clear that since we had to reconstruct part of the state to achieve our control, we certainly incurred an increase in the cost. This increase can be computed by using $u(t)$ given by (3.2.10), in the cost (3.1.4), and rewriting this cost in terms of $\zeta(t)$.

Namely, (3.1.4) becomes:

$$\hat{\tilde{J}}_2(H) = E \left\{ \int_0^\infty \zeta'(t) \, \tilde{\Omega} \, \zeta(t) \, dt \right\} \qquad (3.2.16)$$

where we have defined:

$$\tilde{\Omega} = \begin{bmatrix} Q + C'(K_1^* + K_2^* H)'R\,(K_1^* + K_2^* H)\,C & C'(K_1^* + K_2^* H)'RK_2^* \\ K_2^{*'}R(K_1^* + K_2^* H)C & K_2^{*'}\,RK_2^* \end{bmatrix}$$

$$(3.2.17)$$

Now from (3.2.14), $\zeta(t)$ is given by

$$\zeta(t) = e^{\Phi t} \zeta_0 \qquad (3.2.18)$$

where $\zeta_0 = \begin{bmatrix} x_0 \\ 0 \end{bmatrix}$ since we assumed $z(0) = 0$. Defining now:

$$Z_0 = E\{\zeta_0 \zeta_0'\} = \begin{bmatrix} \Sigma_0 & 0 \\ 0 & 0 \end{bmatrix}, \text{ then:}$$

$$\hat{\tilde{J}}_2(H) = tr \left\{ \int_0^\infty e^{\Phi't} \tilde{\Omega} e^{\Phi t} dt \, Z_0 \right\} \qquad (3.2.19)$$

so that if $\tilde{\Lambda}$ is the solution of the following Lyapunov equation:

$$\tilde{\Lambda} \Phi + \Phi'\tilde{\Lambda} + \tilde{\Omega} = 0 \qquad (3.2.20)$$

then $\hat{\tilde{J}}_2(H)$ will be given by:

$$\hat{\tilde{J}}_2(H) = tr(\tilde{\Lambda} Z_0)$$

The increase in cost due to the use of a compensator will then be:

$$\Delta \hat{\tilde{J}}(H) = \hat{\tilde{J}}_2(H) - \hat{J}_1^* = tr(\tilde{\Lambda} Z_0) - tr(\Pi \Sigma_0)$$

where $\hat{J}_1^*$ is the optimal cost when we have complete state feedback.

We explicitly showed the dependence of $\hat{\tilde{J}}_2(H)$ on H. Our main goal will be to choose H so that $\Delta \hat{\tilde{J}}_2(H)$ is minimum. However, it turns out that even though the variables $x(t)$ and $z(t)$ are the natural variables to work with, as in practice we will implement $z(t)$, they are not suitable for actual pencil and paper calculations, as they lead to intractable equations. We shall then reformulate the problem in terms of $x(t)$ and the error $e(t) = x_2(t) - \hat{x}_2(t)$.
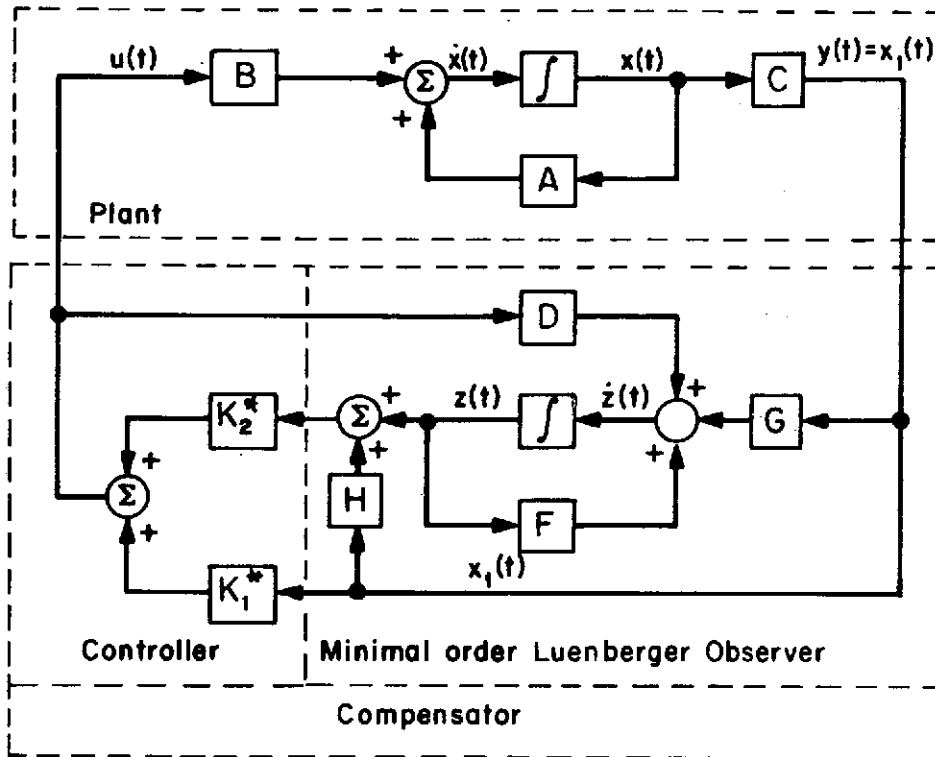
Fig. 3
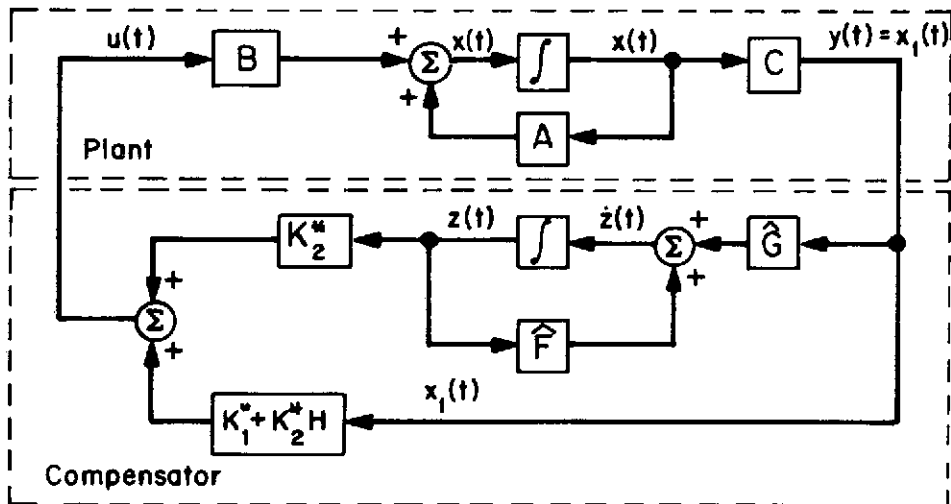
Fig. 4

3.3 Reformulation of the Problem

We will work now in terms of $x(t)$ and $e(t)$. Note that

$$\hat{x}(t) = \begin{bmatrix} x_1(t) \\ \hat{x}_2(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_2(t) - e(t) \end{bmatrix} = [I_{nxn} \quad -N] \begin{bmatrix} x(t) \\ e(t) \end{bmatrix}$$

where we defined:

$$N = \begin{bmatrix} 0_{mx(n-m)} \\ I_{(n-m)x(n-m)} \end{bmatrix} \qquad (3.3.1)$$

We can then rewrite the control $u(t) = K^* \hat{x}(t)$ as:

$$u(t) = K^* x(t) - K^* N e(t) \qquad (3.3.2)$$

Define now the new augmented state:

$$\xi(t) = \begin{bmatrix} x(t) \\ e(t) \end{bmatrix} \qquad (3.3.3)$$

Then we can make use of (3.1.1, 2), (3.2.6) and (3.3.2) to write the over-all closed-loop system in the following form:

$$\dot{\xi}(t) = \Gamma \, \xi(t); \quad \xi(0) = \xi_0 \qquad (3.3.4)$$

where Γ is given by:

$$\Gamma = \begin{bmatrix} A + BK^* & -BK^* N \\ 0 & F \end{bmatrix} \qquad (3.3.5)$$

We note that this procedure is a formal procedure, as obviously we cannot implement $u(t)$ given by (3.3.2), because we don't know $x(t)$ but only $y(t)$. However, we see immediately the theoretical advantage of using $x(t)$

and e(t): Γ given by (3.3.5) shows us clearly that the overall system will have the poles of the optimal closed-loop system assuming we have complete state feedback, associated with the poles relative to F. This wasn't clearly apparent in view of Φ given by (3.2.15).

Now as previously we can evaluate the increase in cost. Using u(t) given by (3.3.2), we can express $\hat{J}_2(H)$ in terms of $\xi(t)$ as:

$$\hat{J}_2(H) = E \left\{ \int_0^\infty \xi'(t) \, \Omega \, \xi(t) \, dt \right\} \qquad (3.3.6)$$

where we have defined:

$$\Omega = \begin{bmatrix} Q + K^{*\prime} \, RK^* & -K^{*\prime} \, RK^* \, N \\ -N' \, K^{*\prime} \, RK^* & N'K^{*\prime} \, RK^* \, N \end{bmatrix} \qquad (3.3.7)$$

Now from what we did previously, it follows from (3.3.4) that:

$$\xi(t) = e^{\Gamma t} \, \xi_0 \qquad (3.3.8)$$

where we have $\xi_0 = \begin{bmatrix} x_0 \\ e_0 \end{bmatrix}$.

Defining now $\Xi_0 = E \{\xi_0 \xi_0'\}$, $\hat{J}_2(H)$ can be rewritten:

$$\hat{J}_2(H) = \mathrm{tr} \left\{ \int_0^\infty e^{\Gamma' t} \, \Omega \, e^{\Gamma t} \, dt \, \Xi_0 \right\} \qquad (3.3.9)$$

so that if Λ is the solution of the Lyapunov equation:

$$\Lambda \, \Gamma + \Gamma'\Lambda + \Omega = 0 \qquad (3.3.10)$$

then $\hat{J}_2(H)$ will be given by:

$$\hat{J}_2(H) = \mathrm{tr}(\Lambda \, \Xi_0) \qquad (3.3.11)$$

We will now evaluate precisely $\hat{J}_2(H)$. We then need to evaluate Ξ_0 and Λ. For the computation of Λ, we will mainly follow Yuksel and Bongiorno, [4].

A – Evaluation of Ξ_0

Recall from (3.2.2) that $\hat{x}_2(t) = z(t) + H x_1(t)$, and from (3.2.5) that $e(t) = \hat{x}_2(t)$ so that $e(t)$ is given by:

$$e(t) = x_2(t) - H x_1(t) - z(t) \qquad (3.3.12)$$

Now x_0 is such that $E\{x_0\} = 0$, so that both $E\{x_{10}\} = 0$ and $E\{x_{20}\} = 0$. Since we have also assumed that the initial state of the compensator is zero $(z(0) = 0)$, we then obtain from (3.3.12)

$$E\{e_0\} = 0 \qquad (3.3.13)$$

so that

$$\Xi_0 = E\{e_0 e_0'\} = H\Sigma_{11}H' + \Sigma_{22} - H\Sigma_{12} - \Sigma_{12}' H' \qquad (3.3.14)$$

where we have partitioned Σ_0 as:

$$\Sigma_0 = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{bmatrix} \qquad (3.3.15)$$

Σ_{11}, Σ_{12}, Σ_{22} having the same dimensions as A_{11}, A_{12}, A_{22}.

Now $E\{x_0 e_0'\}$ will be given by:

$$\Xi_{12} = E\{x_0 e_0'\} = E\left\{ \begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix} e_0' \right\} = \begin{bmatrix} \Sigma_{12} - \Sigma_{11}H' \\ \Sigma_{22} - \Sigma_{12}'H' \end{bmatrix} \qquad (3.3.16)$$

so that we finally obtain:

$$\Xi_0 = E\{\xi_0 \xi_0'\} = \begin{bmatrix} \Sigma_0 & \Xi_{12} \\ \Xi_{12}' & E_0 \end{bmatrix} \tag{3.3.17}$$

where Σ_0 is the initial plant covariance matrix, E_0 is the initial error covariance matrix and Ξ_{12} is given by (3.3.16).

Important Remark. Note that E_0 is a quadratic function of H, and not a simple positive semi-definite matrix.

B – Evaluation of Λ

We will now evaluate Λ, the solution of the Lyapunov algebraic equation (3.3.10) where Ω is given by (3.3.7) and Γ by (3.3.5).

Partition Λ as:

$$\Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}' & \Lambda_{22} \end{bmatrix} \tag{3.3.18}$$

where Λ_{11} is (nxn), Λ_{12} is nx(n-m) and Λ_{22} is (n-m)x(n-m). Then the solution of (3.3.10) requires the solution of the three following equations:

$$\Lambda_{11}(A + BK^*) + (A + BK^*)' \Lambda_{11} + Q + K^{*'} RK^* = 0 \tag{3.3.19}$$

$$(A + BK^*)' \Lambda_{12} + \Lambda_{12}F - \Lambda_{11} BK^*N - K^{*'} RK^*N = 0 \tag{3.3.20}$$

$$\Lambda_{22}F + F' \Lambda_{22} - \Lambda_{12} BK^*N - N'K^{*'} B'\Lambda_{12} + N'K^{*'} RK^*N = 0 \tag{3.3.21}$$

Now by comparison of (3.3.19) and (3.1.10), we see immediately that $\Lambda_{11} = \Pi$. Using $\Lambda_{11} = \Pi$ and $K^* = - R^{-1} B'\Pi$ in (3.3.20) yields:

$$(A + BK^*)' \ \Lambda_{12} + \Lambda_{12}F = 0 \tag{3.3.22}$$

Now we know that $(A + BK^*)$ is stable and we desire a stable F, as F governs the dynamics of the error $e(t)$. We then conclude that $A + BK^*$ and F are non singular, which implies $\Lambda_{12} = 0$.

Using now $\Lambda_{11} = \Pi$ and $\Lambda_{12} = 0$ in (3.3.21) yields:

$$\Lambda_{22}F + F' \ \Lambda_{22} + N'K^{*'} \ RK^*N = 0 \tag{3.3.23}$$

which has a unique positive definite solution, still because of the required stability of F.

Finally, we see that Λ becomes:

$$\Lambda = \begin{bmatrix} \Pi & 0 \\ 0 & \Lambda_{22} \end{bmatrix} \tag{3.3.24}$$

so that using (3.3.17) for Ξ_o, the cost $\hat{J}_2(H)$ can be rewritten:

$$\hat{J}_2(H) = tr(\Lambda \Xi_o) = tr(\Pi \Sigma_o) + tr(\Lambda_{22}E_o) \tag{3.3.25}$$

where Λ_{22} is the solution of (3.3.23) and E_o is given by (3.3.14). Recalling that $\hat{J}_1^* = tr(\Pi \Sigma_o)$ we see immediately that:

$$\Delta\hat{J}_2(H) = \hat{J}_2(H) - \hat{J}_1^* = tr(\Lambda_{22}E_o) \tag{3.3.26}$$

Note that Λ_{22} depends on H through F, and that E_o is a quadratic function of H. Define:

$$M \triangleq N'K^{*'} \ RK^*N \tag{3.3.27}$$

which is independent of H, the problem on hand is then the following:

given $\Delta \hat{J}_2(H) = tr(\Lambda_{22}E_o)$, where:

$$\Lambda_{22}F + F'\Lambda_{22} + M = 0 \qquad (3.3.28)$$

$$F = A_{22} - HA_{12} \text{ and } E_o(H) = H\Sigma_{11}H' + \Sigma_{22} - H\Sigma_{12} - \Sigma'_{12}H' \qquad (3.3.29)$$

find H* such that $\forall H \neq H*$ then

$$\Delta \hat{J}_2(H*) \leq \Delta \hat{J}_2(H)$$

Remarks

1. The approach we have taken here, considering the variables x(t) and e(t) is not original. In particular, Yuksel and Bongiorno [4] showed also that $\Delta \hat{J}_2(H) = tr(\Lambda_{22}E_o)$. However, up to now people have considered E_o as a positive semi-definite matrix independent of H, and consequently have tried to minimize simply $tr(\Lambda_{22})$. Now from (3.3.28), we see immediately that this philosophy leads one to choose F having greater and greater negative eigenvalues. Some authors then asserted that $\Delta \hat{J}_2(H)$ will go to zero as the eigenvalues of F are pushed towards infinity. It was soon shown by counterexamples [8], [9], [10] that this was obviously wrong in some cases. However, no precise solution was given to find the optimal H*.

2. Newman in [5], [6] uses the natural variables x(t) and z(t) to tackle the problem. However, he does not take his plant in the canonical form we have chosen. As a consequence, he ends up with a formidable set of non-linear matrix equations which are not readily solvable, and his conclusion is that the best control is given by $u(t) = K_o\hat{x}(t)$ where $K_o \neq K*$. This conclusion is wrong, and his error was also recently noticed and

corrected by Rom and Sarachik, [7].

3. What we simply noticed is that E_o is in fact a quadratic function of H, and that instead of trying to reduce the "size" of Λ_{22}, we must minimize the function:

$$\Delta \hat{J}_2(H) = tr \quad \{\Lambda_{22}(H) \; E_o(H)\}$$

The problem is now well posed and we can give a precise answer, leading to an optimal design with minimum increase in the cost.

CHAPTER IV

THE OPTIMAL COMPENSATOR

4.1   Statement of the Result

We want now to solve the following minimization problem:

Given Λ_{22} solution of the following Lyapunov equation:

$$\Lambda_{22}F + F'\Lambda_{22} + M = 0 \qquad\qquad (4.1.1)$$

where:

$$F = A_{22} - HA_{12} \qquad\qquad (4.1.2)$$

find H* such that $\bigvee$ H $\neq$ H*

$$\Delta\hat{J}_2(H^*) \leq \Delta\hat{J}_2(H)$$

where

$$\Delta\hat{J}_2(H) = tr(\Lambda_{22}E_o) \qquad\qquad (4.1.3)$$

and E_o is given by:

$$E_o(H) = H\Sigma_{11}H' + \Sigma_{22} - H\Sigma_{12} - \Sigma'_{12}H' \qquad\qquad (4.1.4)$$

Recall that $E_o(H)$ is positive semi-definite, being the initial covariance

matrix of the error.

We will now state the following theorem.

Theorem. 1. Assume that the initial output covariance matrix Σ_{11}

is non-singular. Then there exists a unique minimizing H for the incremental cost (4.1.3), namely:

$$H^* = (WA'_{12} + \Sigma'_{12}) \Sigma_{11}^{-1}$$

where W is the unique positive definite solution of the Riccati-type equation:

$$A_{22}W + WA'_{22} + \Sigma_{22} - (WA'_{12} + \Sigma'_{12}) \Sigma_{11}^{-1} (WA'_{12} + \Sigma'_{12})' = 0$$

2. Assume now that Σ_{11} is singular: then no minimizing H exists, and the incremental cost (4.1.3) can be made arbitrarily small by appropriate choice of H.

We will give now some preliminary results that we will need to prove the theorem.

As we are looking for a stable F, we can write from (4.1.1):

$$\Lambda_{22} = \int_0^\infty e^{(A_{22} - HA_{12})'t} M e^{(A_{22} - HA_{12})t} dt \qquad (4.1.5)$$

Recall that M does not depend on H. The cost $\Delta \hat{J}_2(H)$ can now be rewritten, using the trace identity tr(AB) = tr(BA):

$$\Delta \hat{J}_2(H) = tr \left\{ M \int_0^\infty e^{(A_{22} - HA_{12})t} E_o(H) e^{(A_{22} - HA_{12})'t} dt \right\} \qquad (4.1.6)$$

Note that $\Delta \hat{J}_2(H)$ is a real function of (n-m)xm variables (i.e., the h_{ij}'s, where H = $[h_{ij}]$). A necessary condition for H* to minimize such a function is that $\frac{d}{dH} \Delta \hat{J}_2(H) \Big|_{H^*} = 0$. We shall simply calculate and evaluate the derivative.

A key lemma in this calculation is the following due to Kleinman, [11].

Definition: $g(.)$ is a trace function of the matrix X, if $g(X)$ is of the form:

$$g(X) = tr[G(X)]$$

where $G(.)$ is a continuously differentiable mapping from the space of rxn matrices into the space of nxn matrices.

Lemma: Let $g(X)$ be a trace function. Then if we can write:

$$g(X + \varepsilon \Delta X) - g(X) = \varepsilon \ tr[Q(X) \ \Delta X]$$

as $\varepsilon \rightarrow 0$, where $Q(X)$ is an nxr matrix, X is an rxn matrix, then we have:

$$\frac{d}{dX} g(X) = Q'(X)$$

The calculations we will do follow closely those of Levine, [12]. But let's first show a simple application of this lemma.

Example: Let $G(H) = e^{(A_{22} - HA_{12})t}$. We want to compute:

$$\frac{d}{dH} g(H) = \frac{d}{dH} \left\{ tr[G(H)] \right\}$$

We have

$$G(H + \varepsilon \ \Delta H) = e^{(A_{22} - HA_{12} - \varepsilon \Delta HA_{12})t} \tag{4.1.7}$$

But from p. 175, reference [13], we have that to the first order in ε:

$$G(H + \varepsilon \ \Delta H) = e^{(A_{22} - HA_{12})t} - \varepsilon \int_0^t e^{(A_{22} - HA_{12})(t - \sigma)}$$

$$\Delta HA_{12} \ e^{(A_{22} - HA_{12})\sigma} \ d\sigma \tag{4.1.8}$$

Hence $\dot{\alpha}(\Delta H) = G(H + \varepsilon\Delta H) - G(H)$ is equal to:

$$\alpha(\Delta H) = \int_0^t e^{(A_{22} - HA_{12})(t - \sigma)} \Delta HA_{12} \, e^{(A_{22} - HA_{12})\sigma} \, d\sigma$$

Now since the trace operation commutes with integration, and since $tr(AB) = tr(BA)$, we obtain:

$$tr[\alpha(\Delta H)] = tr\left\{ \int_0^t A_{12} \, e^{(A_{22} - HA_{12})\sigma} \, e^{(A_{22} - HA_{12})(t - \sigma)} \, d\sigma\Delta H \right\}$$

that is:

$$tr[\alpha(\Delta H)] = tr\left\{ A_{12} \, e^{(A_{22} - HA_{12})t} \, \Delta H \right\}$$

therefore:

$$\frac{d}{dH} tr\left\{ e^{(A_{22} - HA_{12})t} \right\} = e^{(A_{22} - HA_{12})'t} \, A'_{12}$$

We can now proceed through the proof of the theorem, with the lemma and example for guidance. Statement 1 is proved in Section 4.2, and Statement 2 of the theorem is proved in Section 4.6.

4.2 Proof of the Result. Σ_{11} non-singular.

Recall from (4.1.6):

$$\Delta\hat{J}_2(H) = tr\left\{ M \int_0^\infty e^{(A_{22} - HA_{12})t} E_o(H) e^{(A_{22} - HA_{12})'t} dt \right\} \quad (4.2.1)$$

where:

$$E_o(H) = H\Sigma_{11} H' + \Sigma_{22} - H\Sigma_{12} - \Sigma'_{12} H' \quad (4.2.2)$$

We want to compute the quantity $\Delta\hat{J}_2(H + \varepsilon\Delta H) - \Delta\hat{J}_2(H)$ in order to apply Kleinman's Lemma.

From (4.2.2) we have to first order in ε:

$$E_o(H + \varepsilon\Delta H) = E_o(H) + \varepsilon\Delta H(\Sigma_{11}H' - \Sigma_{12}) + \varepsilon(\Sigma_{11}H' - \Sigma_{12})' \Delta H' \quad (4.2.3)$$

and from (4.1.8), also to the first order in ε:

$$e^{(A_{22} - HA_{12} - \varepsilon\Delta HA_{12})t} = e^{Ft} - \varepsilon \int_0^t e^{F(t - \sigma)} \Delta HA_{12} e^{F\sigma} d\sigma \quad (4.2.4)$$

so that $\Delta\hat{J}_2(H + \varepsilon\Delta H)$ is given by:

$$\Delta\hat{J}_2(H + \varepsilon\Delta H) = tr\left\{ M \int_0^\infty e^{(F - \varepsilon\Delta HA_{12})t} E_o(H + \varepsilon\Delta H) e^{(F - \varepsilon\Delta HA_{12})'t} dt \right\} \quad (4.2.5)$$

Using (4.2.3) and (4.2.4) we can now rewrite (4.2.5) retaining only the first order terms in ε:

$$\Delta\hat{J}_2(H + \varepsilon\Delta H) = tr\left\{ M \int_0^\infty \left[e^{Ft}E_o(H) e^{F't} + \varepsilon e^{Ft} \Delta H(\Sigma_{11}H' - \Sigma_{12})e^{F't} \right.\right.$$

$$+ \varepsilon e^{Ft}(\Sigma_{11}H' - \Sigma_{12})' \Delta H' e^{F't} - \varepsilon \int_0^t e^{F(t - \sigma)}\Delta HA_{12}e^{F\sigma}d\sigma E_o(H)e^{F't}$$

$$\left.\left. - \varepsilon e^{Ft} E_o(H) \int_0^t e^{F'(t - \sigma)} A'_{12} \Delta H' e^{F'\sigma} d\sigma \right]dt\right\} \quad (4.2.6)$$

Now using $tr(AB) = tr(BA)$ and subtracting (4.2.1) from (4.2.6), we obtain:

$$\Delta\hat{J}_2(H + \varepsilon\Delta H) - \hat{J}_2(H) = \varepsilon\ tr\left\{ M \int_0^\infty \left[2\ (\Sigma_{11}H' - \Sigma_{12})\ e^{F't}\ e^{Ft}\ \Delta H \right.\right.$$

$$- \int_0^t A_{12}\ e^{F\sigma}\ E_0(H)\ e^{F't}\ e^{F(t-\sigma)}\ d\sigma\Delta H -$$

$$\left.\left. - \int_0^t A_{12}\ e^{F(t-\sigma)}\ E_0(H)\ e^{F't}\ e^{F\sigma}\ d\sigma\ \Delta H \right]\ dt \right\}$$

So that we have by application of the lemma:

$$\frac{d}{dH}\ \Delta\hat{J}_2(H) = \left\{ 2 \int_0^\infty e^{F't}\ e^{Ft}\ (\Sigma_{11}H' - \Sigma_{12})'\ dt\right.$$

$$- \int_0^\infty dt \int_0^t d\sigma\ e^{F'(t-\sigma)}\ e^{Ft}\ E_0(H)\ e^{F'\sigma}\ A'_{12}$$

$$\left. - \int_0^\infty dt \int_0^t d\sigma\ e^{F'\sigma}\ e^{Ft}\ E_0(H)\ e^{F'(t-\sigma)}\ A'_{12} \right\}\ M \qquad (4.2.7)$$

Now consider

$$\Theta = \int_0^\infty dt \int_0^t d\sigma\ e^{F'\sigma}\ e^{Ft}\ E_0(H)\ e^{F'(t-\sigma)}$$

and make the following change of variables:

$$\sigma = t - \sigma_1;\ d\sigma = -d\sigma_1;\ \sigma = 0,\ \sigma_1 = t;\ \sigma = t,\ \sigma_1 = 0$$

then Θ becomes $\Theta = \displaystyle\int_0^\infty dt \int_0^t d\sigma_1\ e^{F'(t-\sigma_1)}\ e^{Ft}\ E_0(H)\ e^{F'\sigma_1}$

So that (4.2.7) now becomes:

$$\frac{d}{dH} \, \Delta\hat{J}_2 (H) = \left\{ 2\int_0^\infty e^{F't} \, e^{Ft} (\Sigma_{11}H' - \Sigma_{12})' \, dt \right.$$

$$\left. - \int_0^\infty dt \int_0^t d\sigma \, e^{F'\sigma} \, e^{Ft} \, E_0(H) \, e^{F'(t-\sigma)} A'_{12} \right\} M \qquad (4.2.8)$$

Consider now $\chi = \displaystyle\int_0^\infty dt \int_0^t d\sigma \, e^{F'\sigma} \, e^{Ft} \, E_0(H) \, e^{F'(t-\sigma)}$

we will interchange the order of integrations in χ to obtain:

$$\chi = \int_0^\infty d\sigma \int_\sigma^\infty dt \, e^{F'\sigma} \, e^{Ft} \, E_0(H) \, e^{F'(t-\sigma)}$$

We will now make the change of variables:

$$\tau = t - \sigma; \quad d\tau = dt; \quad t = \sigma, \ \tau = 0; \quad t = \infty, \ \tau = \infty$$

$$\chi = \int_0^\infty d\sigma \int_0^\infty d\tau \, e^{F'\sigma} \, e^{F\sigma} \, e^{F\tau} \, E_0(H) \, e^{F'\tau}$$

or finally

$$\chi = \int_0^\infty e^{F'\sigma} \, e^{F\sigma} \, d\sigma \int_0^\infty e^{F\tau} \, E_0(H) \, e^{F'\tau} \, d\tau$$

so that (4.2.8) can be rewritten:

$$\frac{d}{dH} \, \Delta\hat{J}_2 (H) = 2 \int_0^\infty e^{F't} e^{Ft} dt \left[(\Sigma_{11}H' - \Sigma_{12})' - \int_0^\infty e^{F\tau} \, E_0(H) \, e^{F'\tau} \, d\tau \, A'_{12} \right] M$$

We will now set $\dfrac{d}{dH}\, \Delta\hat{J}_2\,(H) = 0$, to obtain

$$H\Sigma_{11} - \Sigma'_{12} = \int_0^\infty e^{F\tau}\, E_O\,(H)\, e^{F'\tau}\, d\tau\, A'_{12} \tag{4.2.9}$$

Define now

$$W \triangleq \int_0^\infty e^{F\tau}\, E_O\,(H)\, e^{F'\tau}\, d\tau \tag{4.2.10}$$

Then if we assume that Σ_{11} is non-singular, (4.2.9) can be rewritten:

$$H^* = (W\, A'_{12} + \Sigma'_{12})\, \Sigma_{11}^{-1} \tag{4.2.11}$$

We see from (4.2.10) that W satisfies the algebraic Lyapunov equation:

$$FW + WF' + E_O\,(H) = 0 \tag{4.2.12}$$

Now recalling that $E_O\,(H) = H\Sigma_{11}H' + \Sigma_{22} - H\Sigma_{12} - \Sigma'_{12}H'$, that $F = A_{22} - HA_{12}$, and making use of (4.2.11), we can rewrite (4.2.12) as:

$$A_{22}W + WA'_{22} + \Sigma_{22} - (WA'_{12} + \Sigma'_{12})\, \Sigma_{11}^{-1}\, (WA'_{12} + \Sigma'_{12})' = 0 \tag{4.2.13}$$

Note that $E_O\,(H)$ is positive semi-definite, and that we are looking for a stable F, so that W, the solution of the Lyapunov equation (4.2.12), is positive definite, and unique from the linearity of the Lyapunov equation. Equivalently, W is the unique positive definite solution of the Riccati-type equation (4.2.13).

This uniqueness of H^* is of prime importance, as it insures, that given A, B, $C = [I_{mxm} \quad 0_{mx(n-m)}]$ and the cost (3.1.4), there is only one optimal compensator for our problem and that its realization is completely determined by Eq. (3.2.10, 11). We will see in Chapter V that when the output matrix C of the plant does not have this canonical form, then a unique optimal compensator no longer exists, but rather a class of optimal compensators (all of them leading to the same increase in the cost) can be precisely determined.

We can compute now the minimum increase in the cost, which will be given by:

$$\Delta \hat{J}_2^* = tr(\Lambda_{22}^* \ E_o^*) \tag{4.2.14}$$

where Λ_{22}^* is the solution of

$$\Lambda_{22} \ F^* + F^{*'} \ \Lambda_{22} + M = 0 \tag{4.2.15}$$

in which $F^* = A_{22} - H^* A_{12}$ and where E_o^* in (4.2.14) is given by:

$$E_o^* = H^* \Sigma_{11} \ H^{*'} + \Sigma_{22} - H^* \Sigma_{12} - \Sigma_{12}' \ H^{*'} \tag{4.2.16}$$

Important Remarks. 1. Note that H^*, through W and (4.2.11) depends uniquely on the initial state covariance matrix Σ_o, (i.e., Σ_{11}, Σ_{12}, Σ_{22}). We will see clearly with an example, that if we take two different initial state covariance matrices, one "small" and another "large", reflecting two different degrees of knowledge of the initial state, then $F^* = A_{22} - H^* A_{12}$, which governs the dynamics of the error $x_2(t) - \hat{x}_2(t)$, will have

much faster dynamics when we take a large Σ_o. This corresponds to what one would expect, the initial error (large in this case) being quickly reduced to allow the observer to still track closely $x_2(t)$ and achieve a good control.

2. The independence of H* with respect to the closed loop dynamics of the optimal system assuming complete state feedback, justifies completely the sort of separation property we assumed in our design, namely: estimate the state, and then apply the control $u(t) = K*\hat{x}(t)$.

In other words, we can change Q and R in the cost (3.1.4) so as to have a better and better regulation of the state, but if we still assume the same initial state covariance matrix for the plant, then H* will be the same for all designs. The estimator-observer part of the design is independent of the control part, showing clearly that an equivalent of the separation property holds for this problem.

3. Note that there may be solutions of (4.2.12) which are not solutions (4.2.10). We then need to show that there exists a stabilizing H for F = $A_{22} - HA_{12}$; if this is true then W will be the unique positive definite solution of (4.2.13).

We will simply compare our problem with a filtering problem whose solution is known. Namely, consider the following observable, stochastic linear system:

$$\dot{x}_2(t) = A_{22} x_2(t) + \xi(t)$$

$$z(t) = A_{12} x_2(t) + \eta(t)$$

where $\xi(t)$ and $\eta(t)$ are Gaussian processes such that:

$$E\{\xi(t)\} = 0 \qquad E\{\eta(t)\} = 0$$

$$E\{\xi(t)\xi'(\tau)\} = \Sigma_{22}\ \delta(t-\tau);\ E\{\eta(t)\eta'(\tau)\} = \Sigma_{11}\ \delta(t-\tau)$$

$$E\{\eta(t)\xi'(\tau)\} = \Sigma_{12}\ \delta(t-\tau)$$

for all t and τ, and where $\Sigma_{11} > 0$ and $\Sigma_{22} \geq 0$.

Then it can be readily seen that (4.2.13) and (4.2.11) are precisely the equations for the error covariance matrix and gain of the steady-state Kalman filter we would implement to obtain an estimate of $x_2(t)$ from the noisy measurement $z(t)$.

Now, the dynamics of this steady-state Kalman filter would be:

$$\dot{\hat{x}}_2(t) = (A_{22} - HA_{12})\ \hat{x}_2(t) + HA_{12}\ z(t)$$

and it is well known that with the assumptions we have made about Σ_{11}, Σ_{22}, and the observability of the pair $(A_{22},\ A_{12})$, the matrix $F = A_{22} - HA_{12}$ governing the dynamics of the filter is stable ([30], [31]).

We can now return to our problem and conclude that there exists a unique minimizing H given by (4.2.11). This H will lead to a stable $F = A_{22} - HA_{12}$ and consequently insure that (4.2.10) and (4.2.12) have the same solution, if we take for W the unique positive definite solution of (4.2.13).

4.3 Generalizations - Σ_{11} Non-Singular

Up to now, we have considered only the initial conditions: $E\{x_o\} = 0$,

$E\{x_o x_o'\} = \Sigma_o$, $z_o = 0$ (see Chapter III, Paragraphs 3.2 and 3.3). Nevertheless, it may very well happen that we know that x_o is always disturbed in

a particular way so that $E\{x_o\} = \bar{x}_o$, $(\bar{x}_o \neq 0)$. It may also happen that

for a practical reason we cannot start the compensator from $z_o = 0$.

We will see now that the generalizations to handle these particular

initial conditions are very simple. We will still assume that Σ_{11} is not

singular.

Generalization 1: $z_o \neq 0$; $E\{x_o\} = 0$; $E\{x_o x_o'\} = \Sigma_o$

Under these conditions $e_o = x_{20} - Hx_{10} - z_o$ is such that:

$$E\{e_o\} = -z_o; \quad E\{e_o e_o'\} = E_o(H) + Z_o$$

where $E_o(H)$ is as previously and $Z_o = z_o z_o'$. We see that the complete

demonstration goes through replacing $E_o(H)$ by $E_o(H) + Z_o$.

The optimal H is then given by:

$$\tilde{H}^* = (\tilde{W} A_{12}' + \Sigma'_{12}) \Sigma_{11}^{-1} \tag{4.3.1}$$

where $\tilde{W}$ satisfies now:

$$A_{22} \tilde{W} + \tilde{W} A_{22}' + \Sigma_{22} + Z_o - (\tilde{W} A_{12}' + \Sigma'_{12}) \Sigma_{11}^{-1} (\tilde{W} A_{12}' + \Sigma'_{12})' = 0 \tag{4.3.2}$$

It is clearly seen from the previous equation giving $\tilde{W}$, that starting

from $z_o \neq 0$ is in fact equivalent to adding extra uncertainty in the value

of x_{20}. This makes sense because if we know that $E\{x_0\} = 0$, we should

start from $E\{e_0\} = 0$. As we will see in an example, this extra uncertain-

ty will imply a faster response of the optimal compensator, to recover

more rapidly from the initial larger expected error.

The increase in the cost is given in this case by:

$$\Delta \hat{\tilde{J}}_2^* = tr(\tilde{\Lambda}_{22}^* \ \tilde{E}_0^*)$$

where:

$$\tilde{E}_0^* = \tilde{H}^* \ \Sigma_{11} \ \tilde{H}^{*'} + \Sigma_{22} + Z_0 - \tilde{H}^* \ \Sigma_{12} - \Sigma_{12}' \ \tilde{H}^{*'} \qquad (4.3.3)$$

and $\tilde{\Lambda}_{22}^*$ is the solution of:

$$\tilde{\Lambda}_{22} \ \tilde{F}^* + \tilde{F}^{*'} \ \tilde{\Lambda}_{22} + M = 0 \qquad (4.3.4)$$

in which $\tilde{F}^* = A_{22} - \tilde{H}^* \ A_{12}$.

Generalization 2. $E\{x_0\} = \bar{x}_0$, $cov(x_0, x_0) = \Sigma_0$, $\bar{x}_0 \ \bar{x}_0' = X_0$

Assume however that we can start the compensator from $z_0 = \bar{x}_{20} - H \ \bar{x}_{10}$;

this assumption insures $E\{e_0\} = 0$ and $E\{e_0 e_0'\} = E_0$ as previously,

(4.1.4).

The optimal compensator will then be the same as for the general

derivation, due to the fact that $E_0(H)$ is the same. In this case, H^* is

then given by (4.2.11) where W satisfies (4.2.13). The increase in cost

is then also the same and given by (4.2.14).

Note that these results are due to the fact that we assumed the possibility of "tuning" z_o to take the value $z_o = \bar{x}_{20} - H^*\bar{x}_{10}$.

<u>Generalization 3.</u> $z_o \neq 0$ arbitrary, $E\{x_o\} = \bar{x}_o$, $cov(x_o, x_o) = \Sigma_o$, $\bar{x}_o\bar{x}_o' = X_o$, $z_oz_o' = Z_o$.

Under these conditions, we obtain:

$$E\{e_o\} = \bar{e}_o = \bar{x}_{20} - H\bar{x}_{10} - z_o$$

$$E\{e_oe_o'\} = \bar{E}_o(H) = H(\Sigma_{11} + \bar{x}_{10}\bar{x}_{10}')H' + \Sigma_{22} + Z_o + \bar{x}_{20}\bar{x}_{20}' - \bar{x}_{20}z_o'$$

$$- H(\Sigma_{12} + \bar{x}_{10}\bar{x}_{20}' - \bar{x}_{10}z_o') - (\Sigma_{12} + \bar{x}_{10}\bar{x}_{20}' - \bar{x}_{10}z_o')'H'$$

Define now

$$\bar{\Sigma}_{11} = \Sigma_{11} + \bar{x}_{10}\bar{x}_{10}'$$

$$\bar{\Sigma}_{12} = \Sigma_{12} + \bar{x}_{10}\bar{x}_{20}' - \bar{x}_{10}z_o'$$

$$\bar{\Sigma}_{22} = \Sigma_{22} + Z_o + \bar{x}_{20}\bar{x}_{20}' - \bar{x}_{20}z_o'$$

then $\bar{E}_o(H) = H\bar{\Sigma}_{11}H' + \bar{\Sigma}_{22} - H\bar{\Sigma}_{12} - \bar{\Sigma}_{12}'H'$

The answer is then immediate:

$$\bar{H}^* = (\bar{W}A_{12}' + \bar{\Sigma}_{12}')\bar{\Sigma}_{11}^{-1}$$

where $\bar{W}$ satisfies:

$$A_{22} \bar{W} + \bar{W} A_{22}' + \bar{\Sigma}_{22} - (\bar{W} A_{12}' + \bar{\Sigma}_{12}') \bar{\Sigma}_{11}^{-1} (\bar{W} A_{12}' + \bar{\Sigma}_{12}')' = 0$$

The increase in cost is given by:

$$\Delta \bar{J}_2^* = tr(\bar{\Lambda}_{22}^* \bar{E}_o^*)$$

where $\bar{\Lambda}_{22}^*$ is the solution of

$$\bar{\Lambda}_{22} \bar{F}^* + \bar{F}^{*'} \bar{\Lambda}_{22} + M = 0$$

with $\bar{F}^* = A_{22} - \bar{H}^* A_{12}$ and

$$\bar{E}_o^* = \bar{H}^* \bar{\Sigma}_{11} \bar{H}^{*'} + \bar{\Sigma}_{22} - \bar{H}^* \bar{\Sigma}_{12} - \bar{\Sigma}_{12}' \bar{H}^{*'}$$

From this study of different initial conditions, we conclude that whatever the value of $E\{x_o\}$ is, the minimum increase in the cost will be obtained if we can "tune" the initial state of the compensator so as to insure $E\{e_o\} = 0$. If we cannot set z_o to this desirable value, the increase in cost will be bigger, and the optimal compensator faster, to reduce more rapidly the initial bigger expected error.

4.4 An Example. Σ_{11} Non-Singular

Assume we are given the completely controllable, completely observable, non output stabilizable, linear time-invariant system:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \qquad y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

Note that this system is already in the canonical form we need. Assume $E\{x_o\} = 0$ $E\{x_o x_o'\} = \Sigma_o = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and also that $r = 1$ and $Q = \begin{bmatrix} 9 & 0 \\ 0 & 10 \end{bmatrix}$ so that the optimal control assuming complete state feedback is:

$$u^*(t) = \begin{bmatrix} -3 & -4 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

i.e., $K^* = \begin{bmatrix} k_1^* & k_2^* \end{bmatrix} = \begin{bmatrix} -3 & -4 \end{bmatrix}$

leading to the optimal closed-loop system:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

which has two stable modes $s_1 = -3$, $s_2 = -1$.

The optimal cost is

$$\hat{J}_1^* = tr(\Pi\Sigma_o) = 16$$

where Π is given by:

$$\Pi = \begin{bmatrix} 12 & 3 \\ 3 & 4 \end{bmatrix}$$

Suppose now that we don't have access to $x_1(t)$ and $x_2(t)$ but only to $y(t) = x_1(t)$. A dynamic compensator must then be constructed.

1. Assume we start the compensator from $z_o = 0$. The general derivation then applies. From (4.2.13) w is computed to be: $w = 1$, so that from (4.2.11): $h^* = 1$.

From (3.2.3, 4) we then obtain

$$\dot{z}(t) = - z(t) - x_1(t) + u(t) \qquad (4.4.1)$$

and from (3.2.10)

$$u(t) = - 7 x_1(t) - 4 z(t) \qquad (4.4.2)$$

so that the "compact" optimal compensator becomes:

$$\dot{z}(t) = - 5 z(t) - 8 x_1(t)$$
$$u(t) = - 7 x_1(t) - 4 z(t)$$

leading to the overall closed loop system,

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{z}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -7 & 0 & -4 \\ -8 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ z(t) \end{bmatrix}$$

whose modes are $s_1 = -3$, $s_2 = -1$, and $s_3 = -1$. As expected, s_1 and s_2 correspond to the poles of the optimal closed loop system assuming complete state feedback, and s_3 corresponds to the pole associated with $f^* = a_{22} - h^*a_{12} = -1$ governing the error. Namely, we have $\dot{e}(t) = -e(t)$.

We see that for this optimal design the error because of the "small" value of Σ_0 does not decay really fast, contradicting all the heuristic arguments in favor of a fast observer.

The increase in cost is given by:

$$\Delta\hat{J}_2^* = \lambda_{22}^* \; E_0^* \; = 16$$

where $E_0^* = 2$ from (4.2.16), and $\lambda_{22}^* = 8$ from (4.2.15) in which $M = 16$. (Recall $M = N'K^{*'} RK^*N$, and in this example $N = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $K^* = [\, -3 \quad -4]$, $r = 1$).

Now for this simple academic example, we can compute easily $\Delta\hat{J}_2(h) = \lambda_{22}(h) \; E_0(h)$.

From (4.1.1), we have $\lambda_{22}(h) = \dfrac{8}{h}$ and from (4.1.14) $E_0(h) = h^2 + 1$, so that:

$$\Delta J_2(h) = 8(h + \frac{1}{h})$$

Recall that the optimal cost assuming complete state feedback is $\hat{J}_1^* = 16$; we can then very easily evaluate the percentage of increase in the cost due to the fact that we had to use a compensator:

$$\tau(h) = \frac{\Delta \hat{J}_2(h)}{\hat{J}_1^*} = \frac{1}{2}(h + \frac{1}{h})$$

A sketch of $\tau(h)$ as a function of h is shown in Figure 5, page 60.

We see clearly on this sketch that $h^* = 1$ leads to the minimum percentage of increase (in this academic example, 100%), and that a too small or a too big h lead to high increases in the cost, one because the estimation of $x_2(t)$ is very poor leading to a poor control $u(t) = K^*\hat{x}(t)$, the other one because too much effort is spent to do an unnecessarily good estimation of $x_2(t)$.

2. Assume now that for some practical reason, we cannot start the compensator from $z_o = 0$, but instead that $z_o = \sqrt{3}$.

Obviously this is a real disadvantage because now $E\{e_o\} = -\sqrt{3}$ instead of zero. We then impose on the observer part of the compensator to start with a large initial error. We would then expect the optimal compensator in this case to be faster than in the previous design where $E\{e_o\} = 0$. This is precisely what happens.

From (4.3.2) and (4.3.1) we have $\tilde{w} = 2$ and $\tilde{h}^* = 2$ so that $\tilde{f}^* = -2$ and for this design $\dot{e}(t) = -2\ e(t)$.

The equations for the compensator are:

$$\dot{z}(t) = -2\ z(t) - 4\ x_1(t) + u(t)$$
$$u(t) = -11\ x_1(t) - 4\ z(t) \tag{4.4.3}$$

so that we will implement the compact compensator:

$$\dot{z}(t) = -6\,z(t) - 15\,x_1(t)$$

associated with (4.4.3).

The overall optimal system becomes:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{z}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -11 & 0 & -4 \\ -15 & 0 & -6 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ z(t) \end{bmatrix}$$

corresponding to the following modes $s_1 = -3$, $s_2 = -1$, associated with $s_3 = -2$.

The increase in cost is in this case:

$$\Delta \hat{\tilde{J}}^*_2 = 32$$

As previously, we can draw a sketch of the percentage of increase

$$\tilde{\tau}(h) = \frac{\Delta \hat{\tilde{J}}_2(h)}{\hat{J}^*_1} \text{ , (see Figure 6, page 60.)}$$

We have in this case from (4.3.3) and (4.3.4):

$$\tilde{\lambda}_{22}(h) = \frac{8}{h} \quad \text{and} \quad \tilde{E}_o(h) = h^2 + 4$$

so that $\tilde{\tau}(h) = \dfrac{1}{2}\left(h + \dfrac{4}{h}\right)$.

We see that now the minimum percentage of increase is 200%. It is then very important to be able to start the compensator so as to insure $E\{e_o\} = 0$.

Remark. One should not attach too much importance to the numbers in this academic example, which has been constructed so as to insure integer values of all the parameters. We will see in Chapter V another example where the percentage of increase is of the order of 10^{-3}.

3. Assume now that $\Sigma_o = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ but that $z_o = 0$.

The uncertainty in x_{20} is now bigger and one would expect a faster compensator than the one found in 1.

This is exactly what happens, and it can be seen that $w = 2$ and $h^* = 2$ for this example, so that $f^* = -2$.

The increase in the cost is now $\Delta \hat{J}_2^* = 32$.

Percentage of Increase in the Cost ($z_o = 0$)

$$\tau(h) = \frac{\Delta \hat{J}_2(h)}{\hat{J}_1^*}$$

Fig. 5

Percentage of Increase in the Cost ($z_o = \sqrt{3}$)

$$\tilde{\tau}(h) = \frac{\Delta \hat{\tilde{J}}_2(h)}{\hat{J}_1^*}$$
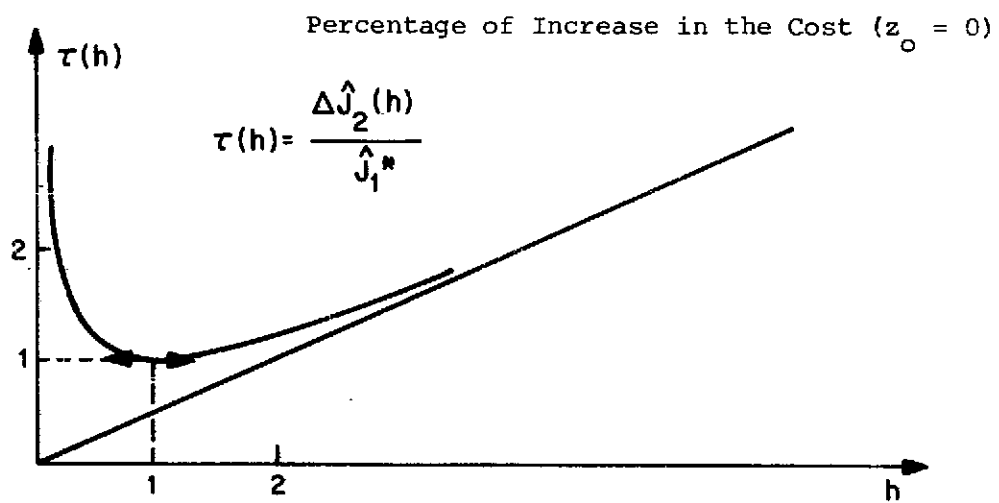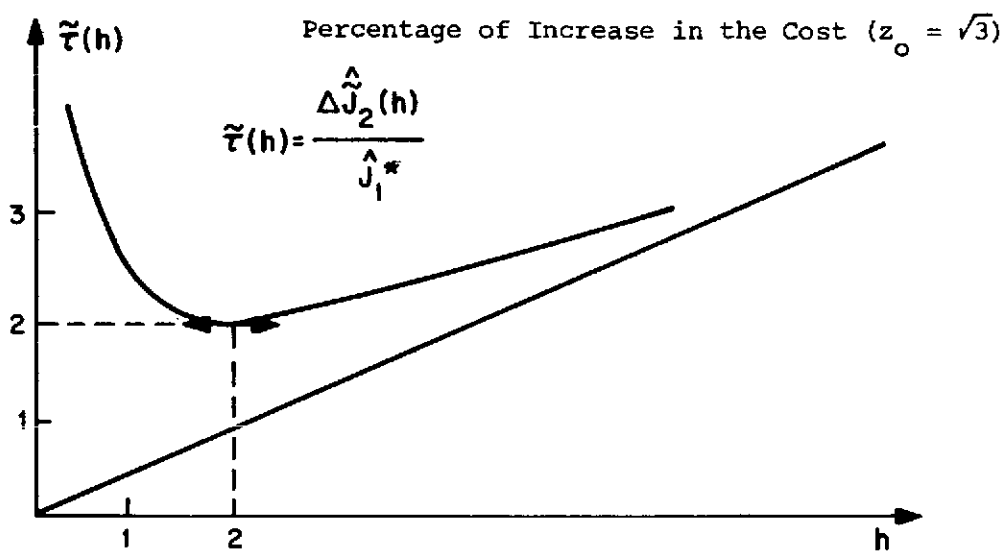
Fig. 6

4.5 Proof of the Result. Σ_{11} Identically Zero

We will return now to the proof of the main theorem, and study the case where Σ_{11} is identically zero, before we tackle the case where Σ_{11} is simply singular.

Assume $\Sigma_{11} = E\{x_{10}x'_{10}\} \equiv 0$. As we also assumed $E\{x_{10}\} = 0$, this means that $x_{10} = 0$ with probability one. We will simply say that we know perfectly that $x_{10} = 0$. Note that under these conditions, the initial covariance matrix takes the form:

$$\Sigma_o = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_{22} \end{bmatrix} \tag{4.5.1}$$

If we still assume $z_o = 0$ for the compensator, then from (4.1.4), we see that now:

$$E_o = \Sigma_{22} \tag{4.5.2}$$

and E_o does not depend any longer on H.

The increase in cost becomes:

$$\Delta \hat{J}_2(H) = \text{tr}(\Lambda_{22} \Sigma_{22}) \tag{4.5.3}$$

where Λ_{22} is still the solution of the Lyapunov equation:

$$\Lambda_{22} F + F' \Lambda_{22} + M = 0 \tag{4.5.4}$$

It should be obvious that the only way to reduce $\Delta \hat{J}_2(H)$ is now to try to make $\text{tr}(\Lambda_{22})$ as "small" as possible. We will show that for this particular case of perfect knowledge of the initial state x_{10}, there is no

minimum for $\Delta\hat{J}_2(H)$, and that the increase in cost can be made arbitrarily small, by choosing F to have greater and greater negative eigenvalues, and consequently choosing a "greater and greater" H.

From (4.1.5), we have

$$
\Lambda_{22} = \int_0^\infty e^{(A_{22} - HA_{12})'t} M e^{(A_{22} - HA_{12})t} dt
$$

so that $\Delta\hat{J}_2(H) = tr(\Lambda_{22} \Sigma_{22})$ becomes:

$$
\Delta\hat{J}_2(H) = tr \left\{ M \int_0^\infty e^{(A_{22} - HA_{12})t} \Sigma_{22} e^{(A_{22} - HA_{12})'t} dt \right\}
$$

We will now give simply a few indications for the demonstration as we will mainly follow what we have previously done.

To the first order in ε, we have:

$$
\Delta\hat{J}_2(H + \varepsilon\Delta H) = tr \left\{ M \int_0^\infty \left[e^{Ft} \Sigma_{22} e^{F't} - \varepsilon \int_0^t e^{F(t - \sigma)} \Delta H A_{12} e^{F\sigma} d\sigma \Sigma_{22} e^{F't} \right. \right.
$$

$$
\left. \left. - \varepsilon e^{Ft} \Sigma_{22} \int_0^t e^{F'(t - \sigma)} A'_{12} \Delta H' e^{F'\sigma} d\sigma \right] dt \right\}
$$

We see immediately that all the corresponding steps will be the same, where we will set $\Sigma_{11} = 0$ and $\Sigma_{12} = 0$ in the equations.

We finally end up with:

$$
\frac{d}{dH} \Delta\hat{J}_2(H) = - 2 \int_0^\infty e^{F't} e^{Ft} dt \int_0^\infty e^{Ft} \Sigma_{22} e^{F't} dt \, A'_{12} M \qquad (4.5.5)
$$

Define U by

$$U = \int_0^\infty e^{F't} \, e^{Ft} \, dt \tag{4.5.6}$$

then U is a solution of the algebraic Lyapunov equation:

$$UF + F'U + I = 0 \tag{4.5.7}$$

Define V by:

$$V = \int_0^\infty e^{Ft} \, \Sigma_{22} \, e^{F't} \, dt \tag{4.5.8}$$

then V is a solution of the algebraic Lyapunov equation:

$$VF' + FV + \Sigma_{22} = 0 \tag{4.5.9}$$

Note that I is positive definite, that Σ_{22} is positive semi-definite, and that we are looking for a stable F, so that solutions U and V are required to be positive definite.

We can now rewrite (4.5.5) as

$$\frac{d}{dH} \Delta \hat{J}_2(H) = -2U \, V \, A'_{12} \, M \qquad \text{(recall } M \geq 0)$$

where $U > 0$, $V > 0$, and are only functions of H through $F = A_{22} - HA_{12}$. We then see clearly that $\Delta \hat{J}_2(H)$ is a monotonically decreasing function of H,

and that no effective minimum exists.

Mathematically, the increase in cost can be made arbitrarily small. However, in practice, we won't implement a compensator which is "too fast", because in this case it will tend to act as a differentiator, and small measurement noise or plant noise (not modeled in this deterministic approach) will give rise to spurious signals.

Generalization. Σ_{11} Identically Zero

Assume $E\{x_o\} = \bar{x}_o$, that is $E\{x_{10}\} = \bar{x}_{10}$ $E\{x_{20}\} = \bar{x}_{20}$, but still $\text{cov}(x_{10}, x_{10}) = 0$ and $\text{cov}(x_{20}, x_{20}) = \Sigma_{22}$. This means that we know perfectly that x_{10} will take the value $\bar{x}_{10}$.

Define $X_{11} = \bar{x}_{10} \bar{x}'_{10}$, $X_{12} = \bar{x}_{10} \bar{x}'_{20}$, $X_{22} = \bar{x}_{20} \bar{x}'_{20}$, $Z_o = z_o z'_o$. We see immediately that the error is such that:

$$E\{e_o\} = \bar{x}_{20} - H \bar{x}_{10} - z_o$$

There are two possibilities here: either we can tune z_o to take the value $z_o = \bar{x}_{20} - H \bar{x}_{10}$ so that $E\{e_o\} = 0$, and the design will be the same as previously (no minimum, and possibility of making $\Delta \hat{J}_2(H)$ as small as we want); or if z_o is arbitrary but fixed, we obtain:

$$E_o(H) = H X_{11} H' + \Sigma_{22} + X_{22} + Z_o - H(X_{12} - \bar{x}_{10} z'_o) - (X_{12} - \bar{x}_{10} z'_o)' H'$$

so that $E_o(H)$ appears again as a quadratic function of H. The answer is immediate: if X_{11} is non-singular (i.e., no component of $\bar{x}_{10}$ is zero), then there will be a unique minimizing H^*.

Define $\hat{X}_{12} = X_{12} - \bar{x}_{10} z'_o$, $\hat{X}_{22} = \Sigma_{22} + X_{22} + Z_o$. Then H^* will be given by:

$$\hat{H}^* = (\hat{W} A_{12}' + \hat{X}_{12}') X_{11}^{-1} \qquad (4.5.10)$$

where $\hat{W}$ satisfies:

$$A_{22} \hat{W} + \hat{W}A_{22} + \hat{X}_{22} - (\hat{W}A_{12}' + \hat{X}_{12}') X_{11}^{-1} (\hat{W}A_{12}' + \hat{X}_{12}')' = 0 \qquad (4.5.11)$$

Note that there exists a unique minimizing H* as soon as $E_o(H)$ is a quadratic function of H, if the quadratic term possesses a positive definite matrix coefficient.

Example: We will show now what happens to the example in (4.4.1), page 55, if we take $\sigma_{11} = 0$, that is $\Sigma_o = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, with still $E\{x_{10}\} = E\{x_{20}\} = 0$.

It is easily seen that in this case $E_o = \sigma_{22} = 1$ and that as previously $\lambda_{22}(h) = \dfrac{8}{h}$, so that:

$$\Delta \hat{J}_2(h) = \frac{8}{h}$$

The percentage of increase becomes now:

$$\tau(h) = \frac{\Delta \hat{J}_2(h)}{\hat{J}_1^*} = \frac{1}{2h}$$

and can be made arbitrarily small by choosing sufficiently large h.

The engineer is then completely free to choose the dynamics of the compensator. Other practical considerations will lead to choice of H*.

Assume now $E\{x_{10}\} = 1$, $E\{x_{20}\} = 0$, $z_o = 0$. It is easily seen that

$E_o(h) = h^2 + 1$ in this case, leading to a unique minimizing h, for

$$\Delta\hat{J}_2(h) = 8\,\frac{h^2 + 1}{h} \,.$$

Namely from (4.5.10, 11) we have $\hat{w} = 1$, $\hat{h}* = 1$ and the optimal

compensator follows.

4.6 Proof of the Result. Σ_{11} is Singular

We will assume now that Σ_{11} is singular but not identically zero.

Without loss of generality, we can rearrange the components of $x_1(t)$ to read:

$$x_1(t) \triangleq \begin{bmatrix} \bar{x}_1(t) \\ \tilde{x}_1(t) \end{bmatrix}$$

where $\bar{x}_1(t)$ is an m' vector (m' < m) and $\tilde{x}_1(t)$ an m-m' vector, such that $E\{\bar{x}_{10}\} = 0$ and $\tilde{x}_{10}$ is perfectly known to be zero. Accordingly Σ_{11} becomes

$$\Sigma_{11} = \begin{bmatrix} \bar{\Sigma}_{11} & 0 \\ 0 & 0 \end{bmatrix}$$

where $\bar{\Sigma}_{11}$ is an m'xm' non-singular matrix.

As a consequence, $E\{x_{10} x'_{20}\} = \Sigma_{12}$ becomes:

$$\Sigma_{12} = \begin{bmatrix} \bar{\Sigma}_{12} \\ 0 \end{bmatrix}$$

where $\bar{\Sigma}_{12}$ is an m'x(n-m) matrix.

The global initial state covariance matrix becomes now:

$$\Sigma_0 = \begin{bmatrix} \bar{\Sigma}_{11} & 0 & \bar{\Sigma}_{12} \\ 0 & 0 & 0 \\ \bar{\Sigma}'_{12} & 0 & \Sigma_{22} \end{bmatrix}$$

Given this new partitioning of Σ_{11} and Σ_0, it is interesting to partition H and A accordingly.

Define $H = [H_1 \quad H_2]$ where H_1 is $(n-m) \times m'$ and H_2 is $(n-m) \times (m-m')$; define also:

$$A_{12} \stackrel{\Delta}{=} \begin{bmatrix} \bar{A}_{12} \\ \tilde{A}_{12} \end{bmatrix}$$

where $\bar{A}_{12}$ is $m' \times (n-m)$ and $\tilde{A}_{12}$ is $(m-m') \times (n-m)$.

Note that with these new definitions $E_o(H)$ becomes, (as can be readily seen):

$$E_o(H) = H_1 \bar{\Sigma}_{11} H_1' + \Sigma_{22} - H_1 \bar{\Sigma}_{12} - \bar{\Sigma}_{12}' H_1' \qquad (4.6.1)$$

so that in fact $E_o(H)$ is simply a function of H_1. From this simple fact, we can immediately predict that there will be no effective minimum in this case either.

Assume first that $H = [H_1 \quad H_2]$ <u>is fixed</u>. We can still solve for W given by (4.2.12) where now $E_o(H) = E_o(H_1)$; namely we have:

$$FW + WF' + E_o(H_1) = 0 \qquad (4.6.2)$$

in which $F = A_{22} - H_1 \bar{A}_{12} - H_2 \tilde{A}_{12}$

Assuming $W > 0$ is found from (4.6.2), let's see what happens to the necessary condition for a minimum given by (4.2.9):

$$H \Sigma_{11} - \Sigma_{12} = W A_{12}' \qquad (4.2.9)$$

Namely we have now:

$$\begin{bmatrix} H_1 \ \bar{\Sigma}_{11} - \bar{\Sigma}'_{12} & 0 \end{bmatrix} = \begin{bmatrix} W\bar{A}'_{12} & W\tilde{A}{}'_{12} \end{bmatrix} \tag{4.6.3}$$

We no longer have the possibility of satisfying (4.6.3), except in the very special case: $\tilde{A}_{12} \equiv 0$. We will briefly study this case before we proceed for the general case.

<u>Special Case:</u> $\tilde{A}_{12} \equiv 0$. Note that F becomes accordingly $F = A_{22} - H_1\bar{A}_{12}$ and does not depend on H_2. As from (4.6.1), H_2 does not appear in E_o, we will simply set $H_2 \equiv 0$. H* is now simply given by $H^* = [H_1^* \quad 0]$, where we have:

$$H_1^* = (W_1 \ \bar{A}'_{12} + \bar{\Sigma}'_{12}) \ \bar{\Sigma}_{11}^{-1}$$

W_1 being given by:

$$A_{22} \ W_2 + W_1 \ A'_{22} + \Sigma_{22} - (W_1 \ \bar{A}'_{12} + \bar{\Sigma}'_{12}) \ \bar{\Sigma}_{11}^{-1} \ (W_1 \ \bar{A}'_{12} + \bar{\Sigma}'_{12})' = 0$$

and an effective minimum exists for this very special case.

<u>General Case:</u> We will now return to the general case, and clearly show the different roles played by H_1 and H_2, by computing separately $\frac{\partial}{\partial H_1} \Delta\hat{J}_2 (H_1, \ H_2)$ and $\frac{\partial}{\partial H_2} \Delta\hat{J}(H_1, \ H_2)$.

These derivatives may be written directly:

$$\frac{\partial}{\partial H_1} \Delta \hat{J}_2 (H_1, H_2) = 2\int_0^\infty e^{F't} e^{Ft} \, dt \left[(H_1' \ \Sigma_{11} - \bar{\Sigma}_{12}') \right.$$

$$\left. -\int_0^\infty e^{F\tau} E_o (H_1) e^{F'\tau} \, d\tau \ \bar{A}_{12}' \right] M \qquad (4.6.4)$$

$$\frac{\partial}{\partial H_2} \Delta \hat{J}_2 (H_1, H_2) = - 2 \int_0^\infty e^{F't} e^{Ft} \, dt \int_0^\infty e^{F\tau} E_o (H_1) e^{F'\tau} \, d\tau \ \tilde{A}_{12}' M$$
$$(4.6.5)$$

where:

$$F = A_{22} - H_1 \bar{A}_{12} - H_2 \tilde{A}_{12}$$

$$E_o (H_1) = H_1 \bar{\Sigma}_{11} H_1' + \Sigma_{22} - H_1 \bar{\Sigma}_{12} - \bar{\Sigma}_{12}' H_1'$$

As expected, we see that there is no way to set $\frac{\partial}{\partial H_2} \Delta \hat{J}_2 (H_1, H_2)$ to be equal to zero. This derivative is always negative, and as a consequence $\Delta \hat{J}_2 (H_1, H_2)$ is a monotonic decreasing function of H_2.

On the other hand, assume that a particular H_2 has been chosen. Then from (4.6.4), we see that for this particular H_2, there exists a unique H_1^* for which $\left. \frac{\partial}{\partial H_1} \Delta \hat{J}_2 (H_1, H_2) \right|_{H_1^*} = 0$.

This suggests clearly that for any arbitrary H_2, say H_{2a}, there exists an "optimal" H_{1a}^* such that $\forall H_1 \neq H_{1a}^*$ we have:

$$\Delta \hat{J}_2 (H_{1a}^*, H_{2a}) \leq \Delta \hat{J}_2 (H_1, H_{2a})$$

This "partially minimizing" H_{1a}^* is easily computed by defining $A_{22a} =$

$A_{22} - H_{2a} \tilde{A}_{12}$, so that we have immediately:

$$H^*_{1a} = (W_a \bar{A}'_{12} + \bar{\Sigma}'_{12}) \bar{\Sigma}^{-1}_{11} \qquad (4.6.6)$$

where W_a satisfies:

$$W_a A'_{22a} + A_{22a} W_a + \Sigma_{22} - (W_a \bar{A}'_{12} + \bar{\Sigma}'_{12}) \bar{\Sigma}^{-1}_{11} (W_a \bar{A}'_{12} + \bar{\Sigma}'_{12})' = 0$$

$$(4.6.7)$$

The important fact to remember is that when Σ_{11} is singular, but not identically zero, not any arbitrary choice of a "large" H will lead to a small cost. The procedure is thus to fix H_{2a} on the basis of engineering judgment and then choose H^*_{1a} solution of (4.6.6, 7).

We will soon give a way to choose a reasonable H_{2a}, but first we will show on a three-dimensional sketch, how $\Delta \hat{J}_2 (H_1, H_2)$ behaves as a function of H_1 and H_2. See Figure 7, page 77.

From the sketch, it is clear that in any hyperplane of constant H_1, $\Delta \hat{J}_2 (H_1, H_2)$ is a monotone decreasing function of H_2. On the contrary in any hyperplane of constant H_2, $\Delta \hat{J}_2 (H_1, H_2)$ possesses a single minimum for $H_1 = H^*_{1a}$.

Proposed Method for a Choice of H_{2a}.

We propose now a possible way of choosing H_{2a}. Our problem comes from the fact that Σ_{11} is singular: the idea is to consider a fake Σ_{11} which is no longer singular. In other words, instead of:

$$\Sigma_{11} = \begin{bmatrix} \bar{\Sigma}_{11} & 0 \\ 0 & 0 \end{bmatrix}$$

take the following fake initial covariance matrix

$$\Sigma_{11f} = \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \sigma^2 I \end{bmatrix}$$

where we can adjust σ to obtain a desirable optimal response using the first design (Σ_{11} non-singular). An initial guess for σ^2 could be $\sigma^2 = $ norm $\bar{\Sigma}_{11}$, corresponding to saying that the fake standard deviation of $\tilde{x}_{10}$ has approximately the same size as the one of $\bar{x}_{10}$. (Recall that $x_{10} = \begin{bmatrix} \bar{x}_{10} \\ \tilde{x}_{10} \end{bmatrix} = \begin{bmatrix} \bar{x}_{10} \\ 0 \end{bmatrix}$ and that we want to give a fake covariance matrix to $\tilde{x}_{10}$.)

Now Σ_{11f} will lead to a unique H* through formulas (4.2.10, 11). Write now $H_f^* = [H_1^*, H_2^*]$ and pick $H_{2a} = H_2^*$. Using formulas (4.6.6, 7), it is easy to compute the value of H_{1a}^* for the real design.

This method is not the unique one, and the designer has entire freedom to choose H_2 so as to satisfy certain desired properties of the closed loop response of the system. The important fact to remember here, is that, once H_2 has been chosen, then H_1 is uniquely determined by (4.6.6, 7).

Example. We will now show by an example the procedure to follow.

Consider the following completely controllable, completely observable, non-output stabilizable, time-invariant plant:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(t)$$

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$

1. Assume first that we have access to the complete state; then a cost (3.1.4) with:

$$Q = \begin{bmatrix} 36 & 0 & 0 \\ 0 & 13 & 0 \\ 0 & 0 & 14 \end{bmatrix} \qquad \text{and } r = 1$$

leads to the optimal control:

$$u^*(t) = [-6 \quad -5 \quad -6] \begin{bmatrix} x_1(t) \\ x_2(t) \\ \hat{x}_3(t) \end{bmatrix}$$

and the optimal closed loop system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ -6 & -5 & -6 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$

corresponding to the stable modes $s_1 = -3$, $s_2 = -2$; $s_3 = -1$. Moreover, if we assume $E\{x_o\} = 0$ and

$$\Sigma_o = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

then the optimal cost is given by:

$$\hat{J}_1^* = \text{tr}(\Pi \, \Sigma_o) = 42$$

where Π is given by:

$$\Pi = \begin{bmatrix} 30 & 6 & 6 \\ 6 & 18 & 5 \\ 6 & 5 & 6 \end{bmatrix}$$

2. Now in practice, we don't have access to $x_3(t)$, so that we will design an observer to estimate $x_3(t)$, and then apply the control:

$$u(t) = [-6 \quad -5 \quad -6] \begin{bmatrix} x_1(t) \\ x_2(t) \\ \hat{x}_3(t) \end{bmatrix}$$

Now it is immediately seen that Σ_{11} is singular. $\Sigma_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. We will then take a fake Σ_{11} to pick the value of h_{2a}.

Consider $\Sigma_{11f} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (i.e., $\sigma^2 = 1$).

Under these conditions, using formulas (4.2.10, 11) we obtain $W = +1$, $H_f^* = [1 \quad 1]$. We will then pick $h_{2a} = +1$. Putting this value of h_2 in (4.6.6, 7) leads immediately to $h_{1a}^* = \sqrt{3} - 1$. So that we will then implement a compensator with

$$H = [\sqrt{3} - 1, 1]$$

For this design

$$F = 0 - [\sqrt{3} - 1 \quad 1] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -\sqrt{3}$$

so that we will have $\dot{e}(t) = -\sqrt{3} \quad e(t)$.

Remark. Note that for Σ_{11f}, the dynamics of the error would have

been $\dot{e}(t) = -2\ e(t)$. The actual compensator is slower. This is exactly what we expected since, in fact, the real initial state covariance matrix is such that the initial expected error is smaller in the real design. The observer expends then less work in recovering from the initial error, and "slows down" its speed.

As a verification of our results, we will try to minimize directly $\Delta\hat{J}_2(h_1,\ h_2)$ for this simple example.

$$\Delta\hat{J}_2(h_1,\ h_2) = \lambda_{22}(h_1,\ h_2)\ E_o(h_1)$$

From (4.6.1), we have

$$E_o(h_1) = h_1^2 + 2$$

and noting that $F = -(h_1 + h_2)$ we have

$$\lambda_{22}(h_1,\ h_2) = \frac{18}{h_1 + h_2} \quad \text{(from 4.2.15 with M = 36)}$$

so that we obtain

$$\Delta\hat{J}_2(h_1,\ h_2) = 18\ \frac{h_1^2 + 2}{h_1 + h_2}$$

$$\frac{\partial}{\partial h_1}\ \Delta\hat{J}_2(h_1,\ h_2) = \frac{18}{(h_1 + h_2)^2}\ (h_1^2 + 2h_1h_2 - 2)$$

$$\frac{\partial}{\partial h_2} \Delta\hat{J}_2 (h_1,\ h_2) = -18 \frac{h_1^2 + 2}{(h_1 + h_2)^2}$$

As expected, $\Delta\hat{J}_2(h_1,\ h_2)$ is a monotonic decreasing function of h_2, and a choice of $h_{2a} = 1$ leads to $h_{1a}^* = \sqrt{3} - 1$.

The increase in the cost for our design, $(H = [\sqrt{3} - 1,\ 1]$ is then:

$$\Delta\hat{J}_2^* = 36(\sqrt{3} - 1)$$

Note that this increase could be made arbitrarily small by choosing h_2 greater and greater and choosing accordingly $h_{1a}^* = \sqrt{2 + h_{2a}^2} - h_{2a}$. It can easily be seen that under these conditions, $\Delta\hat{J}_2^*(h_{1a}^*,\ h_{2a}) = 36 (\sqrt{2 + h_{2a}^2} - h_{2a})$ is a monotonic decreasing function of h_{2a}. (Recall that h_{2a} must be positive here, for stability of F).

Remark. The generalizations to handle particular initial conditions are straightforward. If $E_o(H)$ turns out to be a quadratic function of H, whose quadratic term possesses a positive definite coefficient matrix, then a unique minimizing H^* exists, and the optimal compensator follows, with specified dynamics.
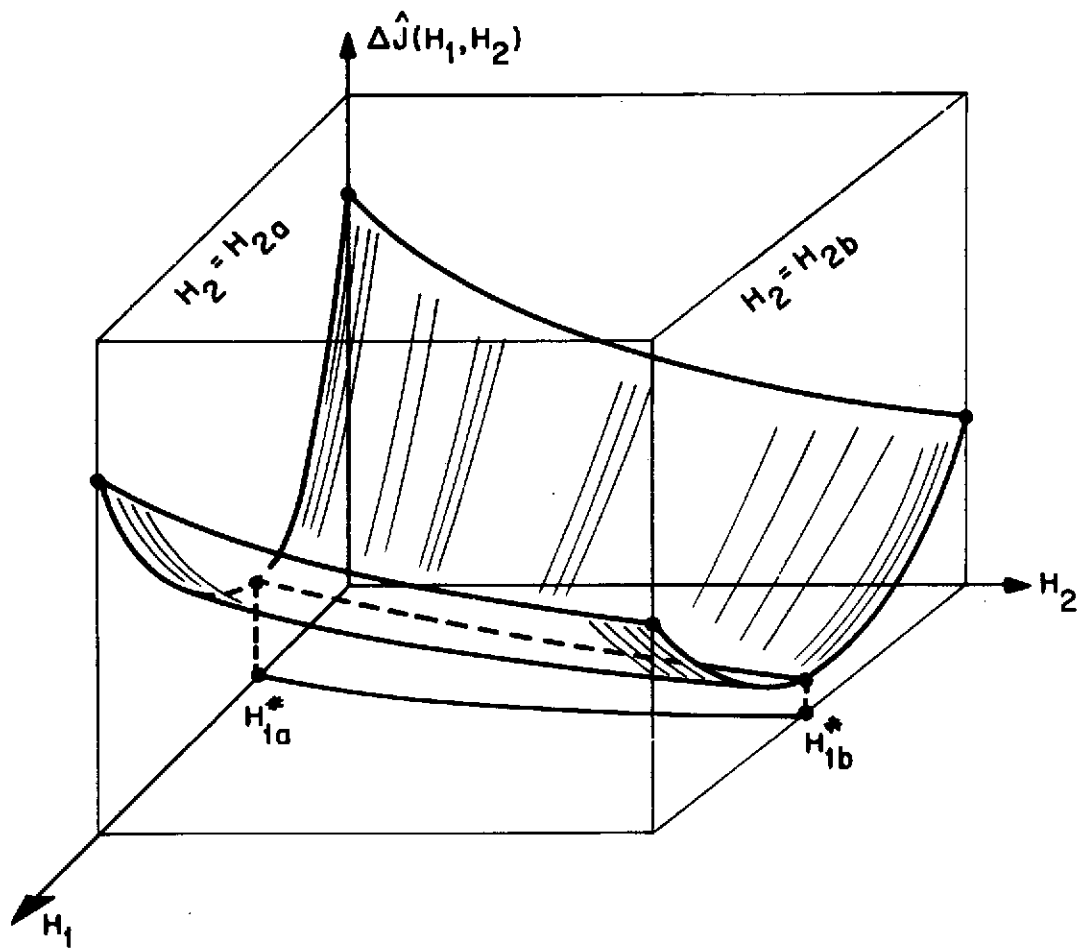
Fig. 7

A Geometric Interpretation of $\Delta\hat{J}(H_1, H_2)$

CHAPTER V

THE CLASS OF OPTIMAL COMPENSATORS

5.1 Introduction

We derived in Chapters III and IV the solution for the optimal control of an n-th order plant having n-m inaccessible states, where the control is to be generated from a compensator of order n-m, and where the cost to be minimized is:

$$\hat{J} = E\left\{ \int_0^\infty \{x'(t)\ Q\ x(t)\ +\ u'(t)\ R\ u(t)\}\ dt \right\} \qquad (5.1.1)$$

We showed that the equivalent of the separation property holds for this problem, namely that the optimal compensator can be decomposed in two totally independent parts: a minimal-order Luenberger observer with optimally specified dynamics, and a controller where we used the optimal gains for the linear regulator problem.

Now the uniqueness of the solution was mainly due to the fact that we used a canonical form for the output matrix of the plant, namely:

$$C = [\,I_{mxm} \qquad O_{mx(n-m)}\,] \qquad (5.1.2)$$

We saw in Chapter II how to go from any arbitrary n-th order plant with output matrix $\tilde{C}$ of rank m,

$$\dot{\bar{x}}(t) = \bar{A} \ \bar{x}(t) + \bar{B} \ \bar{u}(t); \ \bar{y}(t) = \bar{C} \ \bar{x}(t) \qquad (5.1.3)$$

to an n-th order canonical plant:

$$\dot{x}(t) = A \ x(t) + B \ u(t); \ y(t) = C \ x(t) \qquad (5.1.4)$$

where C is given by (5.1.2). We also showed how to transform an optimization problem for the initial plant into an equivalent optimization problem for the canonical plant. It is interesting now to return to the initial plant and see what happens to the optimal compensator.

Recalling that the similarity transformation P to go from (5.1.3) to (5.1.4) is not unique, it is expected that the optimal compensator for the initial plant will not be unique. However, we will define precisely the class of all optimal compensators.

We will now show precisely how to go back to the initial plant, then prove that the total cost for the overall canonical plant is the same as the total cost for the overall optimal initial plant, whatever P matrix we use to go from one to another. We will then define the class of all optimal compensators for the initial plant, and finally provide a numerical example.

5.2 Equality of the Total Costs

Recall from Chapter II that we go from (5.1.3) to (5.1.4) by using the similarity transformation:

$$x(t) = P \bar{x}(t) \qquad (5.2.1)$$

where we have defined:

$$P = \begin{bmatrix} \bar{C} \\ T \end{bmatrix} \qquad (5.2.2)$$

such that C given by (5.1.2) becomes:

$$C = \bar{C} \, P^{-1} \qquad (5.2.3)$$

Recall that the nonuniqueness of P is due to the fact that T is arbitrary as long as its row vectors span the orthogonal complement of the range of $\bar{C}$. Recall also the relations between the two systems:

$$A = P \bar{A} \, P^{-1}; \quad B = P \bar{B}; \quad C = \bar{C} \, P^{-1} \qquad (5.2.4)$$

$$Q = (P')^{-1} \bar{Q} \, P^{-1}; \quad R = \bar{R}; \quad K* = \bar{K}* \, P^{-1} \qquad (5.2.5)$$

In the following, any symbol with an overbar will be related to the initial given plant, any other symbol without this overbar will be related to a particular canonical plant associated with a given P transformation.

The overall optimal canonical system is given by (3.3.4) and (3.3.5)

that is:

$$\dot{\xi}(t) = \Gamma^* \, \xi(t) \tag{5.2.6}$$

where we have defined:

$$\Gamma^* = \begin{bmatrix} A + BK^* & - BK^* \, N \\ 0 & F^* \end{bmatrix} \tag{5.2.7}$$

for which $F = A_{22} - H^* A_{12}$ governs the dynamics of the error $e(t) = x_2(t) - \hat{x}_2(t)$.

Define now the transformation:

$$P = \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} P^{-1} & 0 \\ 0 & I \end{bmatrix} \tag{5.2.8}$$

then if $\bar{\xi}(t) = \begin{bmatrix} \bar{x}(t) \\ \bar{e}(t) \end{bmatrix}$ we see that

$$\bar{\xi}(t) = P^{-1} \, \xi(t) \tag{5.2.9}$$

so that we can easily come back to the initial overall optimal system associated with the particular P matrix used. Namely, we have:

$$\dot{\bar{\xi}}(t) = \bar{\Gamma}^* \, \bar{\xi}(t) \tag{5.2.10}$$

where $\bar{\Gamma}^*$ is given by:

$$\bar{\Gamma}* = P^{-1} \Gamma* P = \begin{bmatrix} \bar{A} + \bar{B}\bar{K}* & -\bar{B}\bar{K}* \ P^{-1} \ N \\ 0 & F* \end{bmatrix}$$

A simple look at this matrix shows that there will be as many optimal compensators as there are P's.

Now the total cost associated with (5.2.) is given by:

$$\hat{J}*_2 = \text{tr}(\Lambda* \ \Xi*_o) \tag{5.2.11}$$

where Λ is given by (3.3.24) and Ξ_o by (3.3.17) all quantities being computed for the optimal H.

On the other hand, the total cost associated with (5.2.10) is given by:

$$\hat{\bar{J}}*_2 = \text{tr}(\Lambda* \ \Xi*_o) \tag{5.2.12}$$

where $\bar{\Lambda}*$ is the solution of:

$$\Lambda \ \bar{\Gamma}* + \bar{\Gamma}*' \ \Lambda + \bar{\Omega} = 0 \tag{5.2.13}$$

where $\bar{\Omega}$ is given by:

$$\bar{\Omega} = \begin{bmatrix} \bar{Q} + \bar{K}*' \ \bar{R} \ \bar{K}* & -\bar{K}*' \ \bar{R} \ \bar{K}* \ P^{-1} \ N \\ -N'(P')^{-1}\bar{K}*' \ \bar{R} \ \bar{K}* & \bar{K}*' \ \bar{R} \ \bar{K}* \end{bmatrix} \tag{5.2.14}$$

It is a simple matter now, using formulas (5.2.8, 9) and the different relations between the two systems, to see that:

$$\Xi_o = P \; \bar{\bar{\Xi}}_o \; P'; \quad \bar{\Omega} = P' \; \Omega \; P \tag{5.2.15}$$

so that from (5.2.13) follows:

$$\bar{\Lambda} = P' \; \Lambda \; P \tag{5.2.16}$$

Now with these formulas, a very simple calculation, similar to the one we did in Chapter II, yields:

$$\hat{J}_2^* = \text{tr}(\Lambda^* \; \Xi_o^*) = \text{tr}(\bar{\Lambda}^* \; \bar{\Xi}_o^*) = \hat{\bar{J}}_2^*$$

In other words, we simply showed that whatever similarity transformation we used to go from the initial plant to the canonical plant, the total costs and consequently the increases in the costs (as we already proved (Chapter II) that the constant parts of the costs were equal), are the same. We proved here a first desirable property of the design; although for an arbitrary plant there is a class of optimal compensators rather than a unique one; all the compensators in the class lead to the same increase in the cost. We will return now to the variables x(t) and z(t), and give precisely the compensator we would implement around the initial plant.

5.3   Optimal Compensator.   Initial Plant

We will return now to the practical variables $x(t)$ and $z(t)$, as we will finally implement $z(t)$ and not $e(t)$.

From (3.2.3), recall that the optimal minimal-order Luenberger observer is given by:

$$\dot{z}(t) = F^* \, z(t) + G^* \, x_1(t) + D^* \, u(t) \tag{5.3.1}$$

$$\hat{x}_2(t) = z(t) + H^* \, x_1(t) \tag{5.3.2}$$

where we have defined

$$F^* = A_{22} - H^* A_{12}$$

$$G^* = F^* \, H^* + (A_{21} - H^* A_{11}) \tag{5.3.3}$$

$$D^* = B_2 - H^* B_1$$

Now the implemented control is:

$$u(t) = K_1^* \, x_1(t) + K_2^* \, \hat{x}_2(t) = (K_1^* + K_2^* \, H^*) \, x_1(t) + K_2^* \, z(t) \tag{5.3.4}$$

So that we can use $u(t)$ given by (5.3.4) in (5.3.1), to yield the "compact" form of the compensator:

$$\dot{z}(t) = \hat{F}^* \, z(t) + \hat{G}^* \, x_1(t) \tag{5.3.5}$$

where we have:

$$\hat{F}^* = F^* + D^* K_2^*$$

$$\hat{G}^* = G^* + D^*(K_1^* + K_2^* H^*)$$

(5.3.6)

Now recall that:

$$x(t) \overset{\Delta}{=} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} \bar{C} \\ T \end{bmatrix} \bar{x}(t)$$

so that $x_1(t) = \bar{C} \bar{x}(t)$ is the same for all designs whatever P is, and that $x_2(t) = T \bar{x}(t)$ will depend on the choice of T.

It is then intuitively clear in view of (5.3.4), that $x_1(t)$ being the same for all designs, the quantity $K_1^* + K_2^* H^*$ should be an invariant, (i e., independent of the P transformation), although $K^* = \bar{K}^* P^{-1}$ and as such K_1^*, K_2^* depend on P, and H also. We will clearly demonstrate this result in Section 5.4.

Let us now proceed a step further. Define as we did in Chapter III, Section 3.2:

$$\zeta(t) = \begin{bmatrix} x(t) \\ z(t) \end{bmatrix}$$

then the overall optimal canonical plant is

$$\dot{\zeta}(t) = \Phi^* \zeta(t)$$

where we have from (3.2.15):

$$\Phi^* = \begin{bmatrix} A + B(K_1^* + K_2^* H^*) C & BK_2^* \\ \hat{G}^* C & \hat{F}^* \end{bmatrix}$$

Using P defined in (5.2.8), consider

$$\bar{\zeta}(t) = P^{-1} \zeta(t) = \begin{bmatrix} \bar{x}(t) \\ \bar{z}(t) \end{bmatrix}$$

we can then return to the initial given plant. Namely, we have

$$\dot{\bar{\zeta}}(t) = \bar{\Phi}^* \bar{\zeta}(t) \tag{5.3.7}$$

where $\bar{\Phi}^*$ satisfies:

$$\bar{\Phi}^* = P^{-1} \Phi^* P = \begin{bmatrix} \bar{A} + \bar{B}(K_1^* + K_2^* H^*) \bar{C} & \bar{B} K_2^* \\ \hat{G}^* \bar{C} & \hat{F}^* \end{bmatrix} \tag{5.3.8}$$

We will then implement in practice the following compensator around the given initial plant:

$$\dot{\bar{z}}(t) = \hat{F}^* \bar{z}(t) + \hat{G}^* \bar{y}(t) \tag{5.3.9}$$

$$u(t) = (K_1^* + K_2^* H^*) \bar{y}(t) + K_2^* \bar{z}(t) \tag{5.3.10}$$

Note that $\hat{F}^*$, $\hat{G}^*$ K_2^* depend on the particular similarity transformation P we have taken to go from the initial plant to the canonical plant and back. Now we have already suggested that $(K_1^* + K_2^* H^*)$ should be an invariant: in view of (5.3.10), it must be so and we will prove it soon. Another desirable property should be that $\hat{F}^*$ have the same dynamics whatever P we choose.

Define now:

$$\bar{Y}(s) = L[\bar{y}(t)]$$

$$\bar{Z}(s) = L[\bar{z}(t)] \qquad\qquad (5.3.11)$$

$$U(s) = L[u(t)]$$

where L denotes the Laplace transform; then the overall transfer function of the compensator is:

$$\Psi(s) = \frac{U(s)}{\bar{Y}(s)} = K_2^*(sI - \hat{F}^*)^{-1}\hat{G}^* + (K_1^* + K_2^* H^*) \qquad (5.3.12)$$

Clearly this transfer function should also be independent of the P matrix we have used to go to the canonical plant and back. (We insist that all the matrices in (5.3.12) depend on P).

We will now prove these desirable properties rigorously, and precisely determine the class of all optimal compensators for a given arbitrary plant.

5.4   The Class of Optimal Compensators for a Given Plant

Recall that the P matrix to go from (5.1.3) to (5.1.4) is given by:

$$P \triangleq \begin{bmatrix} \bar{C} \\ T \end{bmatrix} \tag{5.4.1}$$

where $\bar{C}$ is the mxn output matrix of full rank m of the initial plant, and

T is an arbitrary (n-m)xn matrix such that P is non-singular.

Partition now the inverse of P, namely:

$$P^{-1} = [\; \tilde{C} \quad \tilde{T} \;] \tag{5.4.2}$$

where $\tilde{C}$ is nxm and $\tilde{T}$ is nx(n-m).

$\tilde{C}$ and $\tilde{T}$ are respectively of full rank m and n-m as P^{-1} is obviously non-singular, they also both depend on the particular arbitrary T we have chosen in (5.4.1).

Now from the condition $P^{-1} P = I_n$ we obtain:

$$\tilde{C} \; \bar{C} + \tilde{T} \; T = I_n \tag{5.4.3}$$

and from the condition $P \, P^{-1} = I_n$, we have also:

$$\bar{C} \; \tilde{C} = I_m ; \quad T \; \tilde{T} = I_{n-m}$$

$$\bar{C} \; \tilde{T} = 0_{mx(n-m)} ; \quad T \; \tilde{C} = 0_{(n-m)xm} \tag{5.4.4}$$

It is now a simple matter to compute explicitly the A and B matrices for

a canonical plant, as a function of $\bar{A}$, $\bar{B}$ and the P transformation. Namely,

we have:

$$A = P \bar{A} P^{-1} = \begin{bmatrix} \bar{C} \bar{A} \tilde{C} & \bar{C} \bar{A} \tilde{T} \\ T \bar{A} \tilde{C} & T \bar{A} \tilde{T} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad (5.4.5)$$

$$B = P \bar{B} = \begin{bmatrix} \bar{C} \bar{B} \\ T \bar{B} \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \qquad (5.4.5bis)$$

the A_{ij}'s and B_i's we have used in Chapters III and IV are then readily

identified. It is clearly seen that there are as many canonical plants

as there are P matrices. However, what we proved is that for each of

them the optimal compensator was <u>unique</u>.

Similarly, we have for the initial covariance matrix:

$$\Sigma = P \bar{\Sigma} P' = \begin{bmatrix} \bar{C} \bar{\Sigma} \bar{C}' & \bar{C} \bar{\Sigma} T' \\ T \bar{\Sigma} \bar{C}' & T \bar{\Sigma} T' \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix} \qquad (5.4.6)$$

The Σ_{ij}'s we have used in Chapters III and IV are also readily identified.

Note that $\Sigma_{11} = \bar{C} \bar{\Sigma} \bar{C}'$ is the same for all transformations P, and corres-

ponds to the <u>initial output covariance matrix of the given initial plant</u>.

What we will do now is simply reconsider the equations for the unique

optimal compensator, associated with a particular canonical plant, and

use in these equations the values of A_{ij} and Σ_{ij} given by (5.4.5) and

(5.4.6).

We will only consider here the case where the initial output covariance

matrix $\Sigma_{11} = \bar{C} \; \tilde{\Sigma} \; \bar{C}'$, is non-singular.

1. Equation for W. From (4.2.13), we have that W is the unique positive definite solution of:

$$A_{22} W + WA'_{22} + \Sigma_{22} - (WA'_{12} + \Sigma'_{12}) \; \Sigma_{11}^{-1} \; (WA'_{12} + \Sigma'_{12})' = 0$$

Using (5.4.5) and (5.4.6), we obtain:

$$T \; \bar{A} \; \tilde{T} \; W + W \; \tilde{T}' \bar{A}' T' + T \; \bar{\Sigma} \; T - (W \; \tilde{T}' \bar{A}' \bar{C}' + T \; \bar{\Sigma} \; \bar{C}')(\bar{C} \; \bar{\Sigma} \; \bar{C}')^{-1}$$

$$(W \; \tilde{T}' \bar{A}' \bar{C}' + T \; \bar{\Sigma} \; \bar{C}')' = 0 \qquad\qquad (5.4.7)$$

Define now

$$\mathcal{W} = \tilde{T} \; w \; \tilde{T}' \qquad\qquad (5.4.8)$$

Note that $\mathcal{W}$ is nxn. A left multiplication of (5.4.7) by $\tilde{T}$ and a right multiplication by $\tilde{T}'$ yields, after use of (5.4.3), the following equation for $\mathcal{W}$

$$\bar{A} \; \mathcal{W} + \mathcal{W} \; \bar{A}' + \bar{\Sigma} - (\mathcal{W} \; \bar{A}' + \bar{\Sigma}) \; \bar{C}' \; (\bar{C} \; \bar{\Sigma} \; \bar{C}')^{-1} \; \bar{C} \; (\mathcal{W} \; \bar{A}' + \bar{\Sigma})' = 0 \quad (5.4.9)$$

Note that this equation involves only quantities related to the <u>initial</u> given plant. As a conclusion, $\mathcal{W}$ does not depend whatsoever on P. This is our first invariant, and $\mathcal{W}$ really characterizes the given plant $(\bar{A}, \bar{B}, \bar{C})$. All the W's for the different canonical plants will then be related to this $\mathcal{W}$ by (5.4.8).

According to the fact that we assumed the pair $(\bar{A}, \bar{C})$ to be completely controllable, and also because $\bar{C} \bar{\Sigma} \bar{C}'$ is positive definite, it can be shown that (5.4.9) has always at least one positive definite solution W.

2. **Equation for H***. We will now consider:

$$H^* = (W A_{12}' + \Sigma_{12}') \Sigma_{11}^{-1} \tag{5.4.10}$$

Using (5.4.5) and (5.4.6), we have:

$$H^* = (W \tilde{T}'\bar{A}'\bar{C}' + T \bar{\Sigma} \bar{C}')(\bar{C} \bar{\Sigma} \bar{C}')^{-1} \tag{5.4.11}$$

Consider now $\tilde{T} H^*$, namely, we have

$$\tilde{T} H^* = (\tilde{T} W \tilde{T}'\bar{A}'\bar{C}' + \tilde{T} T \bar{\Sigma} \bar{C}')(\bar{C} \bar{\Sigma} \bar{C}')^{-1}$$

Using (5.4.8) and (5.4.3), it comes immediately:

$$\tilde{T} H^* = (W \bar{A}'\bar{C}' + \bar{\Sigma} \bar{C}')(\bar{C} \bar{\Sigma} \bar{C}')^{-1} - \tilde{C} \tag{5.4.12}$$

Define now

$$H^* = (W \bar{A}' + \bar{\Sigma}) \bar{C}' (\bar{C} \bar{\Sigma} \bar{C}')^{-1} \tag{5.4.13}$$

we see immediately that H^* (an nxm matrix) is a second invariant as it involves only terms related to the given initial plant.

All the different H* for different canonical plants will then be related by

$$\tilde{T} \, H^* = H^* - \tilde{C} \qquad (5.4.14)$$

We can now proceed and express all the matrices for the optimal compensator in terms of W, H^* and $\tilde{T}$. We will then have precisely determined the class of all optimal compensators.

3. <u>Equation for F*</u>. Consider now $F^* = A_{22} - HA^*_{12}$, which governs the dynamics of the error. We have from (5.4.5) and (5.4.6)

$$F^* = T \, \bar{A} \, \tilde{T} - H^* \, \bar{C} \, \bar{A} \, \tilde{T} \qquad (5.4.15)$$

A left multiplication by $\tilde{T}$ and the use of (5.4.14) and (5.4.3) yields

$$\tilde{T} \, F^* = (\bar{A} - H^* \, \bar{C} \, \bar{A}) \, \tilde{T} \qquad (5.4.16)$$

Define now a new nxn invariant matrix:

$$\bar{F}^* = \bar{A} - H^* \, \bar{C} \, \bar{A} = (I_n - H^* \, \bar{C}) \, \bar{A} \qquad (5.4.17)$$

then recalling that $\tilde{T}$ has full rank (n-m), we can make use of the <u>pseudo-inverse</u> matrix (see Zadeh and Desoer [28]) to write:

$$F^* = \tilde{T}^\dagger \, \bar{F}^* \, \tilde{T} \qquad (5.4.18)$$

where we have defined

$$\tilde{T}{}^{\dagger} = (\tilde{T}{}' \; \tilde{T})^{-1} \; \tilde{T}{}' \tag{5.4.19}$$

and where F^* is given by (5.4.17).

We have then precisely defined the structure of all the F^* matrices as a function of the invariant F^* and the transformation $\tilde{T}$. Namely, given P, $\tilde{T}$ is automatically fixed and F^* follows.

Now recall that $\bar{C}$ and $\tilde{T}$ are respectively of full rank m and n-m. From (5.4.4), we have $\bar{C} \; \tilde{T} = 0$ which implies that $R(\tilde{T})$ is contained in $N(\bar{C})$. Since the null space of $\bar{C}$ has dimension n-m, we see that the range space of $\tilde{T}$ spans it exactly, whatever P matrix we choose (i.e., whatever $\tilde{T}$). We will then say that the dynamics of $F^* = \tilde{T}{}^{\dagger} \; F^* \; \tilde{T}$ are those dynamics of F^* which are associated with the null space of $\bar{C}$, and as such will be the same whatever T we have. For more details on this concept, see [27], and also [29], Appendix II, about pseudo-inverses.

We then have the desirable expected property that all optimal compensators in the class will have the same dynamics.

4. <u>Equation for $\hat{F}^* = F^* + D^* K^*_2$</u>. Recall from (5.4.5bis) that $B_1 = \bar{C} \; \bar{B}$ and $B_2 = T \; \bar{B}$; recall also that $K^* = \bar{K}^* \; P^{-1}$ leading to $K^*_1 = \bar{K}^* \; \bar{C}$ and $K^*_2 = \bar{K}^* \; \tilde{T}$; we then obtain:

$$D^* \; K^*_2 \doteq (B_2 - H^* \; B_1) \; K^*_2$$

$$D^* \; K^*_2 = (T \; \tilde{B} - H^* \; \bar{C} \; \bar{B}) \; \bar{K}^* \; \tilde{T} \tag{5.4.20}$$

A left multiplication of (5.4.20) by $\tilde{T}$ and the use of (5.4.3, 14) yields

$$\tilde{T} \; D^* \; K_2^* = (\bar{B} - H^* \; \bar{C} \; \bar{B}) \; \bar{K}^* \; \tilde{T} \tag{5.4.21}$$

Define a fourth nxr invariant matrix:

$$\mathcal{D}^* = \bar{B} - H^* \; \bar{C} \; \bar{B} = (I_n - H^* \; \bar{C}) \; \bar{B} \tag{5.4.22}$$

Then using the pseudo-inverse of $\tilde{T}$, we have in passing:

$$D^* = \tilde{T}{}^\dagger \mathcal{D}^* \tag{5.4.23}$$

and for our present interest:

$$D^* \; K_2^* = \tilde{T}{}^\dagger \; \mathcal{D}^* \; \bar{K}^* \; \tilde{T} \tag{5.4.24}$$

Note that $\bar{K}^*$ is directly related to the initial given plant. We have now

$$\hat{F}^* = \tilde{T}{}^\dagger \; (F^* + \mathcal{D}^* \; \bar{K}^*) \; \tilde{T} \tag{5.4.24}$$

so that we can define

$$\hat{F}^* = F^* + \mathcal{D}^* \; \bar{K}^* \tag{5.4.25}$$

or using (5.4.17) for F^*, (5.4.22) for $\mathcal{D}^*$

$$\hat{F}* = (I_n - H* \bar{C})(\bar{A} + \bar{B}\bar{K}*) \tag{5.4.26}$$

(Recall that $\bar{A} + \bar{B}\bar{K}*$ governs the dynamics of the optimal given plant assuming complete feedback.)

We can give now the general relation between all the matrices $\hat{F}*$

$$\hat{F}* = \tilde{T}^{\dagger} \hat{F}* \tilde{T}$$

Note that, as mentioned above for $F*$, the dynamics of $\hat{F}*$ are those dynamics of $\hat{F}*$ associated with $N(\bar{C})$ and do not depend on $\tilde{T}$.

5. Equation for $K_1^* + K_2^* H*$

It was intuitively clear from (5.3.10) that this quantity should be invariant. We can now prove it very simply.

$$K_1^* + K_2^* H* = \bar{K}* \tilde{C} + \bar{K}* \tilde{T} H* \tag{5.4.27}$$

Now from (5.4.14): $\tilde{T} H* = H* - \tilde{C}$ so that (5.4.27) becomes:

$$K_1^* + K_2^* H* = \bar{K}* \tilde{C} + \bar{K}* H* - \bar{K}* \tilde{C}$$

that is, we have the fifth rxm invariant matrix:

$$K_1^* + K_2^* H* = \bar{K}* H* \tag{5.4.18}$$

6. <u>Equation for $\hat{G}^* = G^* + D^* (K_1^* + K_2^* H^*)$</u>. It remains only now to determine the value of $\hat{G}^*$ as a function of some other expected invariant. We will first evaluate

$$G^* = F^* H^* + (A_{21} - H^* A_{11}) \qquad (5.4.29)$$

Note first that:

$$A_{21} - H^* A_{11} = T \ \bar{A} \ \tilde{C} - H^* \ \bar{C} \ \bar{A} \ \tilde{C}$$

so that a left multiplication by $\tilde{T}$ and the use of the pseudo-inverse $\tilde{T}{}^\dagger$ yields:

$$A_{21} - H^* A_{11} = \tilde{T}{}^\dagger \ F^* \ \tilde{C} \qquad (5.4.30)$$

where F^* is given by (5.4.17).

It is easy to see now that

$$G^* = \tilde{T}{}^\dagger \ F^* \ \tilde{T} \ H^* + \tilde{T}{}^\dagger \ F^* \ \tilde{C}$$

so that using (5.4.14)

$$G^* = \tilde{T}{}^\dagger \ F^* \ H^* = \tilde{T}{}^\dagger G^* \qquad (5.4.31)$$

where we have defined the new nxm invariant matrix

$$G* = F* \; H* = (I_n - H* \; \bar{C}) \; \bar{A} \; H* \tag{5.4.32}$$

It is now a simple matter to see that

$$\hat{G}* = \tilde{T}{}^{\dagger} \; G* + \tilde{T}{}^{\dagger} \; D* \; \bar{K}* \; H*,$$

namely we have

$$\hat{G}* = \tilde{T}{}^{\dagger} \; \hat{G}* \tag{5.4.33}$$

where we have defined the nxm matrix:

$$\hat{G}* = G* + D* \; \bar{K}* \; H* \tag{5.4.34}$$

or equivalently in view of (5.4.36)

$$\hat{G}* = F* \; H* + D* \; \bar{K}* \; H* \tag{5.4.35}$$

which can be rewritten using (5.4.22) and (5.4.26)

$$\hat{G}* = (I_n - H* \; \bar{C}) (\bar{A} + \bar{B}\bar{K}*) \; H* = \hat{F}* \; H*$$

We insist that all matrices denoted by <u>script</u> letters are directly related

to the initial <u>given</u> plant, as well as those which have overbars. We have

then given a complete picture of the structure of the gains for the class

of all optimal compensators.

We will conclude by showing that as we expected the overall transfer function for all compensators in the class, is indeed the same. Namely, (5.3.12) becomes:

$$\Psi(s) = \bar{K}* \tilde{T}(sI_{n-m} - \tilde{T}{}^\dagger \hat{F}* \tilde{T})^{-1} \tilde{T}{}^\dagger \hat{G}* + \bar{K}* H* \tag{5.4.36}$$

from the fact that $\tilde{T}{}^\dagger \tilde{T} = I_{n-m} = \tilde{T}{}^\dagger I_n \tilde{T}$, we can rewrite (5.4.36) as

$$\Psi(s) = \bar{K}* \tilde{T}[\tilde{T}{}^\dagger (sI_n - \hat{F}*) \tilde{T}]^{-1} \tilde{T}{}^\dagger \hat{G}* + \bar{K}* H* \tag{5.4.37}$$

It can be shown ([27], [29]) that $\Phi(s)$ defined as:

$$\Phi(s) = \tilde{T}[\tilde{T}{}^\dagger (sI_n - \hat{F}*) \tilde{T}]^{-1} \tilde{T}{}^\dagger$$

is independent of $\tilde{T}$, and corresponds as we previously mentioned to those dynamics of $\hat{F}*$ associated with the null space of $\bar{C}$. We then have

$$\Psi(s) = \bar{K}* \Phi(s) \hat{G}* + \bar{K}* H* \tag{5.4.38}$$

This last formula is certainly the most important result in this thesis. We have an explicit formula for the optimal compensator transfer function, involving uniquely quantities related directly to the given arbitrary plant. This result as well as the entire Section 5.4 is original.

The results in [7] duplicate only one special case of Chapter IV, namely when E $\{x_o\}$ = 0, z(0) = 0; no explicit general solution is given when the

plant is not in the canonical form assumed in Chapter IV. We have then

rederived the partial results of [7] by a direct method, and more important,

completely solved the problem of finding the optimal compensator based on

a minimal order Luenberger observer, for any arbitrary given plant.

A numerical example will now be provided.

5.5 An Example

Consider the following non-canonical, completely controllable, completely observable, initial plant:

$$
\begin{bmatrix} \dot{\bar{x}}_1(t) \\ \dot{\bar{x}}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} \bar{x}_1(t) \\ \bar{x}_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \tag{5.5.1}
$$

$$
\bar{y}(t) = [1 \quad 1] \begin{bmatrix} \bar{x}_1(t) \\ \bar{x}_2(t) \end{bmatrix} \tag{5.5.2}
$$

If we assume complete state feedback and take a cost given by (2.1.3), where:

$$
\bar{Q} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \bar{r} = 1 \tag{5.5.3}
$$

then the optimal control is given by

$$
u^*(t) = \bar{K}^* \, \bar{x}(t) = [-2 \quad -1] \begin{bmatrix} \bar{x}_1(t) \\ \bar{x}_2(t) \end{bmatrix} \tag{5.5.4}
$$

Defining

$$
E\{\bar{x}_o \, \bar{x}_o'\} = \bar{\Sigma}_o = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \tag{5.5.5}
$$

the optimal cost is in this case:

$$
\hat{\bar{J}}_1^* = \text{tr}\,(\bar{\Pi}\,\bar{\Sigma}_o) = 12 \tag{5.5.6}
$$

where $\bar{\Pi}$ is given by:

$$\bar{\Pi} = \begin{bmatrix} 6 & 2 \\ 2 & 1 \end{bmatrix} \qquad (5.5.7)$$

Now, in practice, we don't have access to the complete state, so that we want to design an optimal compensator for this plant.

A. Consider now a first similarity transformation P_1,

$$P_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad P_1^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \qquad (5.5.8)$$

such that the initial plant now take the following canonical form:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) \qquad (5.5.9)$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \qquad (5.5.10)$$

$$\Sigma_o = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} \quad \text{and} \quad K^* = \bar{K}^* P_1^{-1} = \begin{bmatrix} -2 & 1 \end{bmatrix} \qquad (5.5.11)$$

It is immediately seen using Chapter IV that the optimal compensator for this canonical plant is uniquely determined by $h^* = 2 - \sqrt{2}$, leading to:

$$\dot{z}(t) = -\sqrt{2}\, z(t) + (2 - 2\sqrt{2})\, x_1(t) + (\sqrt{2} - 1)\, u(t) \qquad (5.5.12)$$

where $u(t)$ is given by:

$$u(t) = - \sqrt{2} \, x_1(t) + z(t) \tag{5.5.13}$$

so that the compact equation for $z(t)$ is

$$\dot{z}(t) = - z(t) - \sqrt{2} \, x_1(t) \tag{5.5.14}$$

Now the overall canonical system will be:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{z}(t) \end{bmatrix} = \begin{bmatrix} -\sqrt{2} & -1 & 1 \\ -\sqrt{2} & -2 & 1 \\ -\sqrt{2} & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ z(t) \end{bmatrix} \tag{5.5.15}$$

so that we can return to the initial plant by defining:

$$P_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad P_1^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.5.16}$$

We then obtain:

$$\begin{bmatrix} \dot{\bar{x}}_1(t) \\ \dot{\bar{x}}_2(t) \\ \dot{\bar{z}}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -\sqrt{2} & -2 - \sqrt{2} & 1 \\ -\sqrt{2} & -\sqrt{2} & -1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \bar{z}(t) \end{bmatrix} \tag{5.5.17}$$

leading to the practical implementation for the initial plant:

$$\dot{\bar{z}}(t) = - \bar{z}(t) - \sqrt{2} \, \bar{y}(t) \tag{5.5.18}$$

$$u(t) = - \sqrt{2}\, \bar{y}(t) + \bar{z}(t) \qquad (5.5.19)$$

whose transfer function is:

$$\Psi_1(s) = \frac{U(s)}{\bar{Y}(s)} = - \sqrt{2} - \frac{\sqrt{2}}{s+1} \qquad (5.5.20)$$

It can easily be shown that, as expected:

$$\Delta \hat{J}_2^* = \Delta \hat{\bar{J}}_2^* = 5\sqrt{2} - 7 \approx 71.10^{-3} \qquad (5.5.21)$$

B. Consider now a second similarity transformation P_2:

$$P_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad P_2^{-1} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix} \qquad (5.5.22)$$

such that the initial plant now takes the following different canonical form:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} -0.5 & 0.5 \\ 1.5 & -1.5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} u(t) \qquad (5.5.23)$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

$$\Sigma_0 = \begin{bmatrix} 5 & -1 \\ -1 & 1 \end{bmatrix} \qquad K^* = \bar{K}^* \, P_2^{-1} = \begin{bmatrix} -1.5 & -0.5 \end{bmatrix}$$

We will then implement the optimal compensator uniquely determined by $h* = -3 + 2\sqrt{2}$, leading to

$$\dot{z}(t) = -\sqrt{2}\, x_1(t) + 4(\sqrt{2} - 1)\, x_1(t) + 2(1 - \sqrt{2})\, u(t) \qquad (5.5.24)$$

$$u(t) = -\sqrt{2}\, x_1(t) - 0.5\, z(t) \qquad (5.5.25)$$

so that the compact equation for $z(t)$ is:

$$\dot{z}(t) = -z(t) + 2\sqrt{2}\, x_1(t) \qquad (5.5.26)$$

The overall canonical system will now be:

$$
\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{z}(t) \end{bmatrix}
=
\begin{bmatrix} -0.5 - \sqrt{2} & 0.5 & -0.5 \\ 1.5 + \sqrt{2} & -1.5 & 0.5 \\ 2\sqrt{2} & 0 & -1 \end{bmatrix}
\begin{bmatrix} x_1(t) \\ x_2(t) \\ z(t) \end{bmatrix}
\qquad (5.5.27)
$$

so that we can return to the initial plant by defining P_2 as:

$$
P_2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\qquad
P_2^{-1} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & -0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\qquad (5.5.28)
$$

We then obtain:

$$
\begin{bmatrix} \dot{\bar{x}}_1(t) \\ \dot{\bar{x}}_2(t) \\ \dot{\bar{z}}(t) \end{bmatrix}
=
\begin{bmatrix} 0 & 1 & 0 \\ -\sqrt{2} & -2 -\sqrt{2} & -0.5 \\ 2\sqrt{2} & 2\sqrt{2} & -1 \end{bmatrix}
\begin{bmatrix} \bar{x}_1(t) \\ \bar{x}_2(t) \\ \bar{z}(t) \end{bmatrix}
\qquad (5.5.29)
$$

leading to the practical implementation for the initial plant:

$$\dot{\bar{z}}(t) = -\bar{z}(t) + 2\sqrt{2}\,\bar{y}(t) \qquad\qquad (5.5.30)$$

$$u(t) = -\sqrt{2}\,\bar{y}(t) - 0.5\,\bar{z}(t) \qquad\qquad (5.5.31)$$

whose transfer function is:

$$\Psi_2(s) = \frac{U(s)}{\bar{Y}(s)} = -\sqrt{2} - \frac{\sqrt{2}}{s+1} \qquad\qquad (5.5.32)$$

It is immediately seen that the increase in cost due to the implementation of this compensator is still given by (5.5.21).

It is clearly seen that

1. $\Psi_1(s) = \Psi_2(s)$ see (5.5.20) and (5.5.32)

2. $(k_1^* + k_2^*\,h^*)_1 = (k_1^* + k_2^*\,h^*)_2 = -\sqrt{2}$

as expected.

We will now make use of Section 5.4 and compute the invariant quantities for the initial non-canonical given plant.

From (5.4.9), we have

$$W = \begin{bmatrix} w_1 & 7 + \sqrt{50} \\ 7 + \sqrt{50} & -7 + \sqrt{50} \end{bmatrix} \qquad\qquad (5.5.33)$$

where for positive definiteness w_1 must be greater than $(7 + \sqrt{50})^3$.

Now it follows that:

$$H^* = \begin{bmatrix} -1 & -\sqrt{2} \\ 2 & -\sqrt{2} \end{bmatrix} \qquad\qquad (5.5.34)$$

and from (5.4.17) that

$$F* = \begin{bmatrix} 0 & -\sqrt{2} \\ 0 & -\sqrt{2} \end{bmatrix} \tag{5.5.35}$$

from (5.4.22) we have also

$$\mathcal{D}* = \begin{bmatrix} 1 + \sqrt{2} \\ -1 + \sqrt{2} \end{bmatrix} \tag{5.5.36}$$

So that from (5.5.4), (5.5.35), (5.5.36) and (5.4.25)

$$\hat{F}* = \begin{bmatrix} -2 -2\sqrt{2} & -1 -2\sqrt{2} \\ 2 -2\sqrt{2} & 1 -2\sqrt{2} \end{bmatrix} \tag{5.4.37}$$

Now, we have from (5.4.32)

$$G* = \begin{bmatrix} 2 -2\sqrt{2} \\ 2 -2\sqrt{2} \end{bmatrix}$$

So that from (5.4.35), we obtain

$$\hat{G}* = \begin{bmatrix} 3\sqrt{2}(\sqrt{2} + 1) \\ 3\sqrt{2}(\sqrt{2} - 1) \end{bmatrix}$$

Noting that $\bar{K}* = [-2 \quad -1]$ and $\bar{K}* H* = 3\sqrt{2}$, it follows that $\psi(s)$ given by (5.4.38) is the same as $\psi(s)$ given by (5.5.32) or (5.5.20), namely

$$\psi(s) = -\sqrt{2} - \frac{\sqrt{2}}{s + 1}$$

It is also possible from the matrices P_1^{-1} and P_2^{-1} to pick the $\tilde{T}_i$'s $i = 1, 2$

and to verify that all the formulas in Section 5.5 hold.

Important Remark. Note that from a practical computational standpoint, it is not advantageous to work with the script matrices, related to the initial given plant. As an example, the Riccati-type equation for $\mathcal{W}$ is of order n, instead of n-m for W.

Knowing that all the compensators will lead to the same optimal transfer function, it is then computationally more efficient to pick a P, go to a canonical plant, compute the canonical optimal compensator, and go back to the initial plant by using the $\mathcal{P}$ matrix defined in (5.2.8).

CHAPTER VI

CONCLUSION

We have revisited and solved in this thesis the problem of finding the optimal n-m order compensator based on a minimal-order Luenberger observer. For any given completely controllable and completely observable plant, we have shown that there exists a class of optimal compensators, specified by a unique optimal transfer function. Moreover, all compensators within this class lead to the same cost and the same overall dynamics for the augmented system plant + compensator: the optimal dynamics for complete state feedback, combined with the optimal dynamics of the minimal order Luenberger observer. The interesting result was that a generalization of the separation property holds for this design, (i.e., design of the observer part of the compensator is completely independent of the controller part).

In case the plant is specified by a canonical output matrix (5.1.2), we have seen that the optimal compensator is unique and that its realization is specified by Equations (4.2.11), (4.2.13), (3.2.10), (3.2.11).

Consider now the linear stochastic system, with perfect measurements:

$$\dot{x}(t) = A\, x(t) + \begin{bmatrix} I_m \\ 0 \end{bmatrix} v(t)$$

$$y(t) = C\, x(t)$$

where $x(t)$, A are partitioned as in (3.2.1) and where C has the canonical form (5.1.2). Assume that $v(t)$ is a white noise process specified by:

$$E\{v(t)\} = 0 \qquad E\{v(t)\ v'(\tau)\} = \Sigma_{11}\ \delta(t - \tau)$$

for all t and τ. It is interesting to note that Equations (4.2.13) and (4.2.11) are exactly the equations for the error covariance matrix and gains associated with the Bryson-Johansen filter we would implement to obtain an estimate of $x_2(t)$ from the perfect measurements $y(t) = x_1(t)$. This justifies completely the approach taken by Rom and Sarachik in [7]. Section 5.4 gives general equations for the case where the plant is not in the canonical form assumed in Section 4.2. Notice also that $\Sigma_{11} = \bar{C}\ \bar{\Sigma}\ \bar{C}' > 0$ is equivalent to the condition that the Bryson-Johansen filter does not contain differentiators. For more details on the Bryson-Johansen filter, see [22] and also Jazwinski [30], example 7.14, page 228.

We will now indicate a possible extension of the design philosophy we used towards compensators of lower dimension.

Given the optimal control

$$u^*(t) = K^*\ x(t) = K_1^*\ x_1(t) + K_2^*\ x_2(t)$$

for the linear regulator problem (see (3.2.1) for the partition of $x(t)$) this thesis was concerned mainly with the reconstruction of $x_2(t)$ from $y(t) = x_1(t)$ in order to implement the control

$$u(t) = K_1^*\ x_1(t) + K_2^*\ \hat{x}_2(t)$$

Having only m available outputs, our compensator was of fixed order n-m.

Now it might be very interesting and rewarding to consider the direct

global estimation of $K_2^* \, x_2(t)$. We see immediately that, as in this case

we are trying to observe linear functionals of the state, the dimensions

of the required observer may be far less than n-m. In particular, when

u(t) is a scalar control, Luenberger [25] has shown that any linear func-

tional of the state can be observed with an observer of order $\nu - 1$. Here

ν (see [25]) is the observability index of the plant, defined as the least

positive integer such that the matrix:

$$[c', \, A' \, c', \, .. \, , \, (A')^{\nu - 1} \, c']$$

has full rank n. Since for any completely observable system $\nu-1 \leq$ n-m,

and for many systems $\nu-1$ is in fact far less than n-m, this approach could

greatly reduce the order of the compensator. When u(t) is vector-valued,

it is also reasonable to anticipate the existence of a lower-order compen-

sator. The problem is then to use the approach taken in this thesis to

find the optimal dynamics for this reduced order observer. We will recall

the three steps of our approach:

- Given an arbitrary plant, go to a canonical form by similarity

transformation. (Chapter II)

- Work out the problem for this new plant. (Chapters III, IV)

- Go back to the initial plant. (Chapter V)

We emphasize once again the importance of working with a canonical

plant if one wants to understand clearly the internal structure of the

compensator (see the formulation by Newman [5], [6] to appreciate the

relevance of this remark).

REFERENCES

1. M. Athans and P. F. Falb, "Optimal control: an introduction to the theory and its applications," McGraw-Hill Book Co., New York, 1966.

2. D. G. Luenberger, "An Introduction to Observers," IEEE Trans. Aut. Contr., Vol. AC-16, No. 6, pp. 596-602, Dec. 1971.

3. W. M. Wonham, "On pole assignment in multi-input controllable linear systems," IEEE Trans. Aut. Contr., Vol. AC-12, pp. 660-665, Dec. 1967.

4. Y. O. Yüksel and J. J. Bongiorno, "Observers for linear multivariable systems with applications," IEEE Trans. Aut. Contr., Special Issue LOG Problem, pp. 603-613, Dec. 1971.

5. M. M. Newman, "Specific optimal control of the linear regulator using a dynamic controller based on the minimal-order Luenberger observer," Int. J. Contr., Vol. 12, No. 1, pp. 33-48, 1970.

6. M. M. Newman, "Optimal and sub-optimal control using an observer when some of the state variables are not measurable," Int. J. Contr., Vol. 9, pp. 281-290, 1970.

7. D. B. Rom and P. E. Sarachik, "The design of optimal compensators for linear constant systems with inaccessible states," IEEE Trans. Aut. Contr., (short papers) Vol. AC-18, pp. 509-512, Oct. 1973.

8. J. J. Bongiorno and D. C. Youla, "On observers in multivariable control systems," Intr. J. Contr., Vol. 8, No. 3, pp. 221-243, 1968.

9. J. J. Bongiorno and D. C. Youla, 'Discussion of "On Observers in multi-variable control systems,"' Intr. J. Contr., Vol. 12, pp. 183-190, July 1970.

10. M. Siswosudarmo, "Canonical Observers in Linear Control Systems," Ph.D. Dissert., Polytechnic Inst. Brooklyn, Brooklyn, N. Y., 1969.

11. D. L. Kleinman, "Suboptimal design of linear regulator systems subject to computer storage limitations," Electronic Systems Laboratory, Report 297, M. I. T. (1967).

12. W. S. Levine, "Optimal output feedback controllers for linear systems," Ph.D. Dissertation, M. I. T., Cambridge, 1969.

13. R. Bellman, "Introduction to Matrix Analysis," McGraw-Hill Book Co., New York, 1960.

14. R. E. Kalman, "Contribution to the theory of optimal control," Bol. Soc. Mat. Mexico, pp. 102-119, 1960.

15. --- "When is a linear system optimal?", J. Basic Engineering, (ASME Trans.), Vol. 86, pp. 1-10, 1964.

16. ---, Ho and Narendra, "Controllability of linear dynamical systems," Contributions to differential equations, Vol. 1, 1962.

17. ---, "Mathematical description of linear dynamical systems," J. SIAM on Contr., Ser. A, Vol. 1, No. 2, pp. 152-192, 1963.

18. W. S. Levine and M. Athans, "On the determination of the optimal constant output-feedback gains for linear multivariable systems," IEEE Trans. Aut. Contr., Vol. AC-15, pp. 44-48, Feb. 1970.

19. R. E. Kalman and R. S. Bucy, "New results in filtering and prediction theory, "J. Basic Engineering, (ASME Trans.), Vol. 83, pp. 95-108, March 1961.

20. E. Tse, "On the optimal control of stochastic linear systems," IEEE Trans. Aut. Contr., Vol. AC-16, pp. 776-785, Dec. 1971.

21. W. M. Wonham, "On the separation theorem control," SIAM J. Contr., Vol. 6, No. 2, pp. 312-328, 1968.

22. A.Bryson and D. Johansen, "Linear filtering for time-varying systems using measurements containing colored noise," IEEE Trans. Aut. Contr., Vol. AC-10, pp. 4-10, Jan. 1965.

23. R. Bucy, "Optimal filtering for correlated noise," J. Math. Anal. Appl., No. 20, pp. 1-8, 1967.

24. D. G. Luenberger, "Observing the state of a linear system," IEEE Trans. Mil. Elec., Vol. MIL-8, pp. 74-80, April 1974.

25. ---, "Observers for multivariable systems," IEEE Trans. Aut. Contr., Vol. AC-11, pp. 190-197, April 1966.

26. W. S. Levine, T. L. Johnson, M. Athans, "Optimal limited state variable feedback controllers for linear systems," IEEE Trans. Aut. Contr., Vol. AC-16, Dec. 1971.

27. M. T. Li, "On output feedback stabilizability of linear systems," IEEE Trans. Aut. Contr. (Tech. notes and corr.) Vol. AC-17, pp. 408-410, June 1972.

28. L. A. Zadeh and C. A. Desoer, "Linear Systems Theory," McGraw Hill Book Co., New York, 1960.

29. M. Aoki, "Optimization of Stochastic Systems," Academic Press, New York, 1967.

30. A. H. Jazwinski, "Stochastic Processes and Filtering Theory," Academic Press, New York, 1970.

31. R. E. Kalman, "New Methods in Wiener Filtering Theory," <u>Proc. Symp.</u>
     <u>Eng. Appl. Random Function Theory and Probability</u> (J. L. Bogdanoff and
     F. Kozin, eds.) Wiley, New York, 1963.