November, 1971

Report ESL-R-462
M.I.T. DSR Projects 76265 and 72486
Grants NGL-22-009-124 and AFOSR-70-1941

ON STOCHASTIC CONTROL AND OPTIMAL MEASUREMENT STRATEGIES

Leslie C. Kramer

Electronic Systems Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139

Department of Electrical Engineering

November, 1971 Report ESL-R-463

ON STOCHASTIC CONTROL AND OPTIMAL MEASUREMENT STRATEGIES

by

LESLIE CRAIG KRAMER

This report is based on the unaltered thesis of Leslie C. Kramer submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology, in October, 1971. This research was carried out at the Decision and Control Sciences Group of the Electronic Systems Laboratory, M. I. T. with partial support provided by AFOSR under grant AFOSR-70-1941 and by NASA grant NGL-22-009-124.

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

ON STOCHASTIC CONTROL AND OPTIMAL MEASUREMENT STRATEGIES

by

LESLIE CRAIG KRAMER

S.B., Massachusetts Institute of Technology
1967

S.M., Massachusetts Institute of Technology
1968

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

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
October, 1971

Signature of Author _Leslie C. Kramer_
October 29, 1971

Certified by _Michael Athans_
Thesis Supervisor

Accepted by ______________________
Chairman, Departmental Committee on Graduate Students

ON STOCHASTIC CONTROL AND OPTIMAL MEASUREMENT STRATEGIES

by

LESLIE CRAIG KRAMER

Submitted to the Department of Electrical Engineering on October 29, 1971 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

This thesis is concerned with the control of stochastic dynamic systems, with particular emphasis on those which have the property that one can influence the quality or nature of the measurements which are made to effect control. Four main areas are discussed.

First, the meaning of stochastic optimality and the means by which dynamic programming may be applied to solve a combined control/measurement problem is discussed. Second, a technique is described by which it is possible to apply deterministic methods, specifically the Minimum Principle, to the study of stochastic problems. Third, the methods described are applied to linear systems with Gaussian disturbances to study the structure of the resulting control system. A useful separation property is shown to hold for linear systems with quadratic cost criteria and Gaussian noise. Fourth, several applications are considered.

Thesis Supervisor: Michael Athans
Title: Associate Professor of Electrical Engineering

ACKNOWLEDGEMENT

I would like to express my deep appreciation for the guidance and counsel given me by Professor Michael Athans during the development of this work and for several years before. It has been a privilege to work with this man. I would also like to thank my thesis readers, Professors Sanjoy Mitter and Arthur Baggeroer for their helpful comments and suggestions.

I am also indebted to those with whom I have worked and from whom I have learned during my years with the Electronic Systems Laboratory: Professors Leonard Gould, Jan Willems and Ian Rhodes as well as the many fellow students with whom I have been associated. The help of the non-teaching staff of E.S.L. must also be acknowledged, with special thanks to Elyse Wolf for much preliminary typing, to the entire Publications Department for typing the final draft, and to John Sherman, Norman Darling, and Arthur Giordani for the drawings.

I would like to acknowledge the financial support I have received from the Electrical Engineering Department in return for teaching: more of an enjoyment than a duty.

I would finally like to thank my wife Judith for her help and her patience.

This research was carried out at the Decision and Control Sciences Group with partial support provided by AFOSR under grant AF-AFOSR-70-1941 and by NASA/AMES under grant NASA NGL-22-009-124.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

TABLE OF CONTENTS (Cont'd)

CHAPTER I

INTRODUCTION AND PROBLEM STATEMENT

1.1 General Perspective

One of the basic motivating forces behind the development of the theory and practice of automatic control systems has been the goal of providing acceptable system performance in the presence of uncertainty. This uncertainty can be of two types: the designer may be ignorant of the true nature of the system he is trying to control, or there may be influences on a (known) system which are of an uncertain nature.* The former type of uncertainty leads to the problem of system identification; the latter situation is generally referred to as a problem in stochastic control. It is this second class of problems to which this thesis is devoted.

The earliest feedback control systems (for example, those used to govern the operation of the water clocks and windmills of antiquity) were built to reduce the undesirable effects of unknown influences on these systems. (M. 1)** These early systems were designed using a trial and error method inspired by physical intuition. One could usually determine which unknown factors would influence system behavior (say, water pressure or wind velocity and direction in the above examples) and try to design into the system some form of compensation to make the overall behavior insensitive to these factors.

---

*Of course, a particular problem might be subject to both types of uncertainty.

**Parentheses denote references collected at the back of this work.

Many modern control systems are also built with the goal of making the overall behavior insensitive to parameter fluctuations: Chemical processes must be insensitive to variations in the composition of raw materials, moving bodies (automobiles or aircraft) must be insensitive to disturbing forces (bumps in the road or wind gusts), and so forth. The obvious difference between the design of the control systems built before about 1940 and those built since then is the large amount of mathematical effort that has been applied in the design of modern systems.

According to the most modern viewpoint, the dynamical system to be controlled is modeled by a set of differential equations or difference equations* in which the uncertain quantities play the role of inputs or drives to the equations of motion. The designer is able to manipulate certain other inputs so as to minimize in some sense the "damage" caused by the uncertain influences. To complete the mathematical structure and allow an analysis which determines the control commands, one must specify how the uncertainties are to be modeled mathematically and specify the performance criterion to be used in selecting the control.

The uncertain quantities can be modeled by using the structure of probability theory or by a set membership technique. In the probabilistic framework, one assumes that there is an underlying probability space $(\Omega, \mathcal{B}, p)$ and that the uncertain quantities can be viewed as random variables or random processes defined on this measure space.** A suitable performance

---

*In fact, even more general representations are possible, but these are not considered here.

**It is assumed that the reader is familiar with axiomatic probability theory from the point of view of measures. The associated terminology will be used freely with few definitions given formally. Important results will be referenced.

criterion is then defined in such a way that it also is a random variable on $(\Omega, \mathcal{A}, p)$, and one chooses the control in order to, say, minimize the criterion of performance in some probabilistic sense. If, on the other hand, one has insufficient information to establish a probabilistic framework for the uncertain quantities, one might perhaps know that these quantities are constrained to take their values in some given set. This point of view has been recently developed, for example, by Bertsekas (B. 1), Bertsekas and Rhodes (B.2), and Schweppe (S. 1), (S. 2). The index of performance in this situation typically involves minimizing the maximum damage that can be done if the uncertain quantities actually turn out to have "unfavorable" values — in essence, a sort of worst case design. A situation in which the sets containing the uncertain quantities have ill-defined boundaries has been considered by Zadeh (Z. 1).

In this thesis, the probabilistic point of view will be taken. In particular, a clear and precise formulation of the stochastic* optimal control problem and the stochastic optimal control/measurement problem will be given and interpreted. Although many authors have considered stochastic control before, there has been a great lack of precision in writings on the subject. Hopefully, one result of this thesis will be the elimination of certain areas of confusion that will be pointed out below.

---

*"Stochastic" is used in this thesis as a synonym for "probabilistic".

1.2 The Measurement Optimization Problem

In deterministic** control problems, the concept of state plays a central role. This is because the state summarizes the entire past history of the behavior of the system; given the value of the state at some time t, no explicit information regarding the inputs to the system or its resulting behavior before time t is required to calculate the future response. This notion is particularly useful in the case of finite dimensional systems: those which may be characterized by a state vector of finite dimension. It is intuitively a great simplification to be able to describe the effects of an infinite dimensional entity (the input time function over an interval) by giving a finite set of numbers.

There are two key ideas that follow from the idea of state: first, the index of performance used in selecting the control is usually a functional of the state trajectory and control trajectory over the interval of time of interest, and perhaps of the final value of the state if that interval is of finite duration. The second idea is that one form of "answer" to the control problem consists of specifying the value of the control at time t as a function (only) of the state at time t, that function perhaps changing with time. This "closed loop" solution approach is valid since the state at time t summarizes all available information about the system at that time.

** "Deterministic" is used here to mean "involving no uncertain quantities". Strictly speaking, there are no deterministic control problems, but such a formulation is a useful fiction in some cases. It is up to the designer to determine which type of formulation to use, based on his overall goals and constraints.

In a deterministic problem, the closed loop approach may be conceptually converted to an entirely equivalent open loop technique. One might imagine "running" the system using the feedback control and recording the values of the control as time evolves. If one then "runs" the system again (starting it in the same initial state as before) and supplies the control by playing back the record rather than by feeding back the state, the same state trajectory will result and the performance index will achieve the same value. This is evidently true since the control input is the only influence effecting the state trajectory. This equivalence, of course, fails when the system is influenced by uncertain inputs which act in addition to the specified control since the uncertain quantities may take on different (sample) values over different runs.

Heuristically, one can argue that in the stochastic case, exact knowledge of the state at time t coupled with the application of a closed-loop control law still represents a valid solution to the control problem since that known value of state summarizes the effects of all past system influences, uncertain or known. Unfortunately, the typical problem statement in the stochastic case includes the restriction that the state cannot be determined precisely. The designer is usually forced to use imperfect measurements of the state as his basis for control. It makes intuitive sense to employ these measurements to estimate the actual value of the state and then base the value of the control on that estimate. This interplay between state estimation and control is a crucial aspect of the theory of stochastic control, and it will be considered in much greater detail below. At this point, it suffices to note that imprecise measurements of the true state of the system must be used to generate the control signals.

In the usual formulation of the stochastic control problem, the measurement equation is taken as given. Suppose, however, as is frequently the case, that one has some influence over the quality of the measurements. One might be able to make a particularly good measurement at some time by making a large effort in some sense. One might be able to suspend measurements for a period to conserve resources for later measurement and control. One might be constrained to make a fixed number of discrete measurements or less, say due to the measurement having a disruptive effect on the system, but freedom over when to make these measurements might be given. One might have the choice of several measurement configurations (e.g., measurement of one state or another) with the restriction that only one configuration be used at a time due to resource limitations. Finally, it might be that the measurement subsystem is time-shared among many dynamic systems, with discrete measurements being made on a particular dynamic system at times $t = 0, n, 2n, 3n, \ldots$ where n is a fixed integer. Intuitively, the larger n, the more dynamic systems that can be observed with the one measurement subsystem, although "enough" measurements must be be carried out on each dynamic system to achieve an overall control objective.

For the reasons above, it makes sense to include a measurement control parameter in the equation representing the observation of the system state. Since taking the measurement typically involves expending resources, it also makes sense to include the measurement control history in the performance index functional. One of the primary concerns of this thesis will be the interrelation among the selection of state estimation, dynamic control, and measurement control techniques. It will be shown that for certain classes of systems, a particularly simple overall control structure results.

1.3 Mathematical System Description

In order to further descuss the issues which this thesis will consider, it is convenient to have an established notation. To this end, let R^i, for any integer i represent the i-dimensional Euclidean vector space and let $\mathcal{T} = \{0, 1, 2, \ldots, T\}$ represent a time index set. Suppose that the plant under consideration evolves according to the vector difference equation

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{u}_t, \underline{\xi}_t) \tag{1.3.1}$$

where for $t \in \mathcal{T}$, $\underline{x}_t \in R^n$ is the state vector,* $\underline{u}_t \in R^m$ is the dynamic control vector, and $\underline{\xi}_t \in R^q$ is a random vector on a given probability space $(\Omega, \mathcal{A}, p)$. Thus, by definition (Loeve, (L.1), p.150) $\underline{\xi}_t(\cdot): \Omega \to R^q$ is a measurable map with respect to the σ-algebra $\mathcal{A}$. It will generally be assumed that the sequence of random vectors $\{\underline{\xi}_0, \underline{\xi}_1, \ldots, \underline{\xi}_{T-1}\}$ is "white", i.e., that $\underline{\xi}_i$ is independent of $\underline{\xi}_j$ for all $i \neq j$. The function $\underline{f}_t(\cdot, \cdot, \cdot)$ is assumed to be a Borel function (see Loeve, (L.1), p.110) from $R^n \times R^m \times R^q$ into R^n.** It is also assumed that the initial state $\underline{x}_0$ is a random vector on $(\Omega, \mathcal{A}, p)$ which is independent of $\underline{\xi}_t$ for all $t \in \mathcal{T}$. From these assumptions, it follows recursively (see (L.1), p.110)

---

*The term "state" is used somewhat loosely here. Exactly what constitutes the state of a stochastic problem will be discussed in Chapter II.

**This is slightly more than is needed here. If the sequence $\{\underline{u}_t\}_{t=0}^{t-1}$ is fixed (open loop), all that is required is that $\underline{f}_t(\cdot, \underline{u}, \cdot)$ be Borel from $R^n \times R^q$ into R^n for all fixed $\underline{u}$. The more restrictive hypothesis will be needed below when closed loop controls are defined.

and the Borel Functions Theorem, p. 154) that for any fixed sequence $\{\underline{u}_0, \underline{u}_1, \ldots, \underline{u}_{T-1}\}$ of dynamic controls, the sequence of states $\{\underline{x}_0, \underline{x}_1, \ldots, \underline{x}_T\}$ is a random process on $(\Omega, \mathcal{S}, p)$, i.e., that the $x_t, t \in \mathcal{T}$, are jointly distributed random vectors.

Suppose now that measurements are made according to the equation

$$\underline{y}_t = \underline{g}_t(\underline{x}_t, \underline{v}_t, \underline{\theta}_t); \; t \in \mathcal{T} \tag{1.3.2}$$

where $\underline{y}_t \in R^r$ is the measurement vector, $v_t \in R^s$ is the measurement control, and $\{\underline{\theta}_0, \underline{\theta}_1, \ldots, \underline{\theta}_T\}$ is a sequence of independent random vectors on $(\Omega, \mathcal{S}, p)$ (discrete white noise) with values in R^p which is independent of $\underline{x}_0$ and $\{\underline{\xi}_t\}_{t=0}^{T-1}$. The function $\underline{g}_t(\cdot, \cdot, \cdot)$ is assumed to be a Borel function from $R^n \times R^s \times R^p$ into R^r. It follows, as above, that $\{\underline{y}_t\}_{t=0}^{T}$ is a random process on $(\Omega, \mathcal{S}, p)$ for any fixed sequence of measurement controls.

For ease of notation, sequences will be specified below without indicating their index sets where no confusion will result, e.g.,

$$\{\underline{x}_t\} = \{\underline{x}_t\}_{t=0}^{T} \text{ and } \{\underline{u}_t\} = \{u_t\}_{t=0}^{T-1}.$$

Suppose now that the following functional of $\{\underline{x}_t\}$, $\{\underline{u}_t\}$, and $\{\underline{v}_t\}$ is defined:

$$J = L_T(\underline{x}_T) + \ell_0(\underline{v}_0) + \sum_{t=0}^{T-1} L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1}) \tag{1.3.3}$$

where all the functions are Borel. Then J is also a random variable on $(\Omega, \mathcal{S}, p)$ for any fixed $\{\underline{u}_t\}$ and $\{\underline{v}_t\}$ and it makes sense to specify an optimization problem in terms of minimizing J in some stochastic sense. Exactly in what sense can be a complicated question which will be considered in Chapter II. At that point, the reasons for choosing the particular form of J displayed in equation (1.3.3) will also become clear.

Of course, additional restrictions must typically be placed on the various functions that have been defined in order to insure the a-priori existence and uniqueness of a solution to the optimization problem. These questions, however, are beyond the scope of this thesis. It will be assumed here that a solution exists, and methods of characterizing it will be discussed. At any rate, the L_t and ℓ_0 functions will generally be taken to be positive definite and perhaps convex.

Up to this point, only open loop control sequences have been considered, that is, fixed sequences $\{\underline{u}_t\}$ and $\{\underline{v}_t\}$ have been mentioned. In order to consider closed loop strategies, one must specify precisely what depends on what. To this end, let the following notation be established: For any variable $\underline{z}_t$, the capital letter with subscript will refer to the entire time history of that variable from the initial time t = 0 up to the specified subscript:

$$Z_t = \{\underline{z}_0, \underline{z}_1, \underline{z}_2, \ldots, \underline{z}_{t-1}, \underline{z}_t\} \tag{1.3.4}$$

In particular, $Y_t = \{\underline{y}_0, \underline{y}_1, \ldots, \underline{y}_t\}$ will denote the set of past measurements. Then the functional form of the allowed control strategies will be of the following type:

$$\underline{u}_t = \underline{\phi}_t (Y_t) \tag{1.3.5}$$

$$\underline{v}_{t+1} = \underline{\psi}_{t+1}(Y_t) \tag{1.3.6}$$

where $\underline{\phi}_t(\cdot)$ and $\underline{\psi}_{t+1}(\cdot)$ are Borel functions from $R^{t\cdot r}$ into R^m and R^s respectively. What this says is that the value of the present dynamic control $\underline{u}_t$ and the next measurement control $\underline{v}_{t+1}$ are determined by the measurements up to now.

Under all the assumptions that have been made, if (1.3.5) and (1.3.6) are substituted into (1.3.1) and (1.3.2) to give

$$\left.\begin{aligned} \underline{x}_{t+1} &= \underline{f}_t(\underline{x}_t, \underline{\phi}_t(Y_t), \underline{\xi}_t) \\ \underline{y}_{t+1} &= \underline{g}_{t+1}(\underline{x}_{t+1}, \underline{\psi}_{t+1}(Y_t), \underline{\theta}_{t+1}) \end{aligned}\right\} \qquad (1.3.7)$$

then it again follows recursively from the fact that $\underline{x}_0$ is a random variable on $(\Omega, \mathcal{J}, p)$ that the sequences $\{\underline{x}_t\}$ and $\{\underline{y}_t\}$ are random processes. Similarly, J is still a random variable which may be minimized in a sense to be specified below by choice of the functions $\underline{\phi}_t$ and $\underline{\psi}_t$ from among the class of Borel maps.

According to the above structural assumptions, one might visualize the following scenario for the evolution of the system:

1. Nature picks the initial state $\underline{x}_0$
2. You pick the measurement control $\underline{v}_0$
3. Nature picks the measurement noise $\underline{\theta}_0$
4. You measure $\underline{y}_0$

   Let $k = 0$

5. Based on the measurement set Y_k, you pick dynamic control $\underline{u}_k$ and measurement control $\underline{v}_{k+1}$
6. Nature picks driving noise $\underline{\xi}_k$
7. Based on $\underline{x}_k$, $\underline{\xi}_k$, and $\underline{u}_k$, the plant generates $\underline{x}_{k+1}$
8. Nature picks measurement noise $\underline{\theta}_{k+1}$
9. You measure $\underline{y}_{k+1}$ and increase the measurement set in size

   Increase k by 1 and go to Step 5

Steps 5-9 repeat until the final state $\underline{x}_T$ has been generated and the final measurement $\underline{y}_T$ has been made.

It will be convenient here to make one additional assumption in the interest of ease of presentation. It will generally be assumed that all random variables induce probability density functions in their range spaces. This assumption may be relaxed with no change in results, but making it allows two practical simplifications: first, expected values can be written as Riemann integrals in Euclidean spaces, and second, the changes in the probability structure as time evolves and measurements are taken can be described by (Riemann) integral equations (see, specifically, Theorem 2.3.1 in Chapter II).

Without the assumption of the existence of probability density functions, expected values of a function $f(\cdot)$ of a random variable $\underline{x}$ must be written $\int_{R^n} f(\underline{x})\, dP(\underline{x})$ where the integral is with respect to the measure P induced in R^n by the random vector $\underline{x}$. This offers no more conceptual insight than the equivalent form utilizing the density if it exists, namely $\int_{R^n} f(\underline{x})\, p(\underline{x})\, d(\underline{x})$, where $p(\underline{x})$ is the density. Without the assumption of the existence of probability density functions, the evolution of the probability structure as measurements occur must be accounted for by considering a family of σ-algebras $\mathcal{S}_{Y_t}$ with respect to which the measurements are measurable. This concept does more harm by introducing complexity than it does good by introducing generality. For these reasons, the formulation using densities will generally be used. The measure approach, of course, will be used when it illustrates a point more clearly.

1.4 Previous Results

Stochastic control has been studied for some time now. Certain results have become "well known" and the topic is considered in several textbooks, for example, those of Bryson and Ho (B.4), Aoki (A.1), Kushner (K.1), (K.2), Wong (W.3), and Astrom (A.2). No attempt will be made here to give a complete bibliography on stochastic control since such texts, with their extensive reference lists, or survey papers such as those by Witsenhausen (W.1) and Athans (A.3) are available. Attention will be restricted here to works dealing with the measurement optimization problem in particular.

Several authors have obtained results for specific examples of the measurement optimization problem introduced in the previous section. In most cases, the index of performance is taken to be the expected value of J given in (1.3.3) (further comment on what type of expectation will be made in Chapter II). The particular case investigated is the situation in which the dynamic and measurement systems are linear in $\underline{x}$ and $\underline{u}$ (although not in $\underline{v}$), the cost functional J is quadratic in $\underline{x}$ and $\underline{u}$, and the noises are white Gaussian, There is a certain amount of confusion in the available derivations, however the right results are obtained. As Witsenhausen (W.1) states in talking of the linear-quadratic-Gaussian problem:

> Much confusion has been abetted by the incredible robustness of this case to conceptual misunderstandings: every reasonable assertion about that case is true and, within wide limits, no amount of confusion can give an incorrect result. In fact, the most confused derivations of the correct results are also among the shortest.

Before discussing the particular contributions of individual authors, the underlying method used by all of them will be described.

Consider as an example a linear stochastic system which has the property that one can influence the signal-to-noise ratio of the measurements:

$$\underline{x}_{t+1} = \underline{A}_t \underline{x}_t + \underline{B}_t \underline{u}_t + \underline{\xi}_t \tag{1.4.1}$$

$$\underline{y}_t = v_t \underline{C}_t \underline{x}_t + \underline{\theta}_t \tag{1.4.2}$$

Here $\{v_t\}$ is a scalar sequence which allows the "signal" in the measurement equation to be boosted with respect to the noise $\underline{\theta}_t$. For example, in a radar system, v_t might represent the fact that by sending out more power in each radar pulse, a stronger echo will result. The initial state vector $\underline{x}_0$ is a Gaussian random vector with zero mean and covariance $\underline{\Sigma}_0$:

$$E\{\underline{x}_0\} = \underline{0}; \quad E\{\underline{x}_0 \underline{x}_0'\} = \underline{\Sigma}_0 \tag{1.4.3}$$

The noise sequences are zero mean, white, Gaussian, mutually independent, and independent of $\underline{x}_0$:

$$E\{\underline{\xi}_t\} = \underline{0}; \quad E\{\underline{\xi}_t \underline{\xi}_s'\} = \underline{\Xi}_t \delta_{ts} \tag{1.4.4}$$

$$E\{\underline{\theta}_t\} = \underline{0}; \quad E\{\underline{\theta}_t \underline{\theta}_s'\} = \underline{\Theta}_t \delta_{ts} \tag{1.4.5}$$

$$E\{\underline{\xi}_t \underline{\theta}_s'\} = \underline{0} \tag{1.4.6}$$

$$E\{\underline{\xi}_t \underline{x}_0'\} = \underline{0}; \quad E\{\underline{\theta}_t \underline{x}_0'\} = \underline{0} \tag{1.4.7}$$

where δ_{ts} is the Kronecker delta, $\underline{\Xi}_t$ is a positive semidefinite matrix for all t, and $\underline{\Theta}_t$ is positive definite.

If a cost functional is given that is a slight extension of the usual quadratic one:

$$J(\underline{u}, v) = \underline{x}_T' \underline{Q}_T \underline{x}_T + \ell_T(v_T) + \sum_{t=0}^{T-1} \underline{x}_t' \underline{Q}_t \underline{x}_t + \underline{u}_t' \underline{R}_t \underline{u}_t + \ell_t(v_t) \tag{1.4.8}$$

where $\underline{Q}_t$ are all positive semidefinite and $\underline{R}_t$ are positive definite, then the optimal dynamic control sequence $\underline{u}^* \triangleq \{\underline{u}_t^*\}$ and the optimal measurement control sequence $v^* \triangleq \{v_t^*\}$ must satisfy*

$$E\{J(\underline{u}^*, v^*)\} \leq E\{J(\underline{u}, v)\} \qquad (1.4.9)$$

for all $\underline{u}$ and v which are admissible.

Suppose now that one starts by assuming a "fixed but arbitrary" time sequence for $\{v_t\}$. The solution for the best $\{\underline{u}_t\}$ follows at once from the well known results of Kalman and the "Separation Theorem", discussed, for example, in Bryson and Ho (B.4):

$$\underline{u}_t^* = -\underline{K}_t \hat{\underline{x}}_{t|t} \qquad (1.4.10)$$

$$\hat{\underline{x}}_{t|t} = \hat{\underline{x}}_{t|t-1} + \underline{S}_t(\underline{y}_t - v_t \underline{C}_t \hat{\underline{x}}_{t|t-1}) \qquad (1.4.11)$$

$$\hat{\underline{x}}_{t|t-1} = \underline{A}_{t-1} \hat{\underline{x}}_{t-1|t-1} + \underline{B}_{t-1} \underline{u}_{t-1}; \; \hat{\underline{x}}_{0|-1} = \underline{0} \qquad (1.4.12)$$

$$\underline{K}_t = [\underline{B}_t' \underline{M}_{t+1} \underline{B}_t + \underline{R}_t]^{-1} \underline{B}_t' \underline{M}_{t+1} \underline{A}_t \qquad (1.4.13)$$

$$\underline{S}_t = \underline{\Sigma}_{t|t-1} v_t \underline{C}_t' [v_t^2 \underline{C}_t \underline{\Sigma}_{t|t-1} \underline{C}_t' + \underline{\Theta}_t]^{-1} \qquad (1.4.14)$$

$$\underline{\Sigma}_{t|t-1} = \underline{A}_{t-1} \underline{\Sigma}_{t-1|t-1} \underline{A}_{t-1}' + \underline{\Xi}_{t-1}; \; \underline{\Sigma}_{0|-1} = \underline{\Sigma}_0 \qquad (1.4.15)$$

$$\underline{\Sigma}_{t|t} = \underline{\Sigma}_{t|t-1} - \underline{S}_t [v_t^2 \underline{C}_t \underline{\Sigma}_{t|t-1} \underline{C}_t' + \underline{\Theta}_t] \underline{S}_t' \qquad (1.4.16)$$

$$\underline{M}_t = \underline{A}_t' \underline{M}_{t+1} \underline{A}_t - \underline{K}_t' [\underline{R}_t + \underline{B}_t' \underline{M}_{t+1} \underline{B}_t] K_t + \underline{Q}_t; \; \underline{M}_T = \underline{Q}_T \qquad (1.4.17)$$

where some of these quantities may be interpreted as follows:

$\hat{\underline{x}}_{t|t}$ = estimate of $\underline{x}_t$ given Y_t

$\hat{\underline{x}}_{t|t-1}$ = estimate of $\underline{x}_t$ given Y_{t-1}

$\underline{\Sigma}_{t|t}$ = covariance of $\underline{x}_t$ given Y_t

---

*The precise meaning of this expression will be fully discussed in Chapter II.

$\Sigma_{t|t-1}$ = covariance of $\underline{x}_t$ given Y_{t-1}

$\underline{S}_t$ = Kalman filter gain matrix

$\underline{K}_t$ = linear regulator feedback matrix

The average value of J using the optimal control sequence $\{\underline{u}_t^*\}$ is then given by

$$E\{J(\underline{u}^*, v)\} = \mathrm{tr}\left[\underline{M}_0\underline{\Sigma}_0 + \sum_{t=0}^{T-1} \underline{M}_{t+1}(\underline{\Xi}_t + \underline{B}_t\underline{K}_t\underline{\Sigma}_{t|t}\underline{A}_t')\right]$$

$$+ \sum_{t=0}^{T} \ell_t(v_t) \qquad (1.4.18)$$

where tr represents the matrix trace operation. Note that the expected cost has some terms which depend on $\{v_t\}$ and some which do not. In particular, reference to equations (1.4.10) through (1.4.17) shows that the evolution of $\underline{\Sigma}_{t|t}$ is influenced by $\{v_t\}$ but that of $\underline{M}_t$ and $\underline{K}_t$ is not. Thus one can now define an auxiliary (deterministic) control problem in which the covariance matrix $\underline{\Sigma}_{t|t}$ plays the role of state and v_t the role of control. One chooses the measurement control sequence to minimize those terms in (1.14.18) which it influences.

The approach outlined above is partially justified by the facts that

1. The $\underline{K}_t$ sequence is independent of $\{v_t\}$
2. The control law, i.e., the structure, thus turns out to be independent of $\{v_t\}$ and $\underline{\Sigma}_{t|t}$ although v was assumed known.

On the other hand, it is not clear whether the same solution will be obtained if one solves simultaneously for $\{\underline{u}_t\}$ and $\{v_t\}$. Also, it is not clear what effect the means of penalizing v has on the problem, specifically, what happens for different $\ell_t(v_t)$ functions or if $\underline{x}_t$, $\underline{u}_t$ and v_t couple in the cost.

The measurement optimization problem as formulated here was first motivated by the work of Athans and Schweppe (A.4) on optimal waveform design. These authors consider optimizing the modulation of a radar signal to allow the most accurate estimation possible of a body being tracked while not violating power and energy constraints on the radar. The modulation signal appears in the overall system of equations for this problem as a measurement control. The radar problem is somewhat different than the one formulated here, however, since no control of the body being observed is done, only state estimation.

Several authors have considered measurement/control optimization problems. All of these use the "a-priori v technique" outlined above as applied to linear systems. Representative of this work are the papers of Kushner (K.3), Meier, Peschon and Dressler (M.2), Aoki and Li (A.5), Sano and Terao (S.5), Athans (A.6) and Cooper and Nahi (C.1). Kushner (K.3) considers the optimal timing of a limited number of measurements of the "measure-no measure" type which can be represented as equation (1.4.2) if v_t is constrained to be either zero or one. His system is linear and discrete in time. Meier, Peschon and Dressler (M.2) consider a problem with a more general cost on measurement and apply dynamic programming. As will be discussed later, there is some question as to precisely how to apply dynamic programming to a measurement optimization problem, although these authors prove a theorem concerning its use on the linear-quadratic-Gaussian problem. The work of Aoki and Li (A.5) is similar to that of Kushner. Sano and Terao (S.5) consider a continuous time system measured discretely in time. Again the "a-priori v technique" is applied. Cooper and Nahi (C.1) apply dynamic programming to solve a problem similar to that of Meier, Peschon and Dressler. Athans (A.6) considers a continuous time system in which the measurement control in effect chooses one of several possible sensor outputs.

The difficulty with the approach taken by all the above authors is that it has not been rigorously justified. As mentioned above, there is no guarantee that the "a-priori v technique" will give the same answer as a method in which the designer solves simultaneously for $\{\underline{u}_t\}$ and $\{v_t\}$. Meier, Peschon and Dressler take a step toward providing such a guarantee in certain cases but their proof, which will be discussed in Chapter IV, is incomplete. Part of the work described here will be a careful proof of the optimality of the "a-priori v technique" for linear-quadratic-Gaussian systems.

1.5 Structure of this Thesis

The work reported here will be organized in the following manner: In Chapter II, the meaning of stochastic optimality will be discussed and the application of dynamic programming to the solution of stochastic optimal control problems will be studied. The question of properly formulating an index of performance that makes sense will be considered and several points of view will be discussed. Some mathematical problems relating to the definition of stochastic control problems will be dealt with, and the whole framework of solution techniques for stochastic control problems will be clarified.

Chapter III will present a new technique for solving stochastic control problems formulated around linear systems. A method will be presented in which the designer may assume the noise sequences are fixed, allowing the use of deterministic techniques such as the minimum principle, with expected values being taken at the end. This new approach will be carefully justified and the limits of its applicability studied.

In Chapter IV, linear systems with quadratic cost criteria and Gaussian disturbances will be considered using the techniques of Chapters II and III. A careful proof of a two-way separation theorem will be given, which will justify the "*a-priori* v technique" described above. In Chapter V, a similar analysis will be done for linear systems with Gaussian noise and general cost functionals. It will be shown by example that the two-way separation theorem which is inherent in the use of the "*a-priori* v technique" does not hold in general, and an analysis will be given to characterize those cases in which it does hold.

Chapter VI will present a problem of practical interest which involves coupled measurement and control. The aim is to further the understanding of stochastic problems of this type.

The final chapter will summarize the results given earlier and present a unified view of the stochastic control concept. Areas of application will be considered and topics for future research will also be discussed.

1.6 Contributions of this Thesis

There are essentially three main areas in which the work reported here contributes to the field of stochastic control. First, there is value in a clear and careful exposition of the issues which arise in defining a stochastic control problem. These issues have often been skipped over in other works, and it is felt that a certain amount of confusion has resulted. Secondly, it is felt that the ideas presented in Chapter III regarding the application of deterministic techniques to stochastic problems are a significant theoretical contribution to the field. The application of these ideas will allow solution (or near solution) of some problems which are simply too difficult for practical answers to be found using conventional techniques. Finally, the particular problems analyzed in Chapters IV-VI provide useful insight into the structure of control systems that can be used for many problems of practical interest.

CHAPTER II

STOCHASTIC OPTIMALITY AND DYNAMIC PROGRAMMING

2.1 The Meaning of Stochastic Optimality

In a deterministic optimal control problem, there is no confusion about the meaning of a solution. A system equation and a cost functional of the following types are given:

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{u}_t); \ \underline{x}_0 \text{ given} \tag{2.1.1}$$

$$J = L_T(\underline{x}_T) + \sum_{t=0}^{T-1} L_t(\underline{x}_t, \underline{u}_t) \tag{2.1.2}$$

and one seeks the sequence $\{\underline{u}_t\}$ (subject perhaps to some constraints) which causes J to achieve its minimum value. This process is conceptually well-defined because J is a real-valued function of the control sequence; it can be viewed as a map from the space of sequences to the space of real numbers, and under suitable conditions, the sequence giving the smallest real number for an answer can be determined. That minimizing sequence may be determined in an open loop manner for the particular initial state given in (2.1.1), or the problem may be imbedded in a class of similar problems with fixed but arbitrary initial states and the solution obtained in a closed loop manner via dynamic programming. Either approach is well-defined a-priori because the map $\{\underline{u}_t\} \rightarrow J[\{\underline{u}_t\}]$ is well defined a-priori.

In the stochastic case, the situation is not so simple. Recall the system and cost equations as defined in Section 1.3:

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{u}_t, \underline{\xi}_t) \tag{2.1.3}$$

$$\underline{y}_t = \underline{g}_t(\underline{x}_t, \underline{v}_t, \underline{\theta}_t) \tag{2.1.4}$$

$$J = L_T(\underline{x}_T) + \ell_0(\underline{v}_0) + \sum_{t=0}^{T-1} L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1}) \tag{2.1.5}$$

If the control sequences $\{\underline{u}_t\}$, $\{\underline{v}_t\}$ are fixed or if they are given functions of the measurements as described in Section 1.3, then $\{\underline{x}_t\}$ and $\{\underline{y}_t\}$ are random processes and J is a random variable. If, on the other hand, the control sequences have not yet been selected or are to be selected by applying an as yet underdetermined function to the measurements, the state and measurement sequences are not random processes and J is not yet a random variable because these quantities have not been described as functions of $\omega \epsilon \Omega$, the underlying probability set. See, for example, Witsenhausen (W. 1). Thus the status of the system, indeed, of the mathematical formulation, is uncertain while the designer is searching for the right control sequences. This situation will be discussed further below.

Once the problem described above is clarified, there remains another: for any control sequences, the index of performance J is a random variable, i.e., a real-valued function of the probability variable ω. The map $[\{\underline{u}_t\}, \{\underline{v}_t\}] \rightarrow J[\{\underline{u}_t\}, \{\underline{v}_t\}]$ must now be viewed as a map from a space of sequences to an infinite dimensional space. Since there is no natural ordering on the space of random variables, one cannot pick the influence sequences that give the "smallest" J, and one must be quite careful to specify the criterion that will be used to make the choice.

The most obvious approach is to define a second map that goes from the space of random variables to the real numbers so that the ordering of the reals induces a partial ordering on the random variables. Then the concatenation of the two previous maps, denoted $\mathcal{J}$, $[\{\underline{u}_t\}, \{\underline{v}_t\}] \rightarrow J[\{\underline{u}_t\}, \{\underline{v}_t\}] \rightarrow \mathcal{J}[\{\underline{u}_t\}, \{\underline{v}_t\}]$, is a map from sequences to reals again and the sequences giving the smallest $\mathcal{J}$ can be picked. Of course, the second map must be chosen in an intuitively satisfying way, a way consistent with the original control design. The approach that is usually taken is to use the expected value as the second map.*

If one determines to simply take $\mathcal{J} = E\{J\}$, one is defeating the purpose of taking measurements, however. That expression must be interpreted as an expectation over all underlying uncertainties, including the uncertain measurement sequence to be observed. What one would like to do is choose the control variables at each instant of time in such a way as to utilize all the information available at that time. This seems to indicate that one should take a conditional expectation of some type, but due care must be exercised. Whenever one is attempting to select a particular parameter in order to minimize another quantity, one must be certain that the quantity to be minimized is a number and not a random variable. This can be confusing if a conditional expectation is being minimized because some conditional expectations are random variables and some are real numbers. In particular, if x and y are random variables, the expression $E\{x | y = Y\}$ represents a number if Y is a fixed number, the

---

*Other maps from the space of random variables to the reals might be considered. For example, Sain (S.3) and Sain and Liberty (S.4) consider minimizing the variance of J subject to the constraint $E\{J\}$ = fixed. Other generalizations are a possible area for further investigation.

expression $z(Y) = E\{x|y = Y\}$ is a real-valued function of Y if Y ranges over the real numbers, but the expression $E\{x|y\}$ represents a random variable if the outcome of the random variable y is not known to be a specific value. It depends on whether the conditioning is done on a random variable or on a known fact.

The end result of these considerations is that the index of performance must change as time evolves to account for the information that is accumulated as time evolves. The precise meaning of this concept will be the subject of the next section.

There is one more point to consider: the actions of the control generator at the present time must intuitively depend not only on the measurements that have been observed up to now, but also on the actions that will be taken in the future. This statement must be interpreted very carefully; it is clear that the present control variables cannot depend on the values of future controls, since these values are not yet known. What is meant is that the values of the present control variables must depend on the structure that will be used to generate future controls.

The problems and solutions introduced in a heuristic way in this section will now be considered more carefully mathematically in the next several sections. The expected value map will be the device used to convert the given stochastic problem to one which has a well-defined order on the cost functional range space. Exactly how to apply this map will be a chief concern of the discussion.

In order to make the discussion as clear as possible, the points raised here will be considered more or less in reverse order. The mathematical complication of the "random-variable-to-be" status of the system will be clarified after the approach used to change the cost functional as time progresses is described. In this way, the mathematical questions will be better motivated. Thus the techniques of the next section will not be fully justified until later in the chapter.

2.2 Dynamic Programming and Stochastic Control

The virtue of the dynamic programming technique of solving an optimization problem is that it "automatically" provides an answer in closed loop form. At each step, the variables being determined are specified as functions of the information available at that step. There is, however, a problem when this technique is applied to a stochastic control problem.

Recall how dynamic programming is applied to the deterministic problem summarized by equations (2.1.1) and (2.1.2). At a given instant of time t, the state is known to be $\underline{x}_t$ and the control $\underline{u}_t$ is sought. It is assumed that regardless of the state $\underline{x}_{t+1}$ that results from the choice of $\underline{u}_t$, the optimal remaining control sequence $\underline{u}^*_{t+1}, \ldots, \underline{u}^*_{T-1}$ will be applied and the resulting minimal cost-to-go from $\underline{x}_{t+1}$ will be realized. Let this minimal cost-to-go from $\underline{x}_{t+1}$ be denoted by $I_{t+1}(\underline{x}_{t+1})$:

$$I(\underline{x}_{t+1}) = \min_{\underline{u}_{t+1}, \ldots, \underline{u}_{T-1}} \left\{ \sum_{k=t+1}^{T-1} L_k(\underline{x}_k, \underline{u}_k) + L_T(\underline{x}_T) \right\} \quad (2.2.1)$$

where $\underline{x}_{t+1}$ is fixed but arbitrary. Then according to the principle of optimality (see Bellman (B.3)), the value of $\underline{u}_t$ is chosen to minimize the cost-to-go from $\underline{x}_t$, given by J_t:

$$J_t = L_t(\underline{x}_t, \underline{u}_t) + I(\underline{x}_{t+1}) \tag{2.2.2}$$

$$= L_t(\underline{x}_t, \underline{u}_t) + I(\underline{f}_t(\underline{x}_t, \underline{u}_t)) \tag{2.2.3}$$

Evidently, the $\underline{u}_t$ that minimizes (2.2.3) is a function of $\underline{x}_t$, and if this value is re-substituted into (2.2.3), the resulting minimal value of J_t is the new $I_t(\underline{x}_t)$. This process can be started at $t = T$ and stepped backwards to $t = 0$, finding the optimal feedback control law $\underline{u}_t^*(\cdot)$ along the way. (The arguments of $\underline{u}_t^*(\cdot)$ are $\underline{x}_t$, of course).

Why does this work? How can this idea be applied to a stochastic problem? There are two points of difficulty that can be cleared up easily, but which are often not mentioned explicitly in treatments of stochastic control. First: what constitutes the state in a stochastic problem and second: what can be done about the coupling caused by present measurements influencing estimates of past quantities?

If the stochastic problem is formulated by using the expected value map to provide the ordering on the space of random variables to which the cost functional belongs, it is relatively clear that the role of state is played by the conditional probability density function,* given the measurements, of that quantity denoted by $\underline{x}_t$ in the equations of motion. The reason is that to calculate the expected value of a function of $\underline{x}_t$, it is only necessary to integrate that function against the probability density function (p.d.f. hereafter) of $\underline{x}_t$. This was recognized by Bellman (B.3)

---

*More generally, the conditional probability space $(\Omega_{Y_t}, \mathcal{S}_{Y_t}, p_{Y_t})$ induced in the basic space $(\Omega, \mathcal{S}, p)$ by the set of measurements Y_t.

when dynamic programming was first developed. The idea was more fully investigated by Striebel (S.6). Thus, if one takes a finite dimensional deterministic system and adds noise, one produces an infinite dimensional problem. The vector that was the plant state of the deterministic problem is no longer the state of the system; its probability density function is. In a few cases, it turns out that the stochastic problem is finite dimensional also: precisely those cases in which the conditional p.d.f. may be completely characterized by a finite set of numbers, e.g., the mean vector and covariance matrix in the Gaussian case. In general, however, the stochastic problem is infinitely more difficult, so to speak.

By the abuse of nomenclature, the quantity $\underline{x}_t$ will continue to be referred to as the "plant state vector" in stochastic as well as deterministic contexts. This terminology is useful and suggestive, but one must bear in mind that it is somewhat imprecise: it sometimes leads to statements such as "the state is the conditional p.d.f. of the plant state."

Although it has not explicitly been mentioned in the two paragraphs above, one should note that the whiteness of the driving and measurement noises has been used in asserting that the state of the problem is the conditional p.d.f. of the plant state. If the noises are white, the statistics of the future noises do not change as measurements are made. If the noises are not white, this is not true and one must include into the state of the problem the conditional densities of the future noises given the current measurements.

In the deterministic case, it makes sense to imbed the given problem in a class of similar problems with varying intervals of evolution and various initial states. This is because if the system winds up in some state $\underline{x}_{t_1}$ at $t = t_1$, the accrued cost from the starting time to t_1 is no longer affected by the state and control trajectories after t_1. Present actions influence only present (directly) and future (indirectly) values of state. Thus although the sub-problem solved at each stage of the dynamic programming technique ignores the previously accrued cost and deals only with the state achieved at present, there is perfect justification for this: the accrued cost is not influenced by the remaining optimization.

In the stochastic case, although the actual value of the accumulated cost at time t_1 is not influenced by present or future actions, the expected value of these costs is. Each measurement causes updates in the estimates of all quantities, past (smoothing), present (filtering), or future (predicting). The decoupling effect of time is lost in a stochastic problem.

The only way to proceed from this point is to force a structure on the problem which allows one to neglect the effect of current measurements on the expected value of past costs while still continuing to consider the effect of current actions on future costs. What must be done is to adopt the point of view that the prime goal of the control system is to minimize the actual values of costs as well as possible, rather than estimating those costs. Since the actual value of the previously accumulated cost is established at present, it might as well be neglected in the remaining optimization. True, its value is unknown, but who cares? The unknown value if fixed and estimating it is of very little importance with regard to optimizing the rest of the trajectory.

Notice that in an "ordinary" stochastic control problem, that is, one without measurement optimization considerations, the fact that present measurements effect estimates of past costs is not particularly important. The choice of the present dynamic control will not influence past costs, even though the estimated value of those costs may change. When there is a measurement control capability, however, it is necessary to neglect the effect that the choice of the present measurement control has on past costs in order to have a problem structure that may be dealt with using dynamic programming.

2.3 The Proper Use of Expectations

Since the expected value of a function of a random variable may be obtained by integrating that function against the p.d.f. of the random variable, the propagation of p.d.f.'s is of interest in a stochastic control problem. The conditional p.d.f. of the state of a discrete dynamic system, in particular, can be propagated by a Chapman-Kolmogorov type of equation whether or not a measurement is made. Specifically, from the results of Jazwinski (J.1), p. 174, the following theorem is quoted:

Theorem 2.3.1. Consider a dynamic system as described in Section 1.3. Let $t \in \mathfrak{J}$ be fixed and let $Y_t = \{\underline{y}_{t_0}, \underline{y}_{t_1}, \ldots, \underline{y}_{t_k}\}$ be a set of observations* taken at times t_j such that

$$0 \leq t_0 < t_1 < \ldots < t_k \leq t.$$

*Note that it is not required that measurements at each and every time t to be included in the set Y_t.

Let the conditional p.d.f. $p(\underline{x}_t | Y_t, U_{t-1}, V_t)$ be known. Then assuming the existence of the indicated densities, if there is no measurement at time t+1, the density propagates according to

$$p(\underline{x}_{t+1} | Y_{t+1}, U_t, V_{t+1}) = \int_{R^n} p(\underline{x}_{t+1} | \underline{x}_t, \underline{u}_t)\, p(\underline{x}_t | Y_t, U_{t-1}, V_t)\, d\underline{x}_t \qquad (2.3.1)$$

where $Y_{t+1} = Y_t$, $U_t = U_{t-1} \cup \{\underline{u}_t\}$, and $V_{t+1} = V_t \cup \{\underline{v}_{t+1}\}$. The value $\underline{v}_{t+1}$ is assumed to specify "no measurement". If, however, a measurement is made, the density propagates according to

$$p(\underline{x}_{t+1} | Y_{t+1}, U_t, V_{t+1}) =$$

$$\frac{p(\underline{y}_{t+1} | \underline{x}_{t+1}, \underline{v}_{t+1}) \int_{R^n} p(\underline{x}_{t+1} | \underline{x}_t, \underline{u}_t)\, p(\underline{x}_t | Y_t, U_{t+1}, V_t)\, d\underline{x}_t}{\int_{R^n} \int_{R^n} p(\underline{y}_{t+1} | \underline{x}_{t+1}, \underline{v}_{t+1})\, p(\underline{x}_{t+1} | \underline{x}_t, \underline{u}_t)\, p(\underline{x}_t | Y_t, U_{t+1}, V_t) \}\, d\underline{x}_t\, dx_{t+1}} \qquad (2.3.2)$$

where now $Y_{t+1} = Y_t \cup \{\underline{y}_{t+1}\}$

Remark 1: Equation (2.3.1) is actually the special case of equation (2.3.2) that results if no measurement is made. The p.d.f. thus does not propagate in fundamentally different ways depending on whether or not a measurement is made, in one sense. On the other hand, if one writes equations (2.3.1) and (2.3.2) as operator equations of the following type:

$$p(\underline{x}_{t+1} | Y_{t+1}, U_t, V_{t+1}) = \mathfrak{F}_t [p(\underline{x}_t | Y_t, U_{t-1}, V_t), \underline{u}_t] \qquad (2.3.3)$$

and

$$p(\underline{x}_{t+1} | Y_{t+1}, U_t, V_{t+1}) = \mathfrak{D}_t [p(\underline{x}_t | Y_t, U_{t-1}, V_t), \underline{u}_t, \underline{v}_{t+1}, \underline{y}_t] \qquad (2.3.4)$$

(here (2.3.3) represents (2.3.1) and (2.3.4) represents (2.3.2)) one sees that in the case that a measurement is taken, the a-posteriori density $p(\underline{x}_{t+1}| Y_{t+1}, U_t, V_{t+1})$ is a random variable at time t since the measurement $\underline{y}_t$ has not yet been made, and (2.3.4) depends on $\underline{y}_t$. In contrast to this, reference to (2.3.3) shows that if no measurement is going to be made, the a-posteriori density $p(\underline{x}_{t+1}|Y_{t+1}, U_t, V_{t+1})$ is not random at time t.

Remark 2: Note that the propagation of the conditional p.d.f. only requires knowledge of the values of the controls $\underline{u}_t$ and $\underline{v}_{t+1}$, and not knowledge of the entire control laws $\underline{u}_t(\cdot)$ and $\underline{v}_{t+1}(\cdot)$. This is discussed extensively by Striebel (S.6). In fact, the treatment by Striebel (S.6) and Jazwinski (J.1) do not consider a situation in which measurement control capability is part of the mathamatical structure; however, their proofs go through essentially unchanged in the measurement control case. For this reason, no separate proof is given here.

Remark 3: It is crucial to include the conditioning on U_{t+1} and V_t in the conditional p.d.f. $p(\underline{x}_t|Y_t, U_{t-1}, V_t)$. In fact the conditional p.d.f. of $\underline{x}_t$ given only Y_t (i.e., without knowledge of U_{t-1} and V_t) is not defined. This fact is often overlooked, especially in the case of linear dynamics with Gaussian noise and with Gaussian initial state because in that case, one can characterize the conditional p.d.f. of $\underline{x}_t$ given Y_t by its mean, say $\hat{\underline{x}}_{t|t}$, and its covariance matrix, say $\underline{\Sigma}_{t|t}$. The conditional p.d.f. can be written in terms of $\hat{\underline{x}}_{t|t}$ and $\underline{\Sigma}_{t|t}$ without explicit display of U_{t-1}

and V_t, but one cannot actually calculate $\hat{\underline{x}}_{t|t}$ and $\underline{\Sigma}_{t|t}$ without knowing U_{t-1} and V_t.

Remark 4: To see precisely how $\underline{u}_t$ and $\underline{v}_{t+1}$ enter into, say, (2.3.2), note that to calculate the transition density $p(\underline{x}_{t+1}|\underline{x}_t, \underline{u}_t)$, one uses the system equation (2.1.3) and the density of $\underline{\xi}_t$. This involves $\underline{u}_t$. To calculate the conditional measurement density $p(\underline{y}_{t+1}|\underline{x}_{t+1}, \underline{v}_{t+1})$, one uses the measurement equation (2.1.4) and the density of $\underline{\theta}_{t+1}$. This involves $\underline{v}_{t+1}$. See Papoulis (P.1), p.118 ff. for a discussion of the required calculations.

2.3.1 Minimization of the Unconditional Expectation

Suppose one determines simply to minimize the unconditional expected value of J in (2.1.5):

$$\mathcal{J}_1 = E\{J(U_{T-1}, V_T)\} \tag{2.3.5}$$

where $U_{T-1} \triangleq \{\underline{u}_t\}_{t=0}^{T-1}$ and $V_T \triangleq \{\underline{v}_t\}_{t=0}^{T}$.

This is the requirement usually given by other authors. The logical way to interpret this minimization is to assume that no measurements will be taken. This is because if measurements are made, it would only make sense to utilize them in some way, and the expectation to be minimized would then be conditioned on the measurements. This situation will be considered below. For the moment, it will simply be noted that if one means to take conditional expectation, one should say so.

Under the assumption of no measurements, the p.d.f. of $\underline{x}_t$ propagates according to (2.3.1) where the information set Y_t is always the null set and the initial condition is the given a-priori density of $\underline{x}_0$. The expected value of J is given by

$$\mathcal{J}_1(U_{T-1}, V_T) = \ell_0(\underline{v}_0) + \int_{R^n} L_T(\underline{x}_T)\, p(\underline{x}_T | U_{T-1})\, d\underline{x}_T$$

$$+ \sum_{t=0}^{T-1} \int_{R^n} L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1})\, p(\underline{x}_t | U_{t-1})\, d\underline{x}_t \tag{2.3.6}$$

Note that the logical way to view the sequences $\{\underline{u}_t\}$ and $\{\underline{v}_t\}$ is as deterministic sequences. Since the cost is not conditioned on anything (except the a-priori known statistics so to speak) it makes no sense to have the influence sequences depend on anything but the a-priori knowledge, which means they are deterministic with respect to the evolution of the system. One may view equations (2.3.1) and (2.3.6) as a deterministic infinite dimensional optimal control problem in which the density plays the role of state. The measurement control $\underline{v}_t$ plays a degenerate role since it does not enter into the evolution of the density, but only into the cost.

The control problem defined by (2.3.1) and (2.3.6) can be solved by applying the techniques of deterministic infinite dimensional systems, albeit a difficult task but conceptually feasible. The result is two sequences of vectors $\{\underline{u}_t\}_{t=0}^{T-1}$ and $\{\underline{v}_t\}_{T=0}^{T}$ computed off-line so to speak and applied to the system as it evolves.

Obviously the technique described above is not "the best you can do". Since measurements are possible, it might be better to use them to reduce uncertainty about the system state $\underline{x}_t$ in the hope of making J smaller on the average than the optimal value of $\mathcal{J}_1$. The point to be made, however, is that using such measurements involving redefining the stochastic cost. Many authors assert that they are minimizing $\mathcal{J}_1$ when in fact they are minimizing something else, such as will be defined below.

In order to utilize measurements in real time, it is necessary to force a certain structure on the system. There are several possible structures, as will be shown below, and each leads to solutions that must be interpreted in specific ways. On the other hand, the common goal of all the structures given here is to reduce the stochastic problem to an equivalent deterministic one, just as above, where an infinite dimensional deterministic problem was the end result.

2.3.2 Conditional Open Loop Minimization

Suppose one views the control sequences as deterministic in the sense that they are to depend only on the a-priori statistics, but one chooses these sequences to minimize a conditional expectation. One can then try to formulate an optimization problem bases on the following cost:

$$\mathcal{J}_2 = \ell_0(\underline{v}_0) + E\{L_T(\underline{x}_T) | Y_T, U_{T-1}, V_T\} + \sum_{t=0}^{T-1} E\{L_t(\underline{x}_t, u_t, v_{t+1}) | Y_{t+1}, U_{t+1}, V_t\} \tag{2.3.7}$$

Since measurements are taken as the system evolves, the p.d.f. obeys (2.3.2). If one is to determine $\{\underline{u}_t\}$ and $\{\underline{v}_t\}$ in an open loop a-priori manner, however, this does not lead to a well-defined optimization problem since a-priori, $\underline{y}_t$ is a random process and it enters into the evolution of the p.d.f. This means that the cost associated with this interpretation, given in (2.3.7), is still a random variable, the difficulty mentioned in Section 2.1. Of course, a closed loop approach is desired anyway, so the case just described is included only to point out the problem.

2.3.3 Closed Loop Approaches

What one really would like to do is base current control values on the available observed data. Thus the desired form is

$$\underline{u}_t^* = \underline{\phi}_t^*(Y_t) \tag{2.3.8}$$

$$\underline{v}_{t+1}^* = \underline{\psi}_{t+1}^*(Y_t) \tag{2.3.9}$$

Conceptually, one may view these functions as

$$\underline{u}_t^* = \underline{\phi}_t^* \, (p(x_t | Y_t, U_{t-1}, V_t)) \tag{2.3.10}$$

$$\underline{v}_{t+1} = \underline{\psi}_{t+1}^* \, (p(\underline{x}_t | Y_t, U_{t-1}, V_t)) \tag{2.3.11}$$

since one will be choosing the values to minimize conditional expectations. Notice, however, that there is apparently an important difference between the pair (2.3.8), (2.3.9) and the pair (2.3.10), (2.3.11). In the first pair (2.3.8), (2.3.9), the optimal controls are expressed as functions of only the past measurements Y_t, while in the second pair (2.3.10), (2.3.11), the optimal controls are expressed as functions of the past measurements Y_t and the past controls U_{t-1} and V_t. There is really no difference, however. If, in the second case, the optimal value of $\underline{v}_0$ in terms of a-priori statistics, the optimal values of $\underline{u}_0$ and $\underline{v}_1$ in terms of Y_0 and v_0, the optimal values of $\underline{u}_1$ and $\underline{v}_2$ in terms of $Y_1, \underline{u}_0, \underline{v}_1, v_0$, etc., are all successively substituted into each other, one can express the optimal U_{t-1} and V_t at time t in terms of Y_t only.

Notice that the fact that optimal U_{t-1} and V_t are used is important. If one is in the "middle" of a dynamic programming solution for the optimal controls, say at time t, the optimal control laws $\underline{u}_0^*(\cdot), \ldots, \underline{u}_{t-1}^*(\cdot)$ and $\underline{v}_0^*(\cdot), \ldots, \underline{v}_t^*(\cdot)$ are not yet known. One is forced to hypothesize "fixed but arbitrary" values for U_{t-1} and V_t, and solve for the optimal controls in the form indicated in (2.3.10) and (2.3.11). It is possible to do this because of Theorem 2.3.1, specifically because only values of past controls are needed to insure that the current density is well-defined. See Remarks 2 and 3 associated with Theorem 2.3.1. One cannot put the optimal controls in the desired form (2.3.8) and (2.3.9) until all the optimal control laws in the form of (2.3.10) and (2.3.11) are known. Only then can one successively substitute to eliminate explicit dependence on U_{t-1} and V_t.

The eventual aim is to apply dynamic programming to the problem. The ideas of Section 2.2 will be used to justify neglecting the accumulated cost at a certain step when optimizing the remaining pieces of trajectory. One must be careful, however, to properly deal with the key idea of dynamic programming that, regardless of present actions, the optimal action will be taken in the future. The wrong way will first be described to point out the problem.

Suppose the time is $t = t_1$, somewhere in the center of the trajectory. The actual cost to go is

$$J_{t_1} = L_T(\underline{x}_T) + \sum_{t=t_1}^{T-1} L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1}) \qquad (2.3.12)$$

It is necessary to convert J_{t_1}, which is a random variable (under suitable hypotheses) into a real valued function of $\underline{u}_{t_1}$ and $\underline{v}_{t_1+1}$. Consider the following two approaches:

$$\mathcal{J}_{3,t_1} = E\{L_T | Y_{t_1}, U_{T-1}, V_T\} + \sum_{t=t_1}^{T-1} E\{L_t | Y_{t_1}, U_{t-1}, V_t\} \quad (2.3.13)$$

$$\mathcal{J}_{4,t_1} = E\{L_T | Y_T, U_{T-1}, V_T\} + \sum_{t=t_1}^{T-1} E\{L_t | Y_t, U_{t-1}, V_t\} \quad (2.3.14)$$

Notice the difference in conditioning. The expression for $\mathcal{J}_{3,t_1}$ given in (2.3.13) is a real valued function of $\{\underline{u}_t\}_{t-t_1}^{T-1}$ and $\{\underline{v}_t\}_{t=t_1+1}^{T}$ since Y_{t_1} is a known quantity. The expression in (2.3.14), however, is a random variable since Y_t for $t > t_1$ is a random variable now, at $t = t_1$. Suppose one tries to convert (2.3.14) to a real valued function by applying another expectation, i.e., by re-defining $\mathcal{J}_{4,t_1}$ as the expected value of the quantity called $\mathcal{J}_{4,t_1}$ in (2.3.13):

$$\mathcal{J}_{4,t_1} = E\{E\{L_T | Y_T, U_{T-1}, V_T\} + \sum_{t=t_1}^{T-1} E\{L_t | Y_t, U_{t-1}, V_t\} | Y_{t_1}, U_{t_1-1}, V_{t_1}\} \quad (2.3.15)$$

Then the following theorem is applicable:

<u>Theorem 2.3.2</u>. (See Doob (D.1), p.37 or Loeve (L.1), p.350)

If $\mathcal{B}$ and $\mathcal{B}'$ are sub-σ-algebras of the underlying probability σ-algebra $\mathcal{S}$ such that $\mathcal{B} \subset \mathcal{B}'$, then for any random variable x

$$E\{E\{x|\mathcal{B}'\}|\mathcal{B}\} = E\{x|\mathcal{B}\} = E\{E\{x|\mathcal{B}\}|\mathcal{B}'\} \quad (2.3.16)$$

with probability one.

In terms of jointly distributed random variables $x, y_1, \ldots, y_n$, the interpretation is that

$$E\{E\{x|y_1, y_2, \ldots, y_n\}|y_1, y_2, \ldots, y_{n-1}\}$$
$$= E\{x|y_1, \ldots y_{n-1}\} \tag{2.3.17}$$

since the σ-field $\mathcal{B}$ induced by the random variables $y_1, y_2, \ldots, y_{n-1}$ is included in the σ-field $\mathcal{B}'$ induced by $y_1, y_2, \ldots, y_{n-1}, y_n$. (The inverse image of the set $\{y_1 \in A_1, y_2 \in A_2, \ldots, y_{n-1} \in A_{n-1}\}$ where A_k are Borel sets in the range space $\mathcal{Y}$ of the y_t is identical to the inverse image of $\{y_1 \in A_1, y_2 \in A_2, \ldots, y_{n-1} \in A_n, y_n \in \mathcal{Y}\}$).

Applying this result to $\mathcal{J}_{4, t_1}$ as defined in (2.3.15) shows that $\mathcal{J}_{4, t_1} = \mathcal{J}_{3, t_1}$ (see 2.3.13)) with probability one. The point is that this interpretation does <u>not</u> account for the fact that measurements will be made in the future. The problem as now formulated does not take into account the manner in which $\underline{u}_{t_1+1}$ and $\underline{v}_{t_1+2}$ will depend on Y_{t_1+1}, etc. What is still being neglected is the point mentioned in Section 2.1, that in choosing optimal controls at present, one must use information concerning the control structure that will be utilized in the future. The present formulation more-or-less assumes open-loop operation in the future.

The underlying defect in the approach given above is that it incorrectly evaluates the optimal cost-to-go from the next stage. This "wrong" method was described in some detail because it will be necessary to refer to it in Section 2.4, when open loop feedback strategies are discussed. Now, however, the correct approach will be presented.

Suppose the time is t = T. The data Y_T are available. The cost-to-go reduces to

$$\mathcal{J}_T = E\{L_T(\underline{x}_T)|Y_T, U_{T-1}, V_T\} \tag{2.3.18}$$

There is nothing left to optimize, so the value of $\mathcal{J}_T$ becomes the optimal cost-to-go from T, which will be denoted ϑ_T

$$\vartheta_T = \int_{R^n} L_T(\underline{x}_T)\, p(\underline{x}_T|Y_T, U_{T-1}, V_T)\, d\underline{x}_T \tag{2.3.19}$$

Note that ϑ_T is a function only of Y_T, U_{T-1}, and V_T

Now step backwards one unit of time. Let t=T-1. The cost-to go is

$$\mathcal{J}_{T-1} = E\{L_{T-1}(\underline{x}_{T-1}, \underline{u}_{T-1}, \underline{v}_{T-1}) + \vartheta_T(Y_T, U_{T-1}, V_T)|Y_{T-1}, U_{T-2}, V_{T-1}\} \tag{2.3.20}$$

Writing this out gives

$$\mathcal{J}_{T-1} = \int_{Rn} L_{T-1}(\underline{x}_{T-1}, \underline{u}_{T-1}, \underline{v}_T)\, p(\underline{x}_{T-1}|Y_{T-1}, U_{T-2}, V_{T-1})\, d\underline{x}_{T-1}$$

$$+ \int_{R^m} {}_T(Y_{T-1}\underline{y}_T, U_{T-1}, V_T)\, p(\underline{y}_T|Y_{T-1}, U_{T-2}, V_{T-1}) d\underline{y}_T \tag{2.3.21}$$

Notice the manner in which Y_T is broken up into the known piece Y_{T-1} and the random variable $\underline{y}_T$ in the argument of ϑ_T. This is to be optimized by choice of $\underline{u}_{T-1}$ and $\underline{v}_T$. The p.d..f. $p(\underline{y}_t|Y_{T-1}, U_{T-2}, V_{T-1})$ may be calculated from the measurement equation (2.1.4), the dynamic equation (2.1.3), the statistics of the noises $\underline{\theta}_T$ and $\underline{\xi}_{T-1}$, and the density $p(\underline{x}_{T-1}|Y_{T-1}, U_{T-2}, V_{T-1})$ (see Papoulis, (P.1), p.118 ff). So far, nothing has been done differently here than in the "wrong" method above. The distinction comes when the optimal $\underline{u}_{T-1}$ and $\underline{v}_T$ functions are resubstituted back into (2.3.20) to give the optimal cost-to-go from t= T-1, denoted by $\mathcal{J}_{T-1}(Y_{T-1}, U_{T-2}, V_{T-1})$. The difference is that the expected value of, for example, the term $L_{T-1}(\underline{x}_{T-1}, u_{T-2}, v_{T-1})$ in (2.3.20) is of a different nature than the expected value of the quantity $L_{T-1}(\underline{x}_{T-1}, \underline{\phi}^*_{T-1}(Y_{T-1}), \underline{\psi}^*_T(Y_{T-1}))$ which is obtained by substituting in the optimal feedback functions $\underline{\phi}_{T-1}(\cdot)$ and $\underline{\psi}^*_T$. Comparing the present approach with the "wrong" method associated with (2.3.13) or (2.3.14), one sees that the earlier technique included no provision for specifying that optimal feedback controls would be used in the future, while the present approach does.

Consider now an arbitrary time $t_1 \in \mathcal{T}$. The cost-to-go is of the form

$$\mathcal{J}_{t_1} = E\{L_{t_1}(\underline{x}_{t_1}, \underline{u}_{t_1}, \underline{v}_{t_1+1}) + \mathcal{J}_{t_1+1}(Y_{t_1+1}, U_{t_1}, V_{t_1+1}) | Y_{t_1}, U_{t_1-1}, V_{t_1}\} \quad (2.3.22)$$

By using the optimal return function $\mathcal{J}_{t_1+1}$ in this cost, one includes the fact that optimal feedback controls will be used in the future. The difference between this and the "wrong" approach is that although in neither case are the values of future optimal controls known at $t=t_1$ (because they depend on measurements not yet made), the functions that will be used to generate those controls are assumed known in the correct stochastic dynamic programming technique. While those future control values are random variables at time = t_1, the minimal expected cost-to-go as a function of Y_{t_1} U_{t_1-1}, and V_{t_1} assuming optimal control functions in the future is a well-defined, real-valued function of a real variable. Equation (2.3.22) will thus be taken as the basic recursion relation of stochastic dynamic programming.

2.4 Some Mathematical Questions

In this section three mathematical ideas will be considered. First, the question of whether or not the entire stochastic optimization problem is well-defined before the control laws are specified will be discussed. This issue is quickly resolved once the dynamic programming framework of the previous section is used as the basis for the problem formulation. Second, the dynamic programming technique itself will be reconsidered in a slightly different light. Finally, the notion of open loop feedback optimal controls will be introduced and discussed.

2.4.1 Why Stochastic Dynamic Programming is Well-Defined Mathematically

Consider the stochastic system specified by (2.1.3) - (2.1.4). As a result of the hypotheses made in Section 1.3, the plant state $\underline{x}_t$, the measurements $\underline{y}_t$ and the cost J are all well-defined random variables on $(\Omega, \mathcal{S}, p)$ for any fixed control sequences $\{\underline{u}_t\}$ and $\{\underline{v}_t\}$, or indeed for any specified set of Borel feedback functions $\{\underline{u}_t(Y_t)\}$, $\{\underline{v}_t(Y_{t-1})\}$ as defined in equations (1.3.5) and (1.3.6). In contrast to this complete structure, however, if some of these feedback functions are not specified while searching for a certain optimal control at a certain time, then the formulation breaks down, since the entire set of states $\{\underline{x}_t\}$ and measurements $\{\underline{y}_t\}$ is not then a well-defined random process.

By viewing the solution procedure as in the previous section, however, this difficulty is eliminated. At each step of the solution, one is considering a problem which is well-defined for two reasons: First, the control functions in the future have already been specified, and second, the entire control problem in the past is being neglected. In other words, when one is choosing $\underline{u}_t^*$ and $\underline{v}_{t+1}^*$ as functions of $p(\underline{x}_t | Y_t, U_{t-1}, V_t)$ by minimizing

$$\mathcal{J}_t = E\{L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1}) + \mathcal{V}_{t+1}(Y_{t+1}, U_t, V_{t+1}) | Y_t, U_{t-1}, V_t\} \qquad (2.4.1)$$

the cost-to-go $\mathcal{V}_{t+1}$ is a well-defined random variable since $\{\underline{u}_{t+1}^*(\cdot), \ldots, \underline{u}_{T-1}^*(\cdot)\}$ and $\{\underline{v}_{t+2}^*(\cdot), \ldots, \underline{v}_T^*(\cdot)\}$ have already been specified. By assuming fixed-but-arbitrary U_{t-1} and V_t, and using Theorem 3.3.1 to generate $p(\underline{x}_t | Y_t, U_{t-1}, V_t)$, one ignores the fact that $\{\underline{u}_0(\cdot), \ldots, \underline{u}_{t-1}(\cdot)\}$ and $\{\underline{v}_0(\cdot), \ldots, \underline{v}_t(\cdot)\}$ are not specified.

2.4.2 A Slightly Different Means to the Same End

In a paper soon to appear, Meier, Larson, and Tether (M.3) consider dynamic programming as a solution technique for linear-quadratic-Gaussian problems. Their paper takes essentially the same point of view as that presented in the previous section, arriving at it from a somewhat different direction. It is instructive to compare the approaches.

First, one must note that in the Meier, Larson, and Tether paper, no measurement control capability is included in the problem formulation. Thus, as discussed in Section 2.1 of this thesis, it is quite natural to neglect the accrued cost at time t when optimizing the remainder of the trajectory since present and future controls will not influence past costs, although estimates of those costs may change. In the Meier paper, this point is not mentioned, as indeed it really need not be. It is important, however, in applying dynamic programming to a problem that does include measurement control to realize that one is forcing a certain structure on the problem.

In the Meier paper, the optimal return function is defined by (in the notation of this thesis)

$$\mathcal{J}_t(\underline{x}_t) = \min_{\underline{u}_t(\cdot), \underline{u}_{t+1}(\cdot), \ldots, \underline{u}_{T-1}(\cdot)} E\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{u}_k) | \underline{x}_t\} \tag{2.4.2}$$

where the state $\underline{x}_t$ is assumed known exactly* and the $\underline{u}_k(\cdot)$ functions have their respective arguments $\underline{x}_k$. Notice that by forcing this particular structure into the definition of $\mathcal{J}_t$, use is made of the fact that optimal closed loop controls will be used in the future, the point that was so important in the presentation of the previous section.

---

*Meier, et. al., consider the problem with imprecisely known state by using a transformation technique that will be discussed in the sequel. At present, the difference in their formulation is not important to the discussion.

Meier, et al., next prove the following theorem in order to arrive at the dynamic programming recursion relation:*

Theorem 2.4.1. Let $\underline{x}$ be a random vector. Then

$$\min_{\underline{u}(\cdot)} E\{f[\underline{x}, \underline{u}(\underline{x})]\} = E\{\min_{\underline{u}} f(\underline{x}, \underline{u})\} \tag{2.4.3}$$

Whenever both sides are defined (i.e., both expectations exist).

Remark 1: Note carefully the meaning of (2.4.3). The left-hand side is a functional minimization: One is required to select the entire function, say $\underline{u}^0(\underline{x})$, which when plugged into the expectation gives its minimum value. The right-hand side of (2.4.3), in contrast, is a pointwise minimization: for each $\underline{x}$, select the value $\underline{u}$ that minimizes $f(\underline{x}, \underline{u})$. As $\underline{x}$ varies, call the resulting (minimizing) collection of optimal $\underline{u}$'s the function $\underline{u}^*(\underline{x})$. Note that on the right hand side, the expectation plays no role at all until the function $\underline{u}^*(\cdot)$ has already been specified.

Remark 2: The theorem does not say that $\underline{u}^0(\underline{x}) = \underline{u}^*(\underline{x})$, but only that the expected returns are the same.

Proof of Theorem 2.4.1. As in the remarks, let $\underline{u}^0(\underline{x})$ be the minimizing function of $\underline{x}$ according to the left-hand side interpretation and let $\underline{u}^*(\underline{x})$ be the function of $\underline{x}$ that minimizes $f(\underline{x}, \underline{u})$. Then by definition of $u^0(x)$

*A proof of this theorem (essentially the same as that of Meier, et.al.) is included here since some of the steps in their proof are not justified, although they are correct, and because similar arguments will be used again later in this thesis.

as the function which minimizes $E\{f(\underline{x}, \underline{u}(\underline{x}))\}$, it follows that

$$E\{f(\underline{x}, \underline{u}^0(\underline{x}))\} \leq E\{f(\underline{x}, \underline{u}(\underline{x}))\} \tag{2.4.4}$$

for <u>any</u> choice of $\underline{u}(\underline{x})$ for which the right-hand side of (2.4.4) exists. In particular, if one chooses $\underline{u}(\underline{x}) = \underline{u}^*(\underline{x})$,

$$E\{f(\underline{x}), \underline{u}^0(\underline{x}))\} \leq E\{f(\underline{x}, \underline{u}^*(\underline{x}))\} \tag{2.4.5}$$

The inequality will now be proved in reverse. By definition of $\underline{u}^*(\underline{x})$ as the function of $\underline{x}$ which minimizes $f(\underline{x}, \underline{u})$, it follows that

$$f(\underline{x}, \underline{u}^*(\underline{x})) \leq f(\underline{x}, \underline{u}(\underline{x})) \tag{2.4.6}$$

for any other function $\underline{u}(\underline{x})$, and <u>for all x.</u> In particular,

$$f(\underline{x}, \underline{u}^*(\underline{x})) \leq f(\underline{x}, \underline{u}^0(\underline{x})) \tag{2.4.7}$$

holds for all $\underline{x}$. Multiplying inequality (2.4.7) by the positive p.d.f. of $\underline{x}$ and integrating over $\underline{x}$ preserves the sense of that inequality, giving

$$E\{f(\underline{x}, \underline{u}^*(\underline{x}))\} \leq E\{f(\underline{x}, \underline{u}^0(\underline{x}))\} \tag{2.4.8}$$

Inequalities (2.4.8) and (2.4.4) together prove the theorem.

Q. E. D.

This theorem is used in developing the dynamic programming recursion relation as follows:

From (2.4.2),

$$\mathcal{J}_t(x_t) = \min_{\underline{u}_t(\cdot),\ldots,\underline{u}_{T-1}(\cdot)} \left\{ E\{ L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{u}_k) | \underline{x}_t\} \right\} \tag{2.4.9}$$

$$= \min_{\{\underline{u}_k(\cdot)\}_{k=t}^{T-1}} \left\{ E \left\{ L_T(\underline{x}_T) + L_t(\underline{x}_t, \underline{u}_t) + \sum_{k=t+1}^{T-1} L_k(\underline{x}_k, \underline{u}_k) | \underline{x}_t \right\} \right\} \tag{2.4.10*}$$

$$= \min_{\underline{u}_t(\cdot)} \left\{ L_t(\underline{x}_t, \underline{u}_t) + \min_{\{\underline{u}_k(\cdot)\}_{k=t+1}^{T-1}} E\left\{ L_T(\underline{x}_T) + \sum_{k=t+1}^{T-1} L_k(\underline{x}_k, \underline{u}_k) | \underline{x}_t \right\} \right\} \tag{2.4.11}$$

or, using the theorem above to interchange min and expectation over $\underline{x}_{t+1}$, which is still a random variable,

$$\mathcal{J}_t(\underline{x}_t) = \min_{\underline{u}_t(\cdot)} \left\{ L_t(\underline{x}_t, \underline{u}_t) + E_{\underline{x}_{t+1}} \left\{ \min_{\{\underline{u}_k(\cdot)\}_{k=t+1}^{T-1}} \left\{ E_{\underline{x}_{t+2},\ldots,\underline{x}_T} \left\{ L_T(\underline{x}_T) + \sum_{k=t+1}^{T-1} L_k(\underline{x}_k, u_k) \; | \underline{x}_t \right\} \right\} \right\} \right\} \tag{2.4.12}$$

which in turn may be written

$$\mathcal{J}_t(x_t) = \min_{u_t(\cdot)} \left\{ L_t(x_t, u_t) + E\left\{ \mathcal{J}_{t+1}(x_{t+1}) | \underline{x}_t \right\} \right\} \tag{2.4.13}$$

---

*Since $\underline{x}_t$ is given, $L_t(x_t, u_t)$ can be taken out of the expectation in going from (2.4.10) to (2.4.11).

This recursion is exactly analogous to (2.3.21), but for the differences in formulation.

Suppose now that one tries to apply this approach to precisely the problem formulated in Section 1.3, that is, to a measurement/control problem with imprecisely known state. There are difficulties. If one starts with the definition

$$\mathcal{J}_t = \min_{\{\underline{u}_k(\cdot)\}_{k=t}^{T-1},\ \{\underline{v}_k(\cdot)\}_{k=t+1}^{T}} E\left\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{u}_k, \underline{v}_{k+1}) \middle| Y_t, U_{t-1}, V_t\right\} \tag{2.4.14}$$

one would like to proceed as above and write

$$\mathcal{J}_t = \min_{\{\underline{u}_t, \underline{v}_{t+1}\}} \min_{\{\underline{u}_k(\cdot), \underline{v}_{k+1}(\cdot)\}_{k=t+1}^{T-1}} E\left\{L_T(\underline{x}_T) + L_t(\underline{x}_t, \underline{u}_t, v_{t+1}) + \sum_{k=t+1}^{T-1} L_k(\underline{x}_k, \underline{u}_k, \underline{v}_{k+1}) \middle| Y_t, U_{t-1}, V_t\right\} \tag{2.4.15}$$

If one now tries to pull the terms involving $\underline{x}_t, \underline{u}_t$, and $\underline{v}_{t+1}$ through minimization on future quantities, one finds that this step is impossible since, for a measurement control problem, <u>future measurements influence estimates of present costs</u>. The method cannot be continued. (Note, by the way, that even if this problem were eliminated, there would remain the necessity of proving a theorem analogous to 2.4.1 where $\underline{u}$ was not allowed to depend on the plant state $\underline{x}$, but only on a jointly distributed (output) random variable, say $\underline{y}$, and with expectations taken over both quantities $\underline{x}$ and $\underline{y}$. Such a theorem is not yet available.) Thus the technique outlined in Section 2.3 is a reasonable definition of stochastic optimality, consistent with intuition, which is identical to analytic approaches applicable to less general problems.

One should note, by the way, that using the definition of dynamic programming given in the previous section, there is no need to worry about interchanging expectation operations with minimizations. The cost to be minimized at step t, given in (2.4.1), is a well-defined quantity. One does not have to put it in any other form to determine the optimal control laws.

2.4.3 Open Loop Feedback Optimal Controls

To conclude this section, the concept of open loop feedback optimal controls (abbreviated O.L.F.O.) will be introduced and discussed. This is not a new concept, (see Dreyfus (D.2), Tse and Athans (T.2), Spang (S.7), and Curry (C.2)) but the discussion of subsection 2.3.3 leads naturally to this idea.

Recall that in defining the correct dynamic programming recursion relation

$$\mathcal{J}_t = E\{L_t(\underline{x}_t,\underline{u}_t,\underline{v}_{t+1}) + \mathcal{J}_{t+1}(Y_{t+1},U_t,V_{t+1})\,|\,Y_t,U_{t-1},V_t\} \tag{2.4.16}$$

the following definition was proposed first:

$$\mathcal{J}_{3,t} = E\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k,\underline{u}_k,\underline{v}_{k+1})\,|\,Y_t,U_{t-1},V_t\} \tag{2.4.17}$$

(see equation (2.3.13)). The quantity $\mathcal{J}_{3,t}$ was rejected as a candidate for a basis of dynamic programming because it provided no means for reflecting the fact that optimal controls would be used in the future (i.e., after time t). The formulation finally established, (2.4.16), did allow this fact to be utilized.

In many actual problems of interest, in fact, for any problem involving non-linear plant dynamics or non-linear measurement equations (that is, non-linear in the plant state) the solution for the true stochastic optimal feedback controls is completely impractical from a computational point of view. Although the concept of obtaining the solution via stochastic dynamic programming is well defined, it is generally impossible to carry out the indicated procedures, either analytically or with currently available digital computers and numerical techniques. For this reason, the study of sub-optimal techniques is of great practical importance.

The O.L.F.O. approach is one such technique. It may generally be described as follows: At each time t in the interval of interest:

1. Determine the statistics of the plant state $\underline{x}_t$ based on the available observations Y_t.

2. ASSUME THAT NO MEASUREMENTS WILL BE MADE IN THE FUTURE (i.e., at any time in the index set $\{t+1, t+2, \ldots, T\}$).

3. Determine the optimal control sequence $\{\underline{u}_t^*, \underline{u}_{t+1}^*, \ldots, \underline{u}_{T-1}^*\}$ based upon the currently available data Y_t.

4. At some time t' greater than t, an additional measurement <u>will</u> be made,* although the contrary hypothesis was made at Step 2. Apply the optimal control sequence determined at Step 3, <u>only</u> up to time t', i.e., apply $\{\underline{u}_t^*, \underline{u}_{t+1}, \ldots, u_{t'}\}$.

5. Redefine the set of available data to include the new measurement. Go to Step 1.

---

*It is perfectly possible that t' may be t+1, that is, one step later than present.

Note why the description "open loop feedback optimal control" strategy is appropriate. The method is open loop, because each time a measurement is made, the entire future control time function is determined as though there will be no future measurements. The method is a feedback method because each time the control history is determined, it is as a function of the current data. The dual nature of the method is perhaps most apparent when, for each iteration of Steps 1-5 above, $t' = t+1$. This would have the result that for every t, the entire open loop optimal control sequence from time t to time $T-1$ is computed, but only $\underline{u}_t^*$ is applied. Thus for each t, $\underline{u}_t^*(Y_t)$ is generated, but the star in this context means O.L.F.O. rather than "optimal assuming future closed loop optimal controls."

Two questions arise: What is the advantage of this (sub-optimal) technique, and how would measurement control capability be incorporated into such a problem?

The advantage is basically that the computations involved in finding an O.L.F.O. control sequence are simpler than those required to find a true closed loop optimal. The assumption of no measurements in the future allows one to propagate the statistics of the plant state from time t to time T using equation (2.3.1) (deterministic) rather than (2.3.2) (stochastic). The appropriate cost functional assuming no future measurements is (2.4.17) (with the $\underline{v}_{k+1}$ influence deleted) since in the O.L.F.O. framework, $\{\underline{u}_k\}_{k=t}^{T-1}$ is assumed deterministic given Y_t. Thus the O.L.F.O. approach requires the solution of a sequence of infinite dimentional deterministic problems rather than a truly stochastic infinite dimensional problem.

2 9

The role of measurement controls in an O.L.F.O. problem is somewhat different than in a true closed loop stochastic formulation. Since at each step in an O.L.F.O. solution, the assumption of no future measurements is made to determine the dynamic controls, it is not natural to couple to this determination an evaluation of optimal measurement controls. One must define a subsidiary problem which will allow the determination of when the next measurement should be taken (i.e., when is time t' in Step 4 above) and "how much" measuring to do at time t' (if one has more than only yes-no control over the measurement system). This subsidiary problem might, for example, involve the (deterministic) quantity $p(\underline{x}_k | Y_t)$: the p.d.f. of future states given the measurements already taken and given an O.L.F.O. approach to propagation of that density. One might, for example, determine to measure again if $(p(\underline{x}_k | Y_t)$ gets too "broad" (high variance). The important point is that if one adopts an O.L.F.O. philosophy, one automatically gets a "one-way separation property" in the optimal measurement/control solution: at each stage of the O.L.F.O. solution procedure, one determines the optimal dynamic control independently of the specification of the measurement control.

2.5 Summary

In this chapter, the basic structure of stochastic dynamic programming was discussed. The meaning of stochastic optimality was studied. The notion was developed that the designer must formulate an overall mathematical structure having the property that the map relating the control sequences to the cost is real valued rather than random. It was emphasized that in applying dynamic programming, two points must be considered: first, one must neglect that part of the problem extending from time zero to time t when one is optimizing the trajectory from time t to the terminal time T, and second, one must be careful to include in the dynamic programming structure a provision for utilizing the fact that regardless of present actions, optimal actions will be carried out in the future. The result of these considerations was the basic dynamic recursion equation (2.3.22).

In Section 2.4, some mathematical issues were considered. The potential problem of leaving the mathematical structure of the problem incomplete while searching for control laws was eliminated. The technique used to formulate dynamic programming in Section 2.3 was compared to more conventional approaches, with particular emphasis on the issues introduced by including measurement control capability in the problem. Finally, the notion of open loop feedback optimal control was discussed.

CHAPTER III

DETERMINISTIC TECHNIQUES APPLIED TO STOCHASTIC PROBLEMS

Athans' Axiom

> Every problem can be reformulated so that the Minimum Principle is the key to obtaining the solution.

3.1 General Perspective

The study of deterministic optimal control problems has led to the development of a number of powerful analytic tools, notably the Minimum Principle, (see (P.2), (A.7), (K.5), (A.8)) which are not directly applicable to the solution of stochastic problems. Some techniques, for example, dynamic programming, are applicable to both classes of problems, but it is somehow unfortunate that Minimum Principle ideas cannot be used to gain further insight into the nature of stochastic systems. In this chapter, certain circumstances will be described under which it is possible to use deterministic techniques to approach stochastic problems.

The obvious difficulty with using the Minimum Principle to solve a stochastic problem is that it is basically an open loop technique. When using the Maximum Principle, one assumes a fixed initial state for the system, and one then constructs a two-point boundary value problem involving the plant state of the given dynamic system and the so-called costate dynamic system, which is of the same dimension as the state. The solution of this two-point boundary value problem then gives the optimal control as a function of the state and costate trajectories. Practically speaking, it is not possible, in general, to obtain this solution analytically, and numerical techniques must be used. The result is that one can find the open loop optimal control time function corresponding to any initial state, but not a closed loop solution,* and closed loop solutions are required in stochastic problems. Even if analytic solutions of the two point boundary value problem are obtained, it is generally not possible to eliminate the dependence of the optimal control on the costate, and thus achieve a solution in terms of state (only) feedback. Even if it was possible to express the solution in terms of state feedback, it is not conveient to introduce the concept of an inexactly known state. This, in fact, is the main difficulty in a sense. Deterministic techniques in general and the Minimum Principle in particular assume exact knowledge of all quantities of interest.

It will be shown in this chapter that if one is dealing with a linear system influenced by noises having Gaussian distributions, it is possible to reformulate the stochastic control problem in such a way that the Minimum Principle can be used as a starting point in the solution. The problem of inexact knowledge of the system state is circumvented by reformulating the problem in terms of the conditional mean and conditional covariance of the state, both of which are exactly know quantities which can be generated by a Kalman filter. The problem of inexact a-priori knowledge of what the actual noises will be (i.e., what sample

---

* It is sometimes possible to generate feedback solutions using the Minimum Principle as a starting point. See the treatment of the linear regulator problem in Athans (A.7), and below.

values will occur over a particular run) is circumvented by assuming the noises are deterministic and taking the expected value at the end.

Both of these devices must be carefully justified. Let it be noted here, however, that the reasons for that assuming the noises are Gaussian are to preserve the simple filter structure and to allow reformulation only in terms of means and covariances, parameters which are necessary and sufficient to completely characterize a Gaussian process but not a general random process. The reasons for assuming the system is linear are to preserve the filter structure and to allow proof of a theorem which justifies "*taking the expected value at the end*".

The transformation approach to be introduced in this chapter appears in the paper of Meier, Laison and Tether (M.3), although it is not carefully justified. In particular, those authors "leave to the reader" the proof of Theorem 3.3.1 below, which is crucial to the technique. The idea is closely related to the work of Kailath and Frost (K.6), (K.7), (F.1). Theorem 3.3.2, however, is novel.

3.2 Mathematical Formulation: Linear System with Gaussian Noise

Let the plant state evolve according to the difference equation

$$\underline{x}_{t+1} = \underline{A}_t \underline{x}_t + \underline{B}_t \underline{u}_t + \underline{\xi}_t \qquad (3.2.1)$$

where t is an integer in the index set $\mathcal{J} = \{0, 1, 2, \ldots, T\}$, $\underline{x}_t \in R^n$, $\underline{u}_t \in R^m$, $\underline{\xi}_t \in R^n$, $\underline{A}_t$ is an nxn matrix, and $\underline{B}_t$ is an n x m matrix. The sequence $\{\underline{\xi}_0, \underline{\xi}_1, \ldots, \underline{\xi}_{T-1}\}$ is a random process on $(\Omega, \mathcal{B}, p)$ such that $\underline{\xi}_i$ is independent of $\underline{\xi}_j$, $i \neq j$, and $\underline{\xi}_i$ has a Gaussian distribution with mean zero and covariance matrix Ξ_t, which is positive semidefinite:

$$E\{\underline{\xi}_t\} = \underline{0} \qquad (3.2.2)$$

$$E\{\underline{\xi}_t\underline{\xi}_s'\} = \underline{\Xi}_t\delta_{ts} \qquad (3.2.3)$$

where δ_{ts} is the Kronecker delta. The initial state $\underline{x}_0$ has a Gaussian distribution with mean $\hat{\underline{x}}_0$ and covariance $\underline{\Sigma}_0$:

$$E\{\underline{x}_0\} = \hat{\underline{x}}_0 \qquad (3.2.4)$$

$$E\{(\underline{x}_0 - \hat{\underline{x}}_0)(\underline{x}_0 - \hat{\underline{x}}_0)'\} = \underline{\Sigma}_0 \qquad (3.2.5)$$

The driving noise $\underline{\xi}_t$ is assumed to be independent of $\underline{x}_0$ for all t.

Several possible measurement equations make sense in the linear case. Since the measurement is <u>not</u> required to be linear in the measurement control $\underline{v}_t$, the most general form is

$$\underline{y}_t = \underline{C}_t(\underline{v}_t)\underline{x}_t + \underline{D}_t(\underline{v}_t)\underline{\theta}_t \qquad (3.2.6)$$

where the matrices $\underline{C}_t$ and $\underline{D}_t$, which are functions of $\underline{v}_t$, are respectively r x n and r x p with $\underline{D}_t(\underline{v}_t)$ of full rank for all $\underline{v}_t$. The sequence $\{\underline{\theta}_0, \underline{\theta}_1, \ldots, \underline{\theta}_T\}$ is composed of independent, zero mean, Gaussian random vectors independent of $\{\underline{\xi}_t\}$ and $\underline{x}_0$, with covariance matrices $\underline{\Theta}_t$ of full rank*

$$E\{\underline{\theta}_t\} = \underline{0} \qquad (3.2.7)$$

$$E\{\underline{\theta}_t\underline{\theta}_s'\} = \underline{\Theta}_t\delta_{ts} \qquad (3.2.8)$$

---

* The case where Θ_t and/or $\underline{D}_t$ are not full rank can be treated with a certain amount of additional care. See Tse (T.1). In that case, some matrix inverses that appear below may not exist, and certain complications result.

Several special cases of equation (3.2.6) are of particular interest. Consider for example,

$$\underline{y}_t = v_t \underline{C}_t \underline{x}_t + \underline{\theta}_t \tag{3.2.9}$$

or alternatively

$$\underline{y}_t = v_t [\underline{C}_t \underline{x}_t + \underline{\theta}_t] \;\; ; \;\; v_t \epsilon \{0, 1\} \tag{3.2.10}$$

In both cases, the control v_t is a scalar. The situation represented by equation (3.2.9) can be interpreted more or less as signal-to-noise ratio control: when v_t is large, the "signal" $\underline{C}_t \underline{x}_t$ is boosted with respect to the noise $\underline{\theta}_t$. In equation (3.2.10), the measurement is of the measure-no measure type: if $v_t = 1$, a measurement of fixed "quality" is obtained; if $v_t = 0$, no measurement is made. Another special case of (3.2.6) which is a useful model of certain physical situation is

$$\underline{y}_t = [v_{1,t} \underline{C}_{1,t} + v_{2,t} \underline{C}_{2,t} + \ldots + v_{k,t} \underline{C}_{k,t}] \underline{x}_t + \underline{\theta}_t \tag{3.2.11}$$

where each $v_{i,t}$ is constrained to be either zero or one and the additional constraint

$$\sum_{i=1}^{k} v_{i,t} = 1 \tag{3.2.12}$$

is imposed. This represents a situation in which the measurement control selects one out of k possible measurement configurations. See Athans (A.6).

3.3 An Alternative Approach to Stochastic Control of Linear Systems

3.3.1 Transformation of the System

A fundemental question in stochastic control might be phrased: "How much information is required to solve the problem at hand?". This question is related to the notion of sufficient statistics as discussed by Striebel (S.6) and to the notion of what constitutes the state in a stochastic control problem, as discussed in Chapter II of this thesis.

Given that the state of a stochastic problem at some time t is a conditional p.d.f., the question of how to properly propagate this state arises. Striebel (S.6) showed that the density can be propagated as in equations (2.3.1) or (2.3.2) (she has the analogous measure-theoretic equations) using only the values of dynamic controls $\underline{u}_t$, i.e., without requiring knowledge of complete control laws $\underline{u}_t(\cdot)$. It follows from this idea (see (S.6) and (W.2)) that if one considers a linear system with Gaussian initial state and Gaussian noises as defined in Section 3.2, the conditional p.d.f. of $\underline{x}_t$ given the measurements Y_t will be Gaussian regardless of the form of the feedback (linear or nonlinear) if one knows the values of past controls. Thus at any time t, the density of $\underline{x}_t$ for the system (3.2.1), (3.2.6) is completely characterized by its conditional mean $\hat{\underline{x}}_{t|t}$ and conditional covariance $\underline{\Sigma}_{t|t}$:

$$\hat{\underline{x}}_{t|t} = E\{\underline{x}_t | Y_t, U_{t-1}, V_t\} \tag{3.3.1}$$

$$\underline{\Sigma}_{t|t} = \text{cov}(\underline{x}_t, \underline{x}_t | Y_t, U_{t-1}, V_t) = E\{(\underline{x}_t - \hat{\underline{x}}_{t|t})(\underline{x}_t - \hat{\underline{x}}_{t|t})' | Y_t, U_{t-1}, V_t\} \tag{3.3.2}$$

Note now that it is possible to generate $\hat{\underline{x}}_{t|t}$ and $\underline{\Sigma}_{t|t}$ using a Kalman filter so long as the system is linear and the noises Gaussian. The equations take the form

$$\hat{\underline{x}}_{t+1|t+1} = \underline{A}_t\hat{\underline{x}}_{t|t} + \underline{B}_t\underline{u}_t + \hat{\underline{\xi}}_t \quad (3.3.3)$$

$$\underline{\Sigma}_{t+1|t+1} = \{\underline{I} - (\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}'[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}'$$
$$+ \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']^{-1}\underline{C}_{t+1}\}(\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t) \quad (3.3.4)$$

where

$$\hat{\underline{\xi}}_t = (\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}'[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}'$$
$$+ \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']^{-1}[\underline{y}_{t+1} - \underline{C}_{t+1}(\underline{A}_t\hat{\underline{x}}_{t|t} + \underline{B}_t\underline{u}_t)] \quad (3.3.5)$$

The quantity $\hat{\underline{\xi}}_t$ is the so-called residual or innovation in the Kalman filter. An important property of the filter is that the sequence $\{\hat{\underline{\xi}}_t\}$ is zero mean and "white", with Gaussian statistics. (see Bryson Ho (B.4)). In fact, from (3.3.5), one may calculate the covariance matrix of $\hat{\underline{\xi}}_t$ and thereby specify its density exactly:

$$\text{cov}(\hat{\underline{\xi}}_t, \hat{\underline{\xi}}_t) = E\{\hat{\underline{\xi}}_t\hat{\underline{\xi}}_t'\}$$
$$= \underline{S}_{t+1}[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']\underline{S}_{t+1}'$$
$$(3.3.6)$$

where

$$\underline{S}_{t+1} = (\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}'[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}'_{t|t}\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}'$$
$$+ \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']^{-1} \quad (3.3.7)$$

What has been done up to now is to take the original stochastic system (3.2.1) - (3.2.6) and formulate the appropriate equations for propagating the conditional p. d. f. of $\underline{x}_t$ given in the data. The original n-dimensional system has been replaced by the $n + n^2$ dimensional system (3.3.3) - (3.3.4) (actually $n + \frac{n^2}{2} + \frac{n}{2}$ since $\underline{\Sigma}_{t|t}$ is symmetric for all t). Note that (3.3.3) and (3.3.4) still represent a stochastic system since the mean equation (3.3.3) is driven by the white noise $\hat{\underline{\xi}}_t$. However there is now no measurement equation. The mean $\hat{\underline{x}}_{t|t}$ is known exactly at time t. The measurement aspects of the problem are contained in the dependence of $\underline{C}_t$ and $\underline{D}_t$ on $\underline{v}_t$, which in turn influences the propagation of the covariance and the statistics of $\{\hat{\underline{\xi}}_t\}$.

The original plant state vector $\underline{x}_t$ has thus been supressed in the equations of motion, and inexact knowledge of the "state" in the dynamic equations has been eliminated as a difficulty. It remains to transform the cost functional into a form involving only $\hat{\underline{x}}_{t|t}, \underline{\Sigma}_{t|t}, \underline{u}_t$ and $\underline{v}_t$. This may be done by noting that if $\underline{x}$ is a Gaussian random vector with mean $\hat{\underline{x}}$ and covariance $\underline{\Sigma}$, and if $\underline{u}$ is deterministic, then one can express the expected value of any function $L(\underline{x}, \underline{u})$ for which the expectation exists in terms of $\hat{\underline{x}}$, $\underline{\Sigma}$, and $\underline{u}$. Let script $\mathcal{L}$ functions be defined by

$$E\{L(\underline{x}, \underline{u})\} = \mathcal{L}(\hat{\underline{x}}, \underline{\Sigma}, \underline{u}) \tag{3.3.6}$$

Suppose that the given cost functional is the random variable

$$J = L_T(\underline{x}_T) + \ell_0(\underline{v}_0) + \sum_{t=0}^{T-1} L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1}) \tag{3.3.7}$$

The expected value of this random variable is to be minimized by generating $\{\underline{u}_t\}$ and $\{\underline{v}_t\}$ using feedback as discussed in Chapter II. The claim is that one can equivalently consider the random variable

$$\hat{J} = \mathcal{L}_T(\hat{\underline{x}}_{T|T}, \underline{\Sigma}_{T|T}) + \ell_0(\underline{v}_0) + \sum_{t=0}^{T-1} \mathcal{L}_t(\hat{\underline{x}}_{t|t}, \underline{\Sigma}_{t|t}, \underline{u}_t, \underline{v}_{t+1})$$

(3.3.8)

where this new quantity (a random variable because $\{\hat{\underline{x}}_{t|t}\}$ is a random process a-priori, i.e., at time t = 0) is to be minimized in the same sense, that is, as described in Chapter II. More formally, the following theorem may be stated:

Theorem 3.3.1 Let the following control problem be specified:

$$\left.\begin{aligned} &\text{Dynamics: } \underline{x}_{t+1} = \underline{A}_t\underline{x}_t + \underline{B}_t\underline{u}_t + \underline{\xi}_t \\ &\text{Measurements: } \underline{y}_t = \underline{C}_t(\underline{v}_t)\underline{x}_t + \underline{D}_t(\underline{v}_t)\underline{\theta}_t \\ &\text{Cost: } J = L_T(\underline{x}_T) + \ell_0(\underline{v}_0) + \sum_{t=0}^{T-1} L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1}) \end{aligned}\right\} \quad (3.3.9)$$

where the variables have the meanings given in Section 3.2 and where the optimal controls are to be determined by stochastic dynamic programming as defined in Chapter II. Then the optimal dynamic control functions $\{\underline{u}_t^*(\cdot)\}$ and $\{\underline{v}_t^*(\cdot)\}$ for the system (3.3.9) are identical to those obtained for the following system, also treated by stochastic dynamic programming:

$$\text{Dynamics: } \hat{\underline{x}}_{t+1|t+1} = \underline{A}_t\hat{\underline{x}}_{t|t} + \underline{B}_t\underline{u}_t + \hat{\underline{\xi}}_t$$

$$\underline{\Sigma}_{t+1|t+1} = \{\underline{I} - (\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}'[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']^{-1}\underline{C}_{t+1}'\}(\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t)$$

Measurements: $\hat{x}_{t|t}$ known exactly

$$\text{Cost:} \quad J = \mathcal{L}_T(\hat{\underline{x}}_{T|T}, \underline{\Sigma}_{t|t}) + \ell_0(\underline{v}_0) + \sum_{t=0}^{T-1} \mathcal{L}_t(\hat{\underline{x}}_{t|t}, \underline{\Sigma}_{t|t}, \underline{u}_t, \underline{v}_{t+1}) \tag{3.3.10}$$

$$\text{where } \mathcal{L}_t(\hat{\underline{x}}_{t|t}, \underline{\Sigma}_{t|t}, \underline{u}_t, \underline{v}_{t+1}) = E\{L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1}) | \underline{x}_t = N(\hat{\underline{x}}_{t|t}, \underline{\Sigma}_{t|t})\} \tag{3.3.11*}$$

Proof (by induction): At step T, $\underline{x}_T$ is Gaussian with mean $\hat{\underline{x}}_{T|T}$ and covariance $\underline{\Sigma}_{T|T}$. Thus the cost-to-go associated with system (3.3.9) is

$$\mathcal{J}_T = E\{L_T(\underline{x}_T) | \underline{x}_T = N(\hat{\underline{x}}_{T|T}, \underline{\Sigma}_{T|T})\} \tag{3.3.12}$$

$$\triangleq \mathcal{L}_T(\hat{\underline{x}}_{T|T}, \underline{\Sigma}_{T|T}) \tag{3.3.13}$$

which is exactly the cost-to-go associated with system (3.3.10).

At step t, assume that the costs-to-go associated with (3.3.9) and (3.3.10) from step t+1 are identical functions of $\hat{\underline{x}}_{t+1|t+1}$ and $\underline{\Sigma}_{t+1|t+1}$, say $\mathcal{V}_{t+1}(\underline{x}_{t+1|t+1}, \underline{\Sigma}_{t+1|t+1})$. Then the cost-to-go from step t associated with system (3.3.9) is

$$\mathcal{J}_t = E\{L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1}) + \mathcal{V}_{t+1}(\underline{x}_{t+1|t+1}, \underline{\Sigma}_{t+1|t+1}) | Y_t, U_{t-1}, V_t\} \tag{3.3.14}$$

$$= \mathcal{L}_t(\hat{\underline{x}}_{t|t}, \underline{\Sigma}_{t|t}, \underline{u}_t, \underline{v}_{t+1}) + E\{\mathcal{V}_{t+1}(\underline{x}_{t+1|t+1}, \underline{\Sigma}_{t+1|t+1}) | Y_t, U_{t-1}, V_t\} \tag{3.3.15}$$

---

* The notation $\underline{x} = N(\hat{\underline{x}}, \underline{\Sigma})$ will be used to indicate that $\underline{x}$ is a Gaussian (normal) random vector with mean $\hat{\underline{x}}$ and covariance $\underline{\Sigma}$.

By the innovations theorem of Kailath (K.6), expectations with respect to the noises and measurements in (3.3.9) are equivalent to expectations with respect to the innovations process $\{\hat{\underline{\xi}}_t\}$ in (3.3.10), so (3.3.15) is identical to the cost-to-go associated with system (3.3.10). Since the systems (3.3.9) and (3.3.10) have identical costs-to-go, they have identical optimal control laws. They also have identical optimal costs-to-go at step t, showing that the induction holds and proving the theorem.

Q.E.D.

3.3.2 Using the Transformed System

The previous subsection justified the notion that the original given system with inexact knowledge of the state can be reformulated in terms of the Kalman filter equations involving the conditional mean and conditional covariance of the state, both of which are known. It was noted, however, that the filter equations are still stochastic in nature since the mean equation (3.3.3) is driven by the white noise process $\{\hat{\underline{\xi}}_t\}$. As such, the transformed system still cannot be analyzed using the Minimum Principle since that technique requires full a-priori knowledge of all the drives in the equations of motion, i.e., in the present context, of $\{\hat{\underline{\xi}}_t\}$. The following theorem allows the elimination of this last difficulty:

Theorem 3.3.2 Let the sequence of functions $\{L_t(\underline{x}_t, \underline{u}_t)\}$ be convex Borel functions from R^{n+m} into R. Let $\underline{x}_t$ be the plant state of a stochastic system

$$\underline{x}_{t+1} = \underline{A}_t \underline{x}_t + \underline{B}_t \underline{u}_t + \underline{\xi}_t$$

where $\{\underline{\xi}_t\}$ is a white noise process and where at step t, $\underline{x}_t$ is known exactly. Let $\mathcal{J}_t(\underline{x}_t)$ be defined

$$\mathcal{J}_t(\underline{x}_t) = \min_{\underline{u}_t(\cdot), \underline{u}_{t+1}(\cdot), \ldots, \underline{u}_{T-1}} E\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{u}_k) | \underline{x}_t\} \quad (3.3.16)$$

and let $\hat{\mathcal{J}}_t(\underline{x}_t)$ be defined as follows:

$$\{\hat{\underline{u}}_t(\underline{x}_t, \underline{\xi}_t, \ldots, \underline{\xi}_{T-1})\} = \text{deterministic optimal control functions assuming } \{\underline{\xi}_t\} \text{ fixed, known. Assume these are Borel functions.} \quad (3.3.17)$$

$$\{\bar{\hat{\underline{u}}}_t(\underline{x}_t)\} = E\{\hat{\underline{u}}_t(\underline{x}_t, \underline{\xi}_t, \ldots, \underline{\xi}_{T-1}) | \underline{x}_t\} \text{ (expected value over } \{\underline{\xi}_k\}) \quad (3.3.18)$$

$$\hat{I}_t(\underline{x}_t, \underline{\xi}_t, \ldots, \underline{\xi}_{T-1}) = \text{cost incurred by using the control sequence } \{\bar{\hat{\underline{u}}}_k(\underline{x}_k)\}_{k=t}^{T-1} \quad (3.3.19)$$

$$\hat{\mathcal{J}}_t(\underline{x}_t) = E\{\hat{I}_t(\underline{x}_t, \underline{\xi}_t, \ldots, \underline{\xi}_{T-1}) | \underline{x}_t\} \quad (3.3.20)$$

then

$$\mathcal{J}_t(\underline{x}_t) = \hat{\mathcal{J}}_t(\underline{x}_t) \quad (3.3.21)$$

Remark 1: What this theorem does is compare the costs resulting from two methods of computing a sequence of controls. Equation (3.3.16) defines $\mathcal{J}_t(\underline{x}_t)$ as the minimal cost-to-go using stochastic dynamic programming as in Chapter II. Equation (3.3.20) defines $\hat{\mathcal{J}}_t(\underline{x}_t)$ through the following sequence of steps:

(1) Assume $\{\underline{\xi}_k\}_{k=t}^{T-1}$ is fixed but arbitary

(2) Solve the resulting deterministic control problem in feed back form, obtaining optimal controls $\hat{\underline{u}}_k = \hat{\underline{u}}_k(\underline{x}_k, \underline{\xi}_k, \ldots, \underline{\xi}_{T-1})$ (Equation (3.3.17))

(3) Take the expected value of the $\{\hat{\underline{u}}_k\}$ sequence over the quantities $\{\underline{\xi}_k\}$ which are actually random variables (Equation (3.3.18)). The result is $\{\bar{\hat{\underline{u}}}_k(\underline{x}_k)\}$

(4) Apply the control laws $\{\bar{\hat{\underline{u}}}_k(\underline{x}_k)\}$. Express the resulting cost as a function of $\{\underline{\xi}_k\}$, again assumed fixed but arbitrary, and $\underline{x}_t$. Call the result $\hat{I}_k(\underline{x}_t, \underline{\xi}_t, \ldots, \underline{\xi}_{T-1})$ (Equation (3.3.19))

(5) Take the expected value of that cost:

$$\hat{\mathcal{J}}_t(\underline{x}_t) = E\{\hat{I}_t(\underline{x}_t, \underline{\xi}_t, \ldots, \underline{\xi}_{T-1}) | \underline{x}_t\}.$$

The theorem says that two things are equal: the cost-to-go using true stochastic dynamic programming and the expected value of the cost-to-go obtained by applying controls which are themselves the expected values of the optimal control laws obtained by assuming the noise is fixed but arbitrary.

Remark 2: Note carefully that the theorem does not state that $\underline{u}_t^*(\underline{x}_t) = \bar{\hat{\underline{u}}}_t(\underline{x}_t)$, i.e., that the optimal stochastic control equals the average of the optimal deterministic control assuming the noise fixed. Only the resulting costs are equal, on the average.

Remark 3: A point that must be stressed for future reference is that the theorem stated here has no explicit measurement control features. This theorem will be applied below to systems of the type described by (3.3.10) (Kalman filter systems with transformed costs representing linear stochastic systems of the form of (3.3.9)). The noise $\underline{\xi}_t$ of this theorem will then have statistics which are influenced by the measurement control in combined measurement/dynamic control problems. This fact will be supressed during the deterministic part of the analysis, i.e., when the noise $\{\underline{\xi}_t\}$ is assumed fixed to determine $\hat{\underline{u}}_t(\underline{x}_t, \underline{\xi}_t, \ldots, \underline{\xi}_{T-1})$. However the manner in which the measurement control influences the problem will re-appear when the two expected values of steps 3 and 5 in Remark 1 are taken.

Remark 4: One will note below that the assumption that the noise sequence $\{\underline{\xi}_t\}$ is white is not used explicitly in the proof of the theorem. In fact, this assumption is not necessary to prove the theorem as it stands. This is because the theorem makes a statement concerning the optimal cost-to-go from time t to the end of the interval, a quantity that is well defined whether or not the noise is white.

Consider, however what would occur if the noise was not white and one were to apply this theorem at two different times t_1 and t_2. Let t_2 be greater than t_1. If the noises are not white, then the statistics of the sequence $\{\underline{\xi}_{t_2}, \underline{\xi}_{t_2+1}, \ldots, \underline{\xi}_{T-1}\}$ assuming $\underline{x}_{t_2}$ known will in general be different than the statistics of the same sequence viewed as a subsequence of $\{\underline{\xi}_{t_1}, \underline{\xi}_{t_1+1}, \ldots, \underline{\xi}_{t_2}, \ldots, \underline{\xi}_{T-1}\}$ with $\underline{x}_{t_1}$ known. If the noises are white, the statistics of the shorter sequence will be the same in both cases.

Suppose now that one uses this theorem in the solution of a stochastic optimal control problem with non-white noise. One starts at time t_1 and determines $\{\hat{\bar{\underline{u}}}_t(\cdot)\}_{t=t_1}^{T-1}$. One applies this control law sequence until $t = t_2$. One discards the old solution and resolves for $\{\hat{\bar{\underline{u}}}_t(\cdot)\}_{t=t_2}^{T-1}$ starting at $t = t_2$. This new solution will in general be different from the "unused" part of the old solution since the statistics of the future noises have changed.

This situation is not a violation of the Principle of Optimality (see Bellman, (B.3)). It is merely a result of the fact that the quantity $\underline{x}_t$ is not the entire state of the system. As pointed out in Section 2.2, if the noise is non-white, one must include in the state of the system the statistics of the future noises given the present measurements. Since this information is different at times t_1 and t_2 in the above analysis, which in effect neglects part of the state, it is not unexpected that different control laws will result from analyses at t_1 and t_2.

Remark 5: How will this theorem be useful? Because of this theorem, it becomes possible to use the Minimum Principle or other deterministic techniques to gain insight into stochastic problems. One first converts to the filter system as described in subsection 3.3.1. Next, one assumes a fixed but arbitrary innovations sequence $\{\hat{\underline{\xi}}_t\}$ and analyzes the resulting deterministic system. According to Theorem 3.3.2, the expected cost that would be obtained by using the expected value of the feedback controls derived from deterministic techniques is the same as that obtained using stochastic dynamic programming. Note again that it does not necessarily follow that the stochastic feedback law and the expected value of the deterministic feedback law are identical.* The result only says that the expected costs are the same. In the absence of a theorem regarding the uniqueness of stochastic optimal feedback control laws this is the best that can be hoped for. It does, however, allow one to use deterministic control laws to study a solution at least as good as the true stochastic one.

Remark 6: Consider the implications of Theorem 3.3.2 when the Maximum Principle cannot be manipulated in such a way as to provide the optimal feedback control laws $\hat{\underline{u}}(\underline{x}_t, \underline{\xi}_t, \ldots, \underline{\xi}_{T-1})$ called for in the theorem statement. Suppose that one can only determine an optimal open loop time sequence $\{\hat{\underline{u}}_t, \hat{\underline{u}}_{t+1}, \ldots, \hat{\underline{u}}_{T-1}\}$ given $\underline{x}_t$ and the (fixed) noise sequence $\{\underline{\xi}_k\}_{k=t}^{T-1}$. What can be said about the applicability of the theorem?

The theorem still holds. One can view the process of constructing the two-point boundary value problem called for by the Minimum Principle for any initial state $\underline{x}_t$ and any noise sequence $\{\underline{\xi}_t\}$ as a feedback solution: even though one gets a time function for an answer to the optimal control problem, one gets a particular time function for every initial state and every noise sequence. As discussed in Section 1.2, there is an equivalence between open loop and closed loop optimal controls in the deterministic case. One can average over this process.

* It will be shown in the next chapter that these two feedback laws are identical in the linear-Gaussian-quadratic case.

There is a technical difficulty with using this approach, however. If one takes the average over time functions obtained by considering optimal two-point boundary value problems and applies the average control at time t, $\bar{\hat{\underline{u}}}_t$, there is no guarantee that the average optimal next state $\bar{\hat{\underline{x}}}_{t+1} = \underline{A}_t \bar{\hat{\underline{x}}}_t + \underline{B}_t \bar{\hat{\underline{u}}}_t + \bar{\underline{\xi}}_t$ will occur. When one observes at time t+1 that one is not in the "right" state after applying $\bar{\hat{\underline{u}}}_t$, one must completely resolve the problem, because the optimal controls are not in explicit feedback form.

Proof of Theorem 3.3.2:

I. Definition: A function $\phi(\cdot): R^m \rightarrow R$ is convex if for all $\underline{u}_1, \underline{u}_2 \epsilon R^m$ and for all $\alpha, \beta \epsilon [0, 1]$, $\alpha + \beta = 1$

$$\phi(\alpha \underline{u}_1 + \beta \underline{u}_2) \leq \alpha \phi(\underline{u}_1) + \beta \phi(\underline{u}_2) \quad (3.3.22)$$

II. Lemma 3.3.1: (Loeve, (L. 1), p. 159) If $\phi(\cdot)$ is a convex function* and $\underline{u}$ is a random variable such that $E\{\underline{u}\}$ is finite,

$$E\{\phi(\underline{u})\} \geq \phi(E\{\underline{u}\}) \quad (3.3.23)$$

III. Lemma 3.3.2: Let $\phi(\cdot)$ be a convex function on R^m and let $\psi(\underline{u}): R^m \rightarrow R^m$ be affine, i.e., $\psi(\underline{u}) = \mathcal{L}(\underline{u}) + \underline{c}$ where $\mathcal{L}(\cdot)$ is linear. Then the composition $\phi o \psi$ is convex

Proof:

$$(\phi o \psi)(\alpha u_1 + \beta u_2) \triangleq \phi[\psi(\alpha u_1 + \beta u_2)]$$

---

* It also is stated on p. 159 of Loeve that any convex function is either continuous or not Borel, by the way.

$$= \phi[\mathcal{L}(\alpha\underline{u}_1 + \beta\underline{u}_2) + \underline{c}]$$

$$= \phi[\alpha\mathcal{L}(\underline{u}_1) + \beta\mathcal{L}(\underline{u}_2) + \underline{c}]$$

$$= \phi[\alpha\mathcal{L}(\underline{u}_1) + \beta\mathcal{L}(\underline{u}_2) + (\alpha+\beta)\underline{c}]$$

Since $\alpha+\beta = 1$. Then

$$(\phi o \psi)(\alpha\underline{u}_1 + \beta\underline{u}_2) = \phi[\alpha(\mathcal{L}(\underline{u}_1) + \underline{c}) + \beta(\mathcal{L}(\underline{u}_2) + \underline{c})]$$

$$\leq \alpha\phi[\mathcal{L}(\underline{u}_1) + \underline{c}] + \beta\phi[\mathcal{L}(\underline{u}_2) + \underline{c}]$$

by the convexity of ϕ. This gives

$$(\phi o \psi)(\alpha\underline{u}_1 + \beta\underline{u}_2) \leq \alpha(\phi o \psi)(\underline{u}_1) + \beta(\phi o \psi)(\underline{u}_2)$$

Q.E.D. Lemma 3.3.2

IV. <u>Lemma 3.3.3</u>: Let $\phi_1(\cdot), \phi_2(\cdot), \ldots, \phi_n(\cdot)$ be convex. Then $\phi(\underline{u}) = \phi_1(\underline{u}) + \phi_2(\underline{u}) + \ldots + \phi_n(\underline{u})$ is convex

<u>Proof</u>:

$$\phi(\alpha\underline{u}_1 + \beta\underline{u}_2) = \sum_{i=1}^{n} \phi_i(\alpha\underline{u}_1 + \beta\underline{u}_2)$$

$$\leq \sum \alpha\phi_i(\underline{u}_1) + \beta\phi_i(\underline{u}_2)$$

$$= \alpha\phi(\underline{u}_1) + \beta\phi(\underline{u}_2)$$

Since $\phi_i(\alpha\underline{u}_1 + \beta\underline{u}_2) \leq \alpha\phi_i(\underline{u}_1) + \beta\phi_i(\underline{u}_2)$ for each i

Q.E.D. Lemma 3.3.3

V. <u>Lemma 3.3.4</u>: Consider the deterministic system $\underline{x}_{t+1} = \underline{A}_t\underline{x}_t + \underline{B}_t\underline{u}_t + \underline{\xi}_t$; $t = 0, 1, \ldots, T$, $\{\underline{x}_0, \underline{\xi}_0, \underline{\xi}_1, \ldots, \underline{\xi}_{T-1}\}$ fixed. The map $\{u_k\} \rightarrow \{\underline{x}_k\}$ is affine from $R^{m \cdot T}$ to $R^{n \cdot T}$.

Proof: Immediate from the solution formula

$$\underline{x}_t = \underline{\Phi}_{t,0}\underline{x}_0 + \sum_{s=0}^{t-1} \underline{\Phi}_{t,s+1}(\underline{B}_s\underline{u}_s + \underline{\xi}_s)$$

where $\underline{\Phi}_{t,s} = \underline{A}_t \cdot \underline{A}_{t-1} \cdot \ldots \cdot \underline{A}_s$

Q.E.D. Lemma 3.3.4

Proof of Main Theorem 3.3.2: By definition, $\mathcal{J}_t(\underline{x}_t)$ is the minimal cost-to-go using stochastic dynamic programming. Let $\{\underline{u}_k^*(\cdot)\}_{k=t}^{T-1}$ represent the optimal control laws obtained in that manner. Then

$$\mathcal{J}_t(\underline{x}_t) \triangleq E\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{u}_k^*(\underline{x}_k)) | \underline{x}_t\} \tag{3.3.24}$$

$$\leq E\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{u}_k(\underline{x}_k)) | \underline{x}_t\} \tag{3.3.25}$$

for any other sequence of feedback controls $\{\underline{u}_k(\cdot)\}$ by the optimality of $\{\underline{u}_k^*(\cdot)\}$. In particular, for the control laws $\{\bar{\hat{\underline{u}}}_k(\cdot)\}$,

$$\mathcal{J}_t(\underline{x}_t) \leq E\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \bar{\hat{\underline{u}}}_k(\underline{x}_k)) | \underline{x}_t\} \tag{3.3.26}$$

$$\triangleq E\{\hat{I}_t(\underline{x}_t, \underline{\xi}_t, \underline{\xi}_{t+1}, \ldots, \underline{\xi}_{T-1}) | \underline{x}_t\} \tag{3.3.27}$$

$$= \hat{\mathcal{J}}_t(\underline{x}_t) \tag{3.3.28}$$

So $\mathcal{J}_t(\underline{x}_t) \leq \hat{\mathcal{J}}_t(\underline{x}_t)$. The inequality will now be proved in reverse.

To clarify notation, let $E_x\{\cdot\}$ represent expectation with respect to the random variable x. Then

$$\mathcal{J}_t(\underline{x}_t) \triangleq E\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{u}_k^*(\underline{x}_k)) | \underline{x}_t\} \qquad (3.3.29)$$

$$= E_{\{\underline{\xi}_k\}}\{E\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{u}_k^*(\underline{x}_k)) | \underline{x}_t, \underline{\xi}_t, \ldots, \underline{\xi}_{T-1}\} | \underline{x}_t\} \qquad (3.3.30)$$

by properties of conditional expectation (see Loeve, (L.1), p.341). But given $\{\underline{\xi}_k\}_{k=t}^{T-1}$, the innermost quantity in the expectations in (3.3.30) is minimized for all $\{\underline{\xi}_k\}$ if $\{\underline{u}_k\} = \{\hat{\underline{u}}_k(\underline{x}_k, \underline{\xi}_k, \ldots, \underline{\xi}_{T-1})\}$. Thus

$$\mathcal{J}_t(\underline{x}_t) \geq E_{\{\underline{\xi}_k\}}\{E\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \hat{\underline{u}}_k(\underline{x}_k, \underline{\xi}_k, \ldots, \underline{\xi}_{T-1})) | \underline{x}_t, \{\underline{\xi}_k\}\} | \underline{x}_t\} \qquad (3.3.31)$$

Now the $\hat{\underline{u}}_k(\cdot)$ functions are well-defined as the feedback solution of a certain deterministic optimal control problem. If $\{\underline{\xi}_k\}$ are taken to be random variables, then given $\underline{x}_k$, $\hat{\underline{u}}_k(\underline{x}_k, \underline{\xi}_k, \ldots, \underline{\xi}_{T-1})$ will also be random variables if the $\hat{\underline{u}}_k(\cdot)$ are Borel functions. It has been assumed that they are. If the $\{\hat{\underline{u}}_k\}$ are now considered random variables in their own right for $k \geq t$, (3.3.31) may be written

$$\mathcal{J}_t(\underline{x}_t) \geq E_{\{\underline{\xi}_k\}}\{E_{\{\hat{\underline{u}}_k\}}\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{u}_k) | \underline{x}_t, \{\underline{\xi}_k\}\} | \underline{x}_t\} \quad (3.3.32)$$

Since the L_k *functions* are convex *and the* system is *linear, the* map $\{\underline{\hat{u}}_k\} \rightarrow L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{\hat{u}}_k)$ is convex, by Lemmas 3.3.2 - 3.3.4. Thus by Lemma 3.3.1

$$\mathcal{J}_t(\underline{x}_t) \geq E_{\{\underline{\xi}_k\}}\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, E\{\underline{\hat{u}}_k | \underline{x}_t\}) | \underline{x}_t\}$$

$$= E\{L_T(\underline{x}_T) + \sum_{k=t}^{T-1} L_k(\underline{x}_k, \underline{\bar{\hat{u}}}_k(\underline{x}_k)) | \underline{x}_t\}$$

$$= \hat{\mathcal{J}}_t(\underline{x}_t) \tag{3.3.33}$$

Thus, $\mathcal{J}_t(\underline{x}_t) \geq \hat{\mathcal{J}}_t(\underline{x}_t)$, and the inequality holds both ways, giving

$$\mathcal{J}_t(\underline{x}_t) = \hat{\mathcal{J}}_t(\underline{x}_t) \tag{3.3.34}$$

Q.E.D

3.4 Summary

In this chapter, circumstances under which it is possible to approach a stochastic control problem by using deterministic techniques were described. The special case of a linear system driven by Gaussian noise was introduced as a basis for the analysis.

The first step in the analysis was to consider the Kalman filter equations, which characterize the propagation of the p.d.f. of the plant state given the measurements, in place of the given system equations. The cost functional was expressed in terms of the conditional mean and conditional covariance of the plant state. The fact that these transform actions do not change the solution of the control problem was expressed in Theorem 3.3.1.

The rest of the analysis revolved around Theorem 3.3.2, the main result of this chapter. This theorem states that the average cost incurred by using the expected value of the control law determined by assuming the noises are fixed is equal to the true stochastic optimal cost. As result of this theorem, one can apply deterministic optimal control techniques in solving a stochastic problem, and one is guaranteed that the control law derived using those techniques will result in performance which is as good on the average as that which the true stochastic optimal control will produce.

CHAPTER IV

STRUCTURE OF DYNAMIC AND MEASUREMENT CONTROL SYSTEMS FOR LINEAR-QUADRATIC-GAUSSIAN PROBLEMS

4.1 General Perspective

As was pointed out in Section 1.4, several authors have considered combined measurement/control optimization problems for the special situation in which the dynamic equation is linear in state and control, the measurement equation is linear in state (although not necessarily in measurement control), the noises are white and described by Gaussian statistics, and the cost functional is quadratic in state and dynamic control. The typical approach is to assume a fixed but arbitrary sequence for $\{\underline{v}_t\}$, apply the well known separation theorem and linear regulator techniques (B.4) to the resulting "ordinary" control problem involving $\{\underline{u}_t\}$, collect the terms in the expected cost which depend on $\{\underline{v}_t\}$, and finally solve for the sequence $\{\underline{v}_t\}$ which minimizes those terms. This approach has not been carefully justified; it is not immediately apparent whether the same solution results if one solves simultaneously for $\{\underline{u}_t\}$ and $\{\underline{v}_t\}$. Meier, Peschon and Dressler (M.2) have a theorem stating that indeed the same answer results regardless of the approach. Their proof, however, while basically correct, does not carefully point out what quantities depend on what controls, so that the exact reasons why the "a-priori $\underline{v}$ technique" is optimal are somewhat obscure. They do not consider at all the point discussed in subsection 2.4.2 of this thesis: that one is defining the neglecting of past costs to be optimal. For these reasons, a complete proof will be given below.

A widely applied technique in problem solving is to break a big, complicated problem up into a sequence of smaller, simpler problems. This is the basic motivation for the dynamic programming technique in which a large, multistep optimization is reduced to a sequence of single step problems. In the context of stochastic control problems, another complexity reduction procedure can be used in certain circumstances by applying the Separation Theorem or Certainty Equivalence Principle (W.1). This principle allows the designer to break the overall stochastic control problem into two pieces: an estimation problem involving the determination of the system state and a control problem which is deterministic.* The designer then plugs the "answer" to the estimation problem, i.e., the expected plant state, into the feedback control law of the deterministic control problem and the overall solution is optimal for the overall system.

It is obvious that it is desirable to be able to apply the technique outlined above to a given problem: the overall hard problem is replaced by two easier problems. Considerable effort has been expended in trying to determine precisely when this separation approach is valid. Witsenhausen (W.1) summarizes these efforts. Basically, the easiest solution results in the linear system-quadratic cost-Guassian noise case because the separation property holds, the estimation problem is more-or-less easy to solve using a Kalman Filter (K.4), and the deterministic problem that results is a linear regulator problem, also relatively easy to solve (B.5). It has been shown that in certain other cases, separation techniques can be used, but the resulting two problems may not be so easy. For example,

---

*Not necessarily the obvious one obtained by simply "throwing" away the noise.

Wonham (W.2) considers linear Gaussian systems with a generally non-quadratic cost and the resulting deterministic control problem is not easy, although estimation remains tractable.

The nomenclature attached to the study of stochastic control problems is somewhat confusing with regard to the question of separation theorems. Under the assumption of white noises, Striebel (S.6) showed that the optimal control depends on the data <u>only</u> through the conditional p. d. f. of the plant state given the data. This holds for any system, linear or nonlinear. This itself constitutes a separation theorem since it breaks the overall stochastic control problem into two parts: 1) an estimation problem (calculate the pdf and 2) a control problem (find the feedback function into which the pdf must be substituted).

What is usually called The Separation Theorem (with capitals) is the particular result for linear-Gaussian-quadratic problems that the optimal stochastic control results from plugging the <u>conditional mean</u> of the plant state (output of the Kalman filter) into the feedback law for the <u>same problem with noises set to zero</u>. In addition to this particularly simple structure for the optimal stochastic control, The Separation Theorem states that the solution to the filtering problem (i. e., the equations for determining the conditional mean and conditional covariance) are completely independent of the state and control wieghting matrices which appear in the cost functional. Conversely, the determination of the control <u>law</u> is completely independent of the means and covariance matrices of the stochastic part of the problem.

In this thesis, it becomes necessary to be more precise about the meaning of separation because there is an additional facet of the problem that might separate out: the determination of the optimal measurement control. The various circumstances that can occur are illustrated in Figure 4.1 through 4.4.

Figure 4.1 illustrates the situation which applies to the linear-Gaussian-quadratic case with no measurement control capability. The calculations required to specify the filter (state estimator) are completely independent of those required to specify the dynamic control law. The calculation of the control law and filter parameters are, in turn, independent of the measured data and may be completed a-priori (off-line). This situation will be referred to as complete two-way separation: complete because the filter and control laws are independent of the data, and two-way because the filter and control law calculations are independent.

Figure 4.2 illustrates the structure of the control system for a linear-Gaussian-quadratic problem with measurement control capability. The dynamic control law calculations are independent of the filter and measurement control law calculations and also of the measured data. The filter parameter calculations are coupled back and forth to the measurement control law calculations, but these are both independent of the measured data: the optimal measurement control time sequence can be calculated a-priori (off-line). The separation can still be called complete because all of the control law and filter parameter calculations may be done a-priori (off line). The dynamic control law is still completely independent of the filter calculations (a two-way separation) and of the measurement control calculations (another two-way separation). The filter calculations and the measurement control calculations are tightly coupled, however.

Figure 4.3 illustrates the situation for the linear-Gaussian-non-quadratic case. Separation is still complete: all filter parameters and control law parameters are calculated a-priori. However, the control law which specifies the optimal measurement control $\underline{v}^*_{t+1}$ now has the state estimate at time t, $\hat{\underline{x}}_{t|t}$, as an argument; it no longer depends only on the covariance matrix. It is thus impossile to calculate the optimal measurement control off-line. In addition, the dynamic control-filter separation is now one-way only: the control law depends on the filter but the filter calculations are independent of the control law. The dynamic and measurement control calculations are still decoupled (two-way separation) except, of course, as the measurements influence the filter and the filter the dynamic control law. This will not be considered a direct coupling.

For comparison, the general case is illustrated in Figure 4.4. Everything must be done on line and no separation occurs.*

---

*This may require clarification. If the control laws take as their arguments the entire conditional p.d.f. of the plant state, the control law may be calculated off-line. If, however, as is often the case, the argument of the control law is constrained to be only the current state estimate $\hat{\underline{x}}_{t|t}$, then the control law must be calculated on-line. See Kushner (K.2).

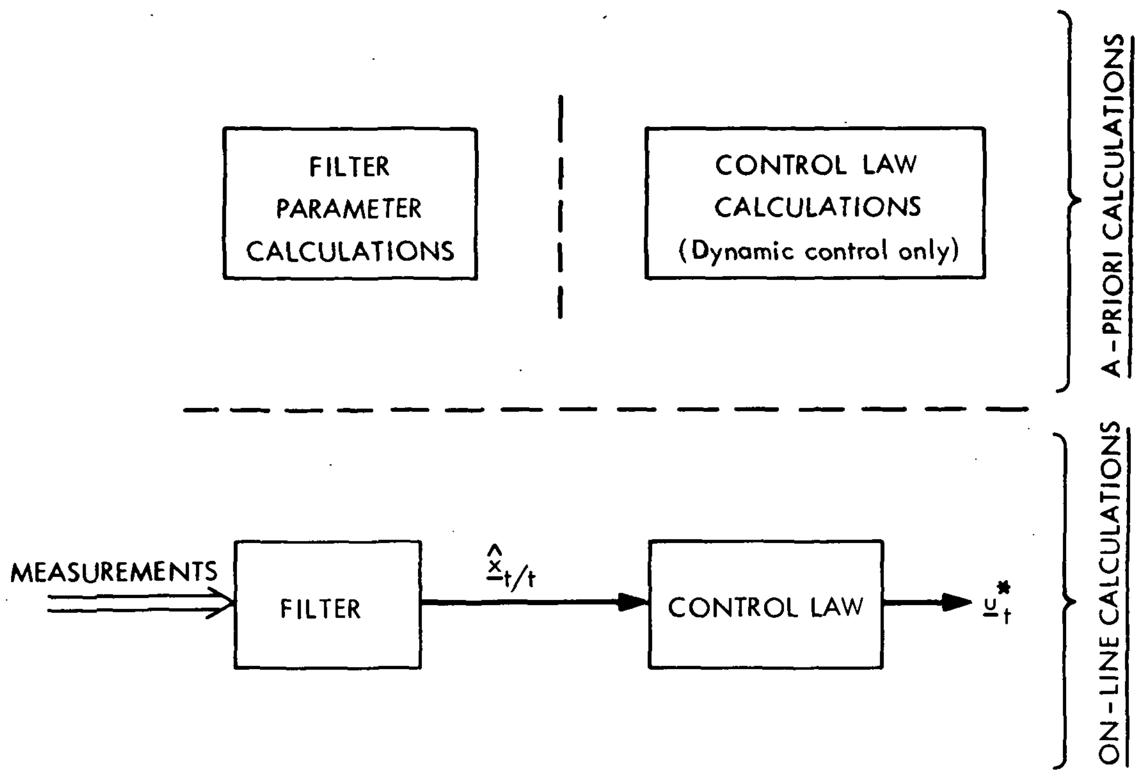

Fig. 4.1 Complete Two-way Separation: Linear-Gaussian-Quadratic with no Measurement Control Capability.

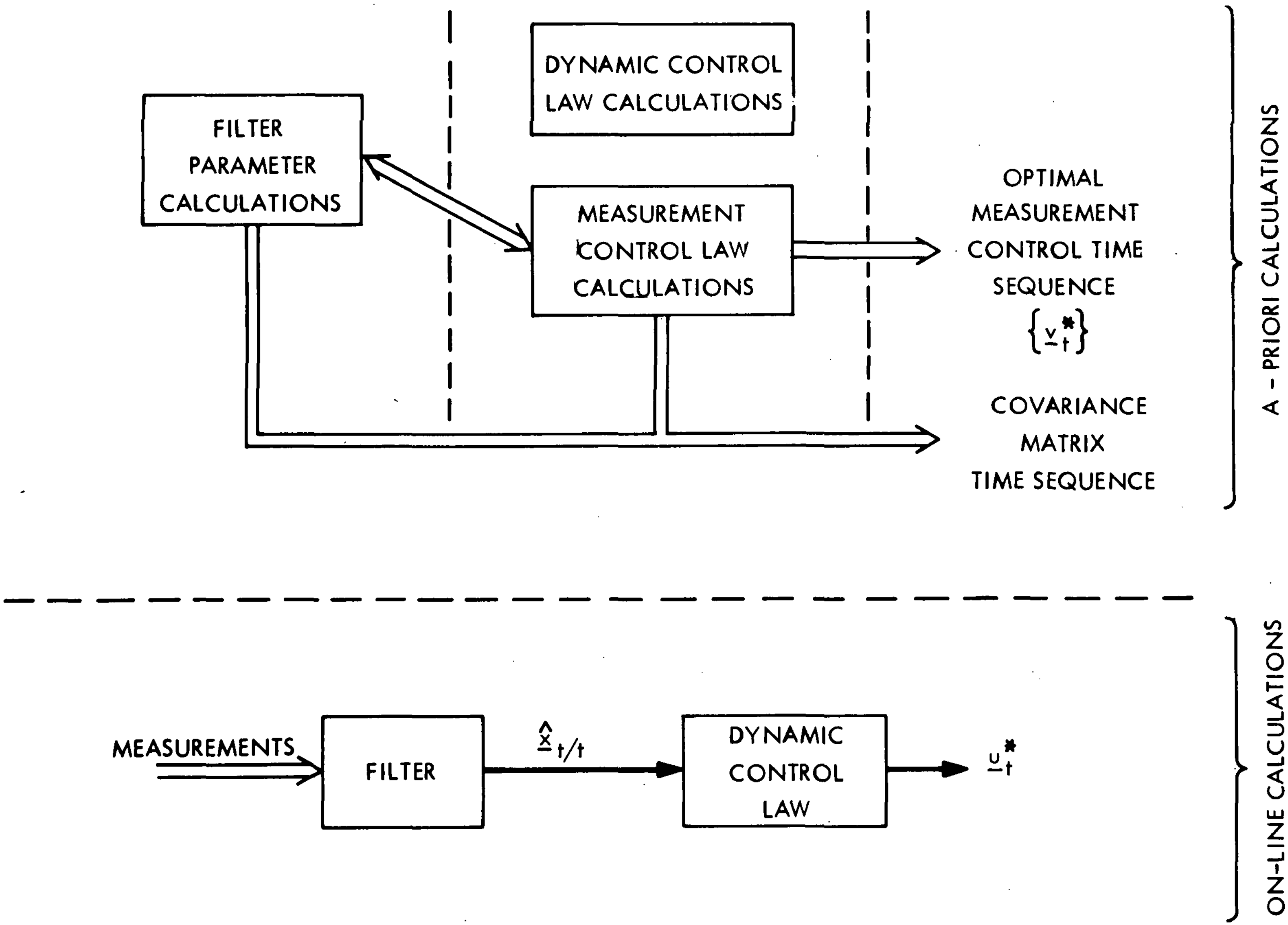

Fig. 4.2 Complete Two-way Dynamic Control-Filter Separation with Complete Two-way Measurement-Dynamic Control Separation: Linear-Gaussian-Quadratic Case.

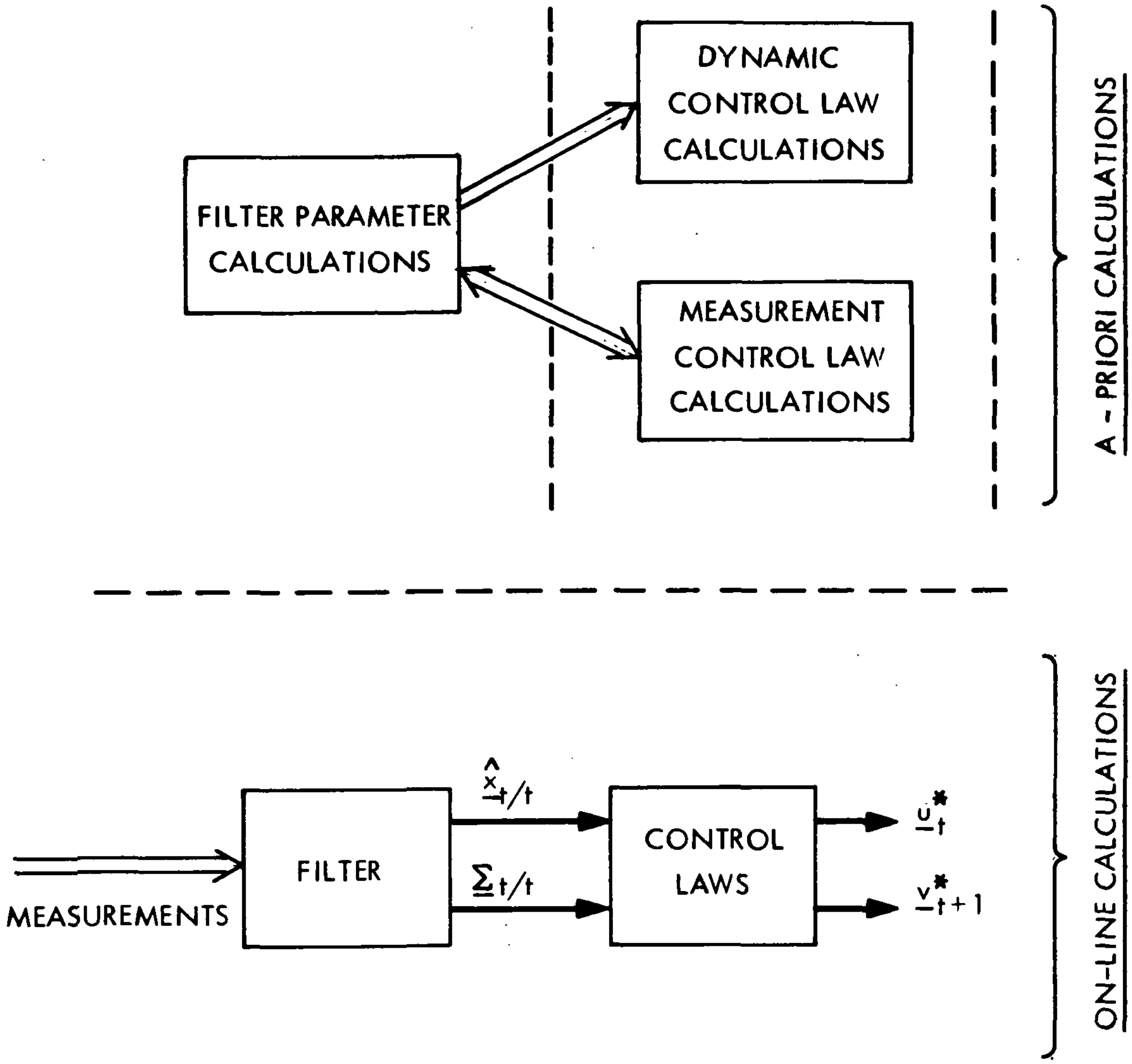

Fig. 4.3 Complete One-way Dynamic Control-Filter Separation with Complete Two-way Measurement-Dynamic Separation: Linear-Gaussian Non-Quadratic Case.

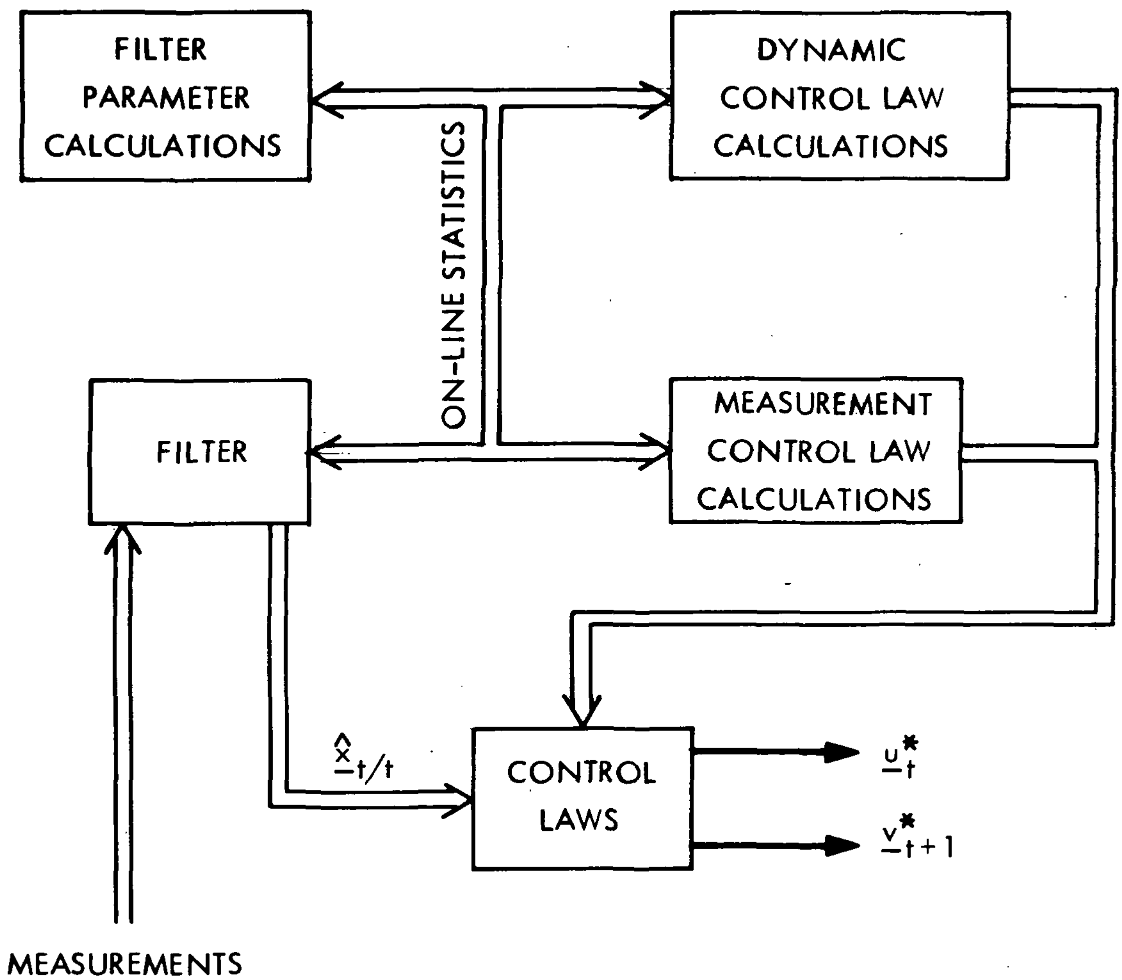

Fig. 4.4 General Case

The present chapter will be devoted to showing that Figure 4.2 indeed describes linear-Gaussian-quadratic systems. Both dynamic programming and deterministic techniques will be used. The next chapter will repeat the analysis for linear-Gaussian-non-quadratic systems to confirm the implications of Figure 4.3.

The equations describing a linear-Gaussian problem with quadratic cost will now be collected here for convenience. Recall from Section 3.2 the equations describing a linear-Gaussian system:

$$\underline{x}_{t+1} = \underline{A}_t \underline{x}_t + \underline{B}_t \underline{u}_t + \underline{\xi}_t \tag{4.1.1}$$

$$\underline{y}_t = \underline{C}_t(\underline{v}_t)\, \underline{x}_t + \underline{D}_t(\underline{v}_t)\, \underline{\theta}_t \tag{4.1.2}$$

The variables have the same meaning as in Section 3.2. The cost functional will be

$$J = \underline{x}_T' \underline{Q}_T \underline{x}_T + \ell_0(\underline{v}_0) + \sum_{t=0}^{T-1} \underline{x}_t' \underline{Q}_t \underline{x}_t + \underline{u}_t' \underline{R}_t \underline{u}_t + \ell_{t+1}(\underline{v}_{t+1}) \tag{4.1.3}$$

where $\underline{Q}_t$ is a symmetric and positive semi-definite matrix and $\underline{R}_t$ is a symmetric and positive definite matrix. The $\ell_t(\cdot)$ functions will be assumed to be scalar positive definite, and convex.

4.2 Results Using Dynamic Programming

Theorem 4.2.1. (Complete, Two-way Separation Theorem)

If the dynamic system is described by equation (4.1.1) with measurement equation (4.1.2) and cost functional (4.1.3), under the assumption that optimal dynamic and measurement control sequences exist, the following properties hold.

A. The optimal measurement control sequence $\{\underline{v}_t^*\}$ may be pre-computed. It is applied open loop and depends only on the system parameters and noise statistics.

B. The optimal dynamic control sequence $\{\underline{u}_t^*\}$ is generated by feedback in real time by $\underline{u}_t^* = -\underline{K}_t \hat{\underline{x}}_{t|t}$ where $\underline{K}_t$ is a pre-computable gain matrix independent of $\{\underline{v}_t^*\}$, and $\hat{\underline{x}}_{t|t}$ is the best linear estimate of $\underline{x}_t$, generated by a Kalman Filter which is "matched" to the measurement profile $\{\underline{v}_t^*\}$.

C. The gain matrix $\underline{K}_t$ is identical to that obtained in the deterministic problem which results if $\underline{\xi}_t$ is identically zero in (4.1.1) and the measurement aspects of the problem are deleted.

Proof: The method of proof will be dynamic programming, applied as discussed in Chapter II. To make the proof as compact as possible, the explicit dependence of $\underline{C}_t$ and $\underline{D}_t$ on $\underline{v}_t$ will be supressed in the notation: $\underline{C}_t$ and $\underline{D}_t$ will appear instead of $\underline{C}_t(\underline{v}_t)$ and $\underline{D}_t(\underline{v}_t)$. However, one must bear in mind that these quantities are influenced by the measurement control. In addition, the symbol $\tilde{Y}_t$ will be used to denote the entire data set at time t, i.e., let $\tilde{Y}_t = Y_t \cup U_{t-1} \cup V_t$.

Terminal Step: Let the time be t= T. The terminal cost-to-go, analogous to equation (2.3.18) is

$$\gamma_T = E\{\underline{x}_T' \underline{Q}_T \underline{x}_T | \tilde{Y}_T\} \tag{4.2.1}$$

where $\breve{Y}_T$ is the fixed, known data base. There is nothing to optimize, and the value of $\mathcal{J}_T$, to be interpreted as $\mathcal{J}_T(\breve{Y}_T)$ in the next step, may be evaluated as follows:

$$\mathcal{J}_T = E\{\underline{x}'_T \underline{Q}_T \underline{x}_T | \breve{Y}_T\}$$

$$= E\{tr[\underline{Q}_T \underline{x}_T \underline{x}'_T] | \breve{Y}_T\} \quad (4.2.2)$$

$$= tr[\underline{Q}_T E\{\underline{x}_T \underline{x}'_T | \breve{Y}_T\}] \quad (4.2.3)$$

Equation (4.2.3) may be expressed only in terms of the mean vector and covariance matrix of $\underline{x}_t$. When conditioned on $\breve{Y}_T$, let

$$\hat{\underline{x}}_{T|T} = E\{\underline{x}_T | \breve{Y}_T\} \quad (4.2.4)$$

$$\underline{\Sigma}_{T|T} = E\{(\underline{x} - \hat{\underline{x}}_{T|T})(\underline{x} - \hat{\underline{x}}_{T|T})' | \breve{Y}_T\} \quad (4.2.5)$$

be the conditional mean and covariance of $\underline{x}_T$, respectively. Then manipulating (4.2.3) gives

$$\mathcal{J}_T = tr[\underline{Q}_T \cdot E\{\underline{x}_T \underline{x}'_T | \breve{Y}_T\}]$$

$$= tr[\underline{Q}_T \cdot E\{(\underline{x}_T - \hat{\underline{x}}_{T|T} + \hat{\underline{x}}_{T|T})(\underline{x}_T - \hat{\underline{x}}_{T|T} + \hat{\underline{x}}_{T|T})' | \breve{Y}_T\}] \quad (4.2.6)$$

$$= tr[\underline{Q}_T \cdot E\{(\underline{x}_T - \hat{\underline{x}}_{T|T})(\underline{x} - \hat{\underline{x}}_{T|T})' + 2(\underline{x}_T - \hat{\underline{x}}_{T|T})\hat{\underline{x}}_{T|T} + \hat{\underline{x}}_{T|T}\hat{\underline{x}}'_{T|T} | \breve{Y}_T\}$$

$$(4.2.7)$$

$$= tr[\underline{Q}_T \cdot [\underline{\Sigma}_{T|T} + \hat{\underline{x}}_{T|T}\hat{\underline{x}}'_{T|T}]] \quad (4.2.8)$$

$$= \hat{\underline{x}}'_{T|T} \underline{Q}_T \hat{\underline{x}}_{T|T} + tr[\underline{Q}_T \underline{\Sigma}_{T|T}] \quad (4.2.9)$$

Arbitrary Step t: In order to analyze an arbitrary step at time t, an expression for $\vartheta_{t+1} = \min E\{\mathcal{J}_{t+1} | \tilde{Y}_{t+1}\}$ must be available to plug into (2.3.22). Using some foresight, a suitable form for ϑ_{t+1} will be suggested and then proved correct by induction.

In (4.2.9), $\vartheta_T(\tilde{Y}_T)$ depends on $\tilde{Y}_T$ through the sufficient statistics $\hat{\underline{x}}_{T|T}$ and $\underline{\Sigma}_{T|T}$. ϑ_T is quadratic in the conditional mean. Since the system is linear and the underlying cost is quadratic, at least in state and dynamic control, this seems a reasonable form in general. It will, of course, be necessary to include a term to represent the contribution of ℓ_t terms. It will be shown that such a term depends on $\tilde{Y}_t$ only through $\underline{\Sigma}_{t|t}$. Then, as a start, assume that*

$$\vartheta_t(\tilde{Y}_t) = \hat{\underline{x}}'_{t|t}\underline{M}_t\hat{\underline{x}}_{t|t} + \underline{m}'_t\hat{\underline{x}}_{t|t} + h_t + I_t(\underline{\Sigma}_{t|t}) \tag{4.2.10}$$

where $\underline{M}_t$ is symmetric. The terminal values are

$$\underline{M}_T = \underline{Q}_T \tag{4.2.11}$$

$$\underline{m}_T = \underline{0} \tag{4.2.12}$$

$$I_T = h_T = 0 \tag{4.2.13}$$

The data $\tilde{Y}_t$ are available. The controls $\underline{u}_t$ and $\underline{v}_{t+1}$ must be chosen to minimize

$$\mathcal{J}_t = E\{\underline{x}'_t\underline{Q}_t\underline{x}_t + \underline{u}'_t\underline{R}_t\underline{u}_t + \ell_{t+1}(\underline{v}_{t+1}) + \vartheta_{t+1}(\tilde{Y}_{t+1}) | \tilde{Y}_t\} \tag{4.2.14}$$

---

* This is slightly more general than the form assumed by Meier, et al. (M.2)

or using (4.2.10)

$$\mathcal{J}_t = E\{\underline{x}_t' \underline{Q}_t \underline{x}_t + \underline{u}_t' \underline{R}_t \underline{u}_t + \ell_{t+1}(\underline{v}_{t+1}) + \hat{\underline{x}}_{t+1|t+1}' \underline{M}_{t+1} \hat{\underline{x}}_{t+1|t+1}$$

$$+ \hat{\underline{x}}_{t+1|t+1}' \underline{m}_{t+1} + h_{t+1} + I_{t+1}(\underline{\Sigma}_{t+1|t+1}) | \tilde{Y}_t\} \tag{4.2.15}$$

To evaluate (4.2.15), notice that assuming $\tilde{Y}_t$ is given, $\underline{x}_t$ is a Gaussian random variable with mean $\hat{\underline{x}}_{t|t}$ and covariance $\underline{\Sigma}_{t|}$ just as at Step T. Since $\underline{x}_{t+1}$ is linearly related to $\underline{x}_t$ and a linear measurement is made, there is no loss in generality in assuming that $\hat{\underline{x}}_{t+1|t+1}$ is generated from $\hat{\underline{x}}_{t|t}$ by using a Kalman Filter. At Step t, $\hat{\underline{x}}_{t+1|t+1}$ is a random variable, since $\underline{y}_{t+1}$ is involved in calculating $\hat{\underline{x}}_{t+1|t+1}$. The form, however, is known. It then follows at once that the propagation equations are:

$$\underline{x}_{t+1} = \underline{A}_t \underline{x}_t + \underline{B}_t \underline{u}_t + \underline{\xi}_t \tag{4.2.16}$$

$$\underline{y}_{t+1} = \underline{C}_{t+1} \underline{x}_{t+1} + \underline{D}_{t+1} \underline{\theta}_{t+1} \tag{4.2.17}$$

$$\hat{\underline{x}}_{t+1|t+1} = \underline{A}_t \hat{\underline{x}}_{t|t} + \underline{B}_t \underline{u}_t + \underline{S}_{t+1} [\underline{y}_{t+1} - \underline{C}_{t+1} (\underline{A}_t \hat{\underline{x}}_{t|} + \underline{B}_t \underline{u}_t)] \tag{4.2.18}$$

Kalman Gain = $\underline{S}_{t+1}$ =

$$(\underline{A}_t \underline{\Sigma}_{t|t} \underline{A}_t' + \underline{\Xi}_t) \underline{C}_{t+1}' \left(\underline{C}_{t+1} [\underline{A}_t \underline{\Sigma}_{t|t} \underline{A}_t' + \underline{\Xi}_t] \underline{C}_{t+1}' + \underline{D}_{t+1} \underline{\Theta}_{t+1} \underline{D}_{t+1}' \right)^{-1} \tag{4.2.19}$$

$$\underline{\Sigma}_{t+1|t+1} = (\underline{I} - \underline{S}_{t+1} \underline{C}_{t+1}) (\underline{A}_t \underline{\Sigma}_{t|t} \underline{A}_t' + \underline{\Xi}_t) \tag{4.2.20}$$

To reduce notational confusion, let

$$\hat{\underline{x}}_{t+1|t} = \underline{A}_t \hat{\underline{x}}_{t|t} + \underline{B}_t \underline{u}_t = E[\underline{x}_{t+1} | \tilde{Y}_t\} \tag{4.2.21}$$

$$\underline{\Sigma}_{t+1|t} = \underline{A}_t \underline{\Sigma}_{t|t} \underline{A}_t' + \underline{\Xi}_t = \text{cov}(\underline{x}_{t+1}, \underline{x}_{t+1} | \tilde{Y}_t) \tag{4.2.22}$$

The filter equations become

$$\hat{x}_{t+1|t+1} = \hat{x}_{t+1|t} + S_{t+1}[y_{t+1} - C_{t+1}\hat{x}_{t+1|t}] \qquad (4.2.23)$$

$$S_{t+1} = \Sigma_{t+1|t}C'_{t+1}[C_{t+1}\Sigma_{t+1|t}C'_{t+1} + D_{t+1}\Theta_{t+1}D'_{t+1}]^{-1} \qquad (4.2.24)$$

$$\Sigma_{t+1|t+1} = [I - S_{t+1}C_{t+1}]\Sigma_{t+1|t} \qquad (4.2.25)$$

The expression for γ_t becomes

$$\gamma_t = E\{x'_t Q_t x_t + u'_t R_t u_t + \ell_{t+1}(v_{t+1}) + I_{t+1}(\Sigma_{t+1|t+1}) + h_{t+1}$$

$$+ [\hat{x}_{t+1|t} + S_{t+1}(y_{t+1} - C_{t+1}\hat{x}_{t+1|t})]'M_{t+1}[\hat{x}_{t+1|t} + S_{t+1}(y_{t+1} - C_{t+1}\hat{x}_{t+1|t})]$$

$$+ M'_{t+1}[\hat{x}_{t+1|t} + S_{t+1}(y_{t+1} - C_{t+1}\hat{x}_{t+1|t})] | \breve{Y}_t\} \qquad (4.2.26)$$

Each term in (4.2.26) will now be considered in turn

$$E\{x'_t Q_t x_t | Y_t\} = \mathrm{tr}\left[Q_t \cdot E\{x_t x'_t | \breve{Y}_t\}\right] \qquad (4.2.27)$$

$$= \mathrm{tr}\left[Q_t[\Sigma_{t|t} + \hat{x}_{t|t}\hat{x}'_{t|t}]\right] \qquad (4.2.28)$$

$$= \hat{x}'_{t|t} Q_t \hat{x}_{t|} + \mathrm{tr}\left[Q_t \Sigma_{t|t}\right] \qquad (4.2.29)$$

The terms $u'_t R_t u_t$ and $\ell_{t+1}(v_{t+1})$ come out of the expectation unchanged since u_t and v_{t+1} are taken to be deterministic given $\breve{Y}_t$. The next term in (4.2.26) is

$$E\{\hat{\underline{x}}'_{t+1|t+1}\underline{M}_{t+1}\hat{\underline{x}}_{t+1|t+1}|\breve{Y}_t\}$$

$$= E\{[\hat{\underline{x}}_{t+1|t} + \underline{S}_{t+1}(\underline{D}_{t+1}\underline{\theta}_{t+1} + \underline{C}_{t+1}(\underline{x}_{t+1} - \hat{\underline{x}}_{t+1|t}))]'\ \underline{M}_{t+1}$$

$$[\hat{\underline{x}}_{t+1|t} + \underline{S}_{t+1}(\underline{D}_{t+1}\underline{\theta}_{t+1} + \underline{C}_{t+1}(\underline{x}_{t+1} - \hat{\underline{x}}_{t+1|t}))]|\breve{Y}_t\} \quad (4.2.30)$$

$$= E\{\hat{\underline{x}}_{t+1|t}\underline{M}_{t+1}\hat{\underline{x}}_{t+1|t} + \underline{\theta}'_{t+1}\underline{D}'_{t+1}\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{D}_{t+1}\underline{\theta}_{t+1}$$

$$+ (\underline{x}_{t+1} - \hat{\underline{x}}_{t+1|t})'\underline{C}'_{t+1}\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{C}_{t+1}(\underline{x}_{t+1} - \hat{\underline{x}}_{t+1|t})$$

$$+ 2\hat{\underline{x}}'_{t+1|t}\underline{M}_{t+1}\underline{S}_{t+1}\underline{C}_{t+1}(\underline{x}_{t+1} - \hat{\underline{x}}_{t+1|t})|\breve{Y}_t\} \quad (4.2.31)$$

where the cross terms in $\underline{\theta}_{t+1}$ are zero by the whiteness and zero mean of the $\{\underline{\theta}_t\}$ process. The last term in (4.2.31) is zero since $\hat{\underline{x}}_{t+1|t}$ is de-deterministic given $\breve{Y}_t$ and it is an unbiased estimate. Continuing from (4.2.31) then gives

$$E\{\hat{\underline{x}}_{t+1|t+1}\underline{M}_{t+1}\hat{\underline{x}}_{t+1|t+1}|\breve{Y}_t\} = \hat{\underline{x}}'_{t+1|t}\underline{M}_{t+1}\hat{\underline{x}}_{t+1|t}$$

$$+ \operatorname{tr}\left[\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1}\right]$$

$$+ \operatorname{tr}\left[\underline{C}'_{t+1}\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\right] \quad (4.2.32)$$

$$= [\underline{A}_t\hat{\underline{x}}_{t|} + \underline{B}_t\underline{u}_t]'\underline{M}_{t+1}[\underline{A}_t\hat{\underline{x}}_{t|} + \underline{B}_t\underline{u}_t] \quad (4.2.33)$$

$$+ \operatorname{tr}\left[\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1} + \underline{C}'_{t+1}\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\right]$$

$$= \hat{\underline{x}}'_{t|t}\underline{A}'_t\underline{M}_{t+1}\underline{A}_t\hat{\underline{x}}_{t|t} + \underline{u}'_t\underline{B}'_t\underline{M}_{t+1}\underline{B}_t\underline{u}_t + 2\hat{\underline{x}}'_{t|t}\underline{A}'_t\underline{M}_{t+1}\underline{B}_t\underline{u}_t$$

$$+ \operatorname{tr}\left[\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1} + \underline{C}'_{t+1}\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{C}_{t+1}\underline{\Sigma}_{t+1}\right]$$

$$(4.2.34)$$

The next term in (4.2.26) is

$$E\{\underline{m}'_{t+1}[\hat{\underline{x}}_{t+1}|_t + \underline{S}_{t+1}(\underline{y}_{t+1} - \underline{C}_{t+1}\hat{\underline{x}}_{t+1}|_t)]|\underset{\sim}{Y}_t\}$$

$$= \underline{m}'_{t+1}\hat{\underline{x}}_{t+1}|_t \tag{4.2.35}$$

$$= \underline{m}'_{t+1}[\underline{A}_t\hat{\underline{x}}_t|_t + \underline{B}_t\underline{u}_t] \tag{4.2.36}$$

Collecting all the terms in the expansion of $\mathcal{J}_t$ now gives

$$\begin{aligned}\mathcal{J}_t = {} & \hat{\underline{x}}'_t|_t\underline{Q}_t\hat{\underline{x}}_t|_t + \mathrm{tr}\underline{Q}_t\underline{\Sigma}_t|_t + \underline{u}'_t\underline{R}_t\underline{u}_t + \ell_{t+1}(\underline{v}_{t+1}) \\ & + \hat{\underline{x}}'_t|_t\underline{A}'_t\underline{M}_{t+1}\underline{A}_t\hat{\underline{x}}_t|_t + \underline{u}'_t\underline{B}'_t\underline{M}_{t+1}\underline{B}_t\underline{u}_t + 2\hat{\underline{x}}'_t|_t\underline{A}'_t\underline{M}_{t+1}\underline{B}_t\underline{u}_t \\ & + \mathrm{tr}[\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1} + \underline{C}'_{t+1}\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{C}_{t+1}\underline{\Sigma}_{t+1}|_t] \\ & + h_{t+1} + I_{t+1}[(\underline{I} - \underline{S}_{t+1}\underline{C}_{t+1})\underline{\Sigma}_{t+1}|_t] + \underline{m}'_{t+1}\underline{A}_t\hat{\underline{x}}_t|_t + \underline{m}'_{t+1}\underline{B}_t\underline{u}_t\end{aligned} \tag{4.2.37}$$

Reference to the propagation equations (4.2.16) - (4.2.25) shows that the only terms in (4.2.37) which depend on $\underline{u}_t$ are those which display $\underline{u}_t$ explicitly. In particular, $\underline{S}_{t+1}$ and $\underline{\Sigma}_{t+1}|_t$ are independent of $\underline{u}_t$ although $\underline{v}_{t+1}$ influences these quantities. Thus to minimize $\mathcal{J}_t$ with respect to $\underline{u}_t$, one can collect terms in $\underline{u}_t$ and obtain

$$\mathcal{J}_{t\,(\underline{u}_t \text{ terms})} = \underline{u}'_t[\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t]\underline{u}_t + \underline{m}'_{t+1}\underline{B}_t\underline{u}_t + 2\underline{x}'_t|_t\underline{A}'_t\underline{M}_{t+1}\underline{B}_t\underline{u}_t \tag{4.2.38}$$

which is minimized by choosing

$$\underline{u}^*_t = -[\underline{R}_t + \underline{B}'_t + \underline{M}_{t+1}\underline{B}_t]^{-1}[\underline{B}'_t\underline{M}_{t+1}\underline{A}_t\underline{x}_t|_t + \tfrac{1}{2}\underline{B}'_t\underline{m}_{t+1}] \tag{4.2.39}$$

Notice that $\underline{v}_{t+1}$ does not influence the above equation in any way, and the simultaneous minimization of $\mathcal{J}_t$ with respect to $\underline{u}_t$ and $\underline{v}_{t+1}$ separates into two parts that may be done in sequence.

Collecting terms in (4.2.37) which involve $\underline{v}_{t+1}$ is somewhat more confusing since $\underline{v}_{t+1}$ influences $\underline{S}_{t+1}$, $\underline{\Sigma}_{t+1|t+1}$, $\underline{C}_{t+1}$ and $\underline{D}_{t+1}$. The result is

$$\mathcal{J}_{t(\underline{v}_{t+1}\text{ terms})} = \ell_{t+1}(\underline{v}_{t+1}) + I_{t+1}[(\underline{I} - \underline{S}_{t+1}\underline{C}_{t+1})\underline{\Sigma}_{t+1|t}]$$
$$+ \operatorname{tr}[\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1}$$
$$+ \underline{C}'_{t+1}\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}] \qquad (4.2.40)$$

Manipulating the trace operation using $\operatorname{tr}[\underline{ABC}] = \operatorname{tr}[\underline{CAB}]$ yields:

$$\mathcal{J}_{t(\underline{v}_{t+1}\text{ terms})} = \ell_{t+1}(\underline{v}_{t+1}) + I_{t+1}[(\underline{I} - \underline{S}_{t+1}\underline{C}_{t+1})\underline{\Sigma}_{t+1|t}]$$
$$+ \operatorname{tr}[(\underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1} + \underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}'_{t+1})\underline{S}'_{t+1}\underline{M}_{t+1}\underline{S}_{t+1}]$$
$$(4.2.41)$$

Substituting the expression for $\underline{S}_{t+1}$ from (4.2.24) into (4.2.21) gives

$$\mathcal{J}_{t(\underline{v}_{t+1}\text{ terms})} = \ell_{t+1}(\underline{v}_{t+1}) + I_{t+1}\left[(\underline{I} - \underline{S}_{t+1}\underline{C}_{t+1})\underline{\Sigma}_{t+1|t}\right]$$
$$+ \operatorname{tr}\Big[(\underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1} + \underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}'_{t+1})(\underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1}$$
$$+ \underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}'_{t+1})^{-1}\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{M}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}'_{t+1}$$
$$(\underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1} + \underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}'_{t+1})^{-1}\Big] \qquad (4.2.42)$$

which simplifies to

$$\mathcal{J}_{t\,(\underline{v}_{t+1}\text{terms})} = \ell_{t+1}(\underline{v}_{t+1}) + I_{t+1}[(\underline{I} - \underline{S}_{t+1}\underline{C}_{t+1})\,\underline{\Sigma}_{t+1}|_t]$$

$$+ \; \mathrm{tr}[\underline{M}_{t+1}\underline{\Sigma}_{t+1}|_t \underline{C}'_{t+1}(\underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1} + \underline{C}_{t+1}\underline{\Sigma}_{t+1}|_t \underline{C}'_{t+1})^{-1}$$

$$\underline{C}_{t+1}\underline{\Sigma}_{t+1}|_t] \qquad (4.2.43)^*$$

In order to minimize this expression with respect to $\underline{v}_{t+1}$, specific forms of ℓ_{t+1} and I_{t+1} must be known. Rather than assuming such forms, the general case will be considered.

In order to complete the inductive proof that (4.2.10) is a correct expression for the optimal cost-to-go from Step t, one must substitute the optimal values of $\underline{u}_t$ and $\underline{v}_{t+1}$ into (4.2.37) and show that the appropriate form results. Since no specific form for ℓ_{t+1} is suggested, let $\underline{v}^*_{t+1}$ represent the general optimal value obtained by minimizing (4.2.43). By inspection of (4.2.43), <u>this value will depend on $\underline{\Sigma}_t|_t$ but will be independent of $\hat{\underline{x}}_t|_t$</u>. Then, substituting $\underline{v}^*_{t+1}$ and $\underline{u}^*_t$ (as given in (4.2.39) into (4.2.37), one obtains

$$\min \mathcal{J}_t = \mathcal{V}_t = \hat{\underline{x}}'_t|_t[\underline{Q}_t + \underline{A}'_t\underline{M}_{t+1}\underline{A}_t]\hat{\underline{x}}_t|_t + \mathrm{tr}\,\underline{Q}_t\underline{\Sigma}_t|_t + \ell_{t+1}(\underline{v}^*_{t+1})$$

$$+ \; \underline{u}^{*\prime}_t[\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t]\,\underline{u}^*_t + 2\hat{\underline{x}}'_t|_t\underline{A}_t\underline{M}_{t+1}\underline{B}_t\underline{u}^*_t$$

$$+ \; \mathrm{tr}[\underline{S}'^*_{t+1}\underline{M}_{t+1}\underline{S}^*_{t+1}\underline{D}^*_{t+1}\underline{\Theta}_{t+1}\underline{D}'^*_{t+1}] + h_{t+1} + I^*_{t+1}$$

(Cont'd. on next page)

---

*The $\underline{S}_{t+1}$ factor is intentionally left in the argument of I_{t+1}. It would do no good to substitute in $\underline{S}_{t+1}$ since the precise form of I_{t+1} is not known at this point.

$$+ \quad \mathrm{tr}[\underline{C}'^{*}_{t+1}\underline{S}'^{*}_{t+1}\underline{M}_{t+1}\underline{S}^{*}_{t+1}\underline{C}^{*}_{t+1}\underline{\Sigma}_{t+1|t}]$$

$$+ \quad \underline{m}'_{t+1}\underline{A}_t\hat{\underline{x}}_{t|t} + \underline{m}'_{t+1}\underline{B}_t\underline{u}^{*}_t \tag{4.2.44}$$

The term I^{*}_{t+1} represents the corresponding term in (4.2.39) with v^{*}_{t+1} substituted in; similarly, $\underline{S}^{*}_{t+1}$ is (4.2.24) with substituted in. The matrices C^{*}_{t+1} and $\underline{D}^{*}_{t+1}$ represent $C_{t+1}(v^{*}_{t+1})$ and $D_{t+1}(v^{*}_{t+1})$. The expression given in (4.2.44) may be reduced to the following, after a considerable amount of algebra:

$$\mathcal{L}_t = \hat{\underline{x}}'_{t|t}[\underline{Q}_t + \underline{A}'_t\underline{M}_{t+1}\underline{A}_t - \underline{A}'_t\underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t\underline{M}_{t+1}\underline{A}_t]\hat{\underline{x}}_{t|t}$$

$$- \frac{3}{2}\underline{m}'_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t\underline{M}_{t+1}\underline{A}_t\hat{\underline{x}}_{t|t}$$

$$- \frac{1}{4}\underline{m}'_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t\underline{m}_{t+1} + \text{terms in } \underline{\Sigma}_{t|t} + h_{t+1} \tag{4.2.45}$$

This is of the form

$$\mathcal{L}_t = \hat{\underline{x}}'_{t|t}\underline{M}_t\hat{\underline{x}}_{t|t} + \underline{m}'_t\hat{\underline{x}}_{t|t} + I_t(\underline{\Sigma}_{t|t}) + h_t \tag{4.2.46}$$

where

$$\underline{M}_t = \underline{Q}_t + \underline{A}'_t[\underline{M}_{t+1} - \underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t\underline{M}_{t+1}]\underline{A}_t \tag{4.2.47}$$

$$\underline{m}'_t = -\frac{3}{2}\underline{m}'_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t\underline{M}_{t+1}\underline{A}_t \tag{4.2.48}$$

$$I_t(\underline{\Sigma}_{t|t}) = \text{remaining terms depending on } \underline{\Sigma}_{t|t} \tag{4.2.49}$$

$$h_t = \text{remaining constant terms} \tag{4.2.50}$$

Note that h_t cannot be expressed in closed form in terms of h_{t+1} since there are typically constant terms in the group "terms in $\underline{\Sigma}_{t|t}$" in (4.2.45). Similarly, the exact nature of I_t cannot be determined until specific ℓ_t functions are specified. Note, however, that since $\underline{m}_T = \underline{0}$ and $\underline{m}_t$ is linearly related to $\underline{m}_{t+1}$, it follows that $\underline{m}_t = \underline{0}$ for $t = 0, 1, \ldots, T$. Thus the optimal dynamic control can be expressed

$$\underline{u}_t^* = - \underline{K}_t \underline{x}_{t|t} \tag{4.2.51}$$

where $\underline{K}_t = (\underline{R}_t + \underline{B}_t' \underline{M}_{t+1} \underline{B}_t)^{-1} \underline{B}_t' \underline{M}_{t+1} \underline{A}_t$. This completes the induction and proves the theorem, because inspection of equations (4.2.51), (4.2.43) and (4.2.50) shows that the optimization of $\underline{v}_{t+1}$ is independent of $\underline{u}_t$ and is influenced only by a-priori knowledge, while the gain matrix used to obtain $\underline{u}_t^*$ from $\underline{x}_{t|t}$ is identical to that which appears in the deterministic case.

Q.E.D.

The equations representing the "answers" in Theorem 4.2.1 are hidden in among others used in the derivation. To facilitate later discussion, the actual solution of the measurement/control optimization problem will be summarized as a second theorem:

Theorem 4.2.2: For the measurement/control optimization problem given by equations (4.1.1) through (4.1.3), the optimal controls may be obtained as follows:

First: Determine off-line the sequences of matrices $\{\underline{M}_t\}$ and $\{\underline{K}_t\}$ according to

$$\underline{M}_t = \underline{Q}_t + \underline{A}_t' [\underline{M}_{t+1} - \underline{M}_{t+1} \underline{B}_t (\underline{R}_t + \underline{B}_t' \underline{M}_{t+1} \underline{B}_t)^{-1} \underline{B}_t' \underline{M}_{t+1}] \underline{A}_t ; \; \underline{M}_T = \underline{Q}_T \tag{4.2.52}$$

$$\underline{K}_t = (\underline{R}_t + \underline{B}_t' \underline{M}_{t+1} \underline{B}_t)^{-1} \underline{B}_t' \underline{M}_{t+1} \underline{A}_t \tag{4.2.53}$$

Second: Determine the optimal measurement control sequence $\{\underline{v}_t^*\}$ by solving off-line the following optimal control problem: minimize J_v given by

$$J_v = \text{tr}\ \underline{M}_0 \underline{\Sigma}_0 + \sum_{t=0}^{T-1} \underline{M}_{t+1}(\underline{\Xi}_t + \underline{B}_t \underline{K}_t \underline{\Sigma}_{t|t} \underline{A}_t') + \sum_{t=0}^{T} \ell_t(\underline{v}_t) \tag{4.2.54}$$

subject to

$$\underline{\Sigma}_{t|t-1} = \underline{A}_{t-1} \underline{\Sigma}_{t-1|t-1} \underline{A}_{t-1}' + \underline{\Xi}_{t-1};\ \underline{\Sigma}_{0|-1} = \underline{\Sigma}_0 \tag{4.2.55}$$

$$\underline{\Sigma}_{t|t} = \underline{\Sigma}_{t|t-1} - \underline{S}_t [\underline{C}_t(\underline{v}_t) \underline{\Sigma}_{t|t-1} \underline{C}_t'(\underline{v})_t + \underline{D}_t(\underline{v}_t) \underline{\Theta}_t \underline{D}_t'(\underline{v}_t)]\ \underline{S}_t' \tag{4.2.56}$$

$$\underline{S}_t = \underline{\Sigma}_{t|t-1} \underline{C}_t'(\underline{v}_t)[\underline{C}_t(\underline{v}_t) \underline{\Sigma}_{t|t-1} \underline{C}_t'(\underline{v}_t) + \underline{D}_t(\underline{v}_t) \underline{\Theta}_t \underline{D}_t'(\underline{v}_t)]^{-1} \tag{4.2.57}$$

The first and second steps are completed a-priori.

Third: While the system evolves, compute on-line the optimal control sequence $\{\underline{u}_t^*\}$ by feedback according to

$$\underline{u}_t^* = -\underline{K}_t \hat{\underline{x}}_{t|t} \tag{4.2.58}$$

$$\hat{\underline{x}}_{t|t-1} = \underline{A}_{t-1} \hat{\underline{x}}_{t-1|t-1} + \underline{B}_{t-1} \underline{u}_{t-1}^*;\ \hat{\underline{x}}_{0|-1} = \hat{\underline{x}}_0 \tag{4.2.59}$$

$$\hat{\underline{x}}_{t|t} = \hat{\underline{x}}_{t|t-1} + \underline{S}_t(\underline{y}_t - \underline{C}_t(v_t^*) \hat{\underline{x}}_{t|t-1}) \tag{4.2.60}$$

Proof: The propagation equations for $\underline{M}_t$ and $\underline{K}_t$ are taken from (4.2.11), (4.2.47) and (4.2.51). These are independent of $\{\underline{v}_t\}$ and $\{\underline{u}_t\}$. Consideration of (4.2.43) shows that these define the optimization of $\{\underline{v}_t\}$ independently of that of $\{\underline{u}_t\}$, and reference to any standard work, for example, Bryson and Ho (B.4), shows that (4.2.54) summarizes the cost resulting from the choice of $\{\underline{v}_t\}$. Equations (4.2.55) - (4.2.57) and (4.2.59) - (4.2.60) are the Kalman Filter equations, the use of which was justified in the proof of Theorem 4.2.1.

Q.E.D.

Several of the special cases of measurement equation (4.1.3) which were introduced in Section 3.2 will now be considered as corollaries.

Corollary 4.2.3. (Signal-to-Noise Ratio Control). For the measurement/dynamic control problem specified by equations (4.1.1)-(4.1.3), if the measurement equation takes the form

$$\underline{y}_t = v_t \underline{C}_t \underline{x}_t + \underline{\theta}_t \tag{4.2.61}$$

Then the deterministic control problem that one must solve to determine the optimal measurement control takes the following form:

State Equations:

$$\underline{\Sigma}_{t+1|t} = \underline{A}_t \underline{\Sigma}_{t|t} \underline{A}_t' + \underline{\Xi}_t ; \quad \underline{\Sigma}_{0|-1} = \underline{\Sigma}_0 \tag{4.2.62}$$

$$\underline{\Sigma}_{t+1|t+1} = \underline{\Sigma}_{t+1|t} - \underline{S}_{t+1}\left[v_{t+1}^2 \underline{C}_{t+1} \underline{\Sigma}_{t+1} \underline{C}_{t+1}' + \underline{\Theta}_{t+1}\right] \underline{S}_{t+1}' \tag{4.2.63}$$

$$\underline{S}_{t+1} = v_{t+1} \underline{\Sigma}_{t+1|t} \underline{C}_{t+1}' \left[v_{t+1}^2 \underline{C}_{t+1} \underline{\Sigma}_{t+1|t} \underline{C}_{t+1}' + \underline{\Theta}_{t+1}\right]^{-1} \tag{4.2.64}$$

Cost Equation:

$$J_v = \mathrm{tr}\left\{\underline{M}_0 \underline{\Sigma}_0 + \sum_{t=0}^{T-1} \underline{M}_{t+1}(\underline{\Xi}_t + \underline{B}_t \underline{K}_t \underline{\Sigma}_{t|t} \underline{A}_t')\right\} + \sum_{t=0}^{T} \ell_t(\underline{v}_t) \tag{4.2.65}$$

where $\underline{M}_t$ and $\underline{K}_t$ satisfy (4.2.52) and (4.2.53).

Corollary 4.2.4. (Measure-No Measure Control). For the measurement/dynamic control problem specified by equations (4.1.1) - (4.1.3), if the measurement equation takes the form

$$\underline{y}_t = v_t\left[\underline{C}_t \underline{x}_t + \underline{\theta}_t\right]; \; v_t \epsilon \{0,1\} \tag{4.2.66}$$

Then the deterministic control problem that one must solve to determine the optimal measurement control takes the following form:

State Equations:

$$\Sigma_{t+1|t} = \underline{A}_t \Sigma_{t|t} \underline{A}'_t + \underline{\Xi}_t; \; \Sigma_{0|-1} = \Sigma_0 \tag{4.2.67}$$

$$\begin{aligned}\underline{\Sigma}_{t+1|t+1} &= \underline{\Sigma}_{t+1|t} \\ &\quad - v_{t+1}\left[\underline{S}_{t+1}(\underline{C}_{t+1}\underline{\Sigma}_{t+1}\underline{C}'_{t+1} + \underline{\Theta}_{t+1})\underline{S}'_{t+1}\right]; \; v_{t+1}\epsilon\{0,1\}\end{aligned} \tag{4.2.68}$$

$$\underline{S}_{t+1} = \underline{\Sigma}_{t+1|t}\,\underline{C}'_{t+1}\left[\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}'_{t+1} + \underline{\Theta}_{t+1}\right]^{-1} \tag{4.2.69}$$

Cost Equation:

Same as (4.2.65)

Corollary 4.2.5. (Sensor Selection). For the measurement/dynamic control problem specified by equations (4.1.1) - (4.1.3), if the measurement equation takes the form(where each $v_{i,t}$ is either 0 or 1):

$$\underline{y}_t = \left[\sum_{i=1}^{k} v_{i,t}\underline{C}_{i,t}\right]\underline{x}_t + \underline{\theta}_t; \; \sum_{i=1}^{k} v_{i,t} = 1 \tag{4.2.70}$$

Then the deterministic control problem that one must solve to determine the optimal measurement control takes the following form:

State Equations:

Let $$\tilde{\underline{C}}_t \triangleq \sum_{i=1}^{k} v_{i,t}\underline{C}_{i,t} \qquad (4.2.71)$$

Then

$$\underline{\Sigma}_{t+1|t} = \underline{A}_t \underline{\Sigma}_{t|t} \underline{A}_t' + \underline{\Xi}_t ; \ \underline{\Sigma}_{0|-1} = \underline{\Sigma}_0 \qquad (4.2.72)$$

$$\underline{\Sigma}_{t+1|t+1} = \underline{\Sigma}_{t+1|t} - \underline{S}_{t+1}\left[\tilde{\underline{C}}_{t+1}\underline{\Sigma}_{t+1|t}\tilde{\underline{C}}_{t+1} + \underline{\Theta}_{t+1}\right]\underline{S}'_{t+1} \qquad (4.2.73)$$

$$\underline{S}_{t+1} = \Sigma_{t+1|t}\tilde{\underline{C}}'_{t+1}\left\{\tilde{\underline{C}}_{t+1}\underline{\Sigma}_{t+1|t}\tilde{\underline{C}}_{t+1} + \underline{\Theta}_{t+1}\right\}^{-1} \qquad (4.2.74)$$

Cost Equation:

Same as (4.2.65)

Proofs of Corollaries: All are special cases of Theorem 4.2.2 and are proved by specializing the appropriate equations in that Theorem.

4.3 Results Using the Transformation Approach

In the previous section, a double separation property was shown to hold in the solution of a measurement/control optimization problem for a linear systems influenced by Gaussian noise and penalized by a cost criterion which is quadratic in the plant state and the dynamic control. The technique of stochastic dynamic programming was used in the proof. In this section, the same result will be proved using the discrete-time Minimum Principle (K.5) in a manner justified by the result of Chapter III of this thesis. Thus, this section will serve as a detailed example of the application of the results of Chapter III.

The analysis of this section will proceed as follows:

1. The given system consisting of linear dynamics (equations (4.1.1), (4.1.2)) and quadratic cost (equation (4.1.3)) will be replaced by the Kalman filter equations and transformed cost equation, as described in Chapter III.

2. The innovation process driving the Kalman filter will be assumed deterministic: fixed but arbitrary.

3. The discrete Minimum Principle will be applied to the resulting deterministic optimal control problem and the results will be manipulated to give a feedback solution for the dynamic control: $\underline{u}_t^* = \underline{u}_t^*(\hat{\underline{x}}_{t|t}, \hat{\underline{\xi}}_t, \hat{\underline{\xi}}_{t+1}, \ldots, \hat{\underline{\xi}}_{T-1})$.

4. The expected value of $\underline{u}_t^*$ will then be taken over the sequence $\{\hat{\underline{\xi}}_k\}_{k=t}^{T-1}$ and it will be shown that the <u>same</u> solution is obtained in this manner as was obtained by dynamic programming.

5. Finally, the complete, two-way separation property will be shown to hold.

The reader is assumed to be familiar with the Minimum Principle and its application. Since this section amounts to a re-proof of Theorem 4.2.2, no formal theorem statement will be given here.

4.3.1 Transformation of the System Equations

Let the linear system with Gaussian noise (equations (4.1.1),(4.1.2)) be replaced by the Kalman Filter equations

$$\underline{\hat{x}}_{t+1|t+1} = \underline{A}_t \underline{\hat{x}}_{t|t} + \underline{B}_t \underline{u}_t + \underline{\hat{\xi}}_t \tag{4.3.1}$$

as described in subsection 3.3.1. Note that the initial condition for equation (4.3.1) is $\underline{\hat{x}}_{0|0}$, the initial estimate of the plant state given the measurement at time zero. This is related to the a-priori estimate of the initial state, denoted $\underline{\hat{x}}_0$, by the usual Kalman filter equation (see (1.4.11)) where $\underline{\hat{x}}_0$ is interpreted as $\underline{\hat{x}}_{0|-1}$.

$$\underline{\hat{x}}_{0|0} = \underline{\hat{x}}_0 + \underline{S}_0 (\underline{y}_0 - \underline{C}_0 (\underline{v}_0) \underline{\hat{x}}_0) \tag{4.3.2}$$

where $\underline{S}_0$ is the Kalman gain at time zero (see equation 4.3.7) below). For uniformity of notation it will be convenient to define the term $\underline{S}_0(\underline{y}_0 - \underline{C}_0 (\underline{v}_0) \underline{\hat{x}}_0)$ to be $\underline{\hat{\xi}}_{-1}$, so that equation (4.3.2) may be written

$$\underline{\hat{x}}_{0|0} = \underline{\hat{x}}_0 + \underline{\hat{\xi}}_{-1} \tag{4.3.3}$$

Associated with the plant state estimate equation (4.3.1) is the conditional covariance equation (see (1.4.14) - (1.4.16))

$$\underline{\Sigma}_{t+1|t+1} = \underline{A}_t \underline{\Sigma}_{t|t} \underline{A}'_t + \underline{\Xi}_t$$

$$- (\underline{A}_t \underline{\Sigma}_{t|t} \underline{A}'_t + \underline{\Xi}_t) \underline{C}'_{t+1} \left[\underline{C}_{t+1} (\underline{A}_t \underline{\Sigma}_{t|t} \underline{A}'_t + \underline{\Xi}_t) \underline{C}'_{t+1} + \underline{D}_{t+1} \underline{\Theta}_{t+1} \underline{D}'_{t+1} \right]^{-1}$$

$$\underline{C}_{t+1} (\underline{A}_t \underline{\Sigma}_{t|t} \underline{A}'_t + \underline{\Xi}_t) \qquad (4.3.4)^*$$

Similarly to the case of equation (4.3.2), the initial condition for (4.3.4) is $\underline{\Sigma}_{0|0}$, the conditional covariance given the measurement at t=0. This is related to the a-priori covariance $\underline{\Sigma}_0$ by

$$\underline{\Sigma}_{0|0} = \underline{\Sigma}_0 - \underline{\Sigma}_0 \underline{C}'_0 \left[\underline{C}_0 \underline{\Sigma}_0 \underline{C}'_0 + \underline{D}_0 \underline{\Theta}_0 \underline{D}'_0 \right]^{-1} \underline{C}_0 \underline{\Sigma}_0 \qquad (4.3.5)$$

Note now that the measurement control sequence $\{\underline{v}_t\}$ has an influence on the statistics of the innovations process $\{\hat{\underline{\xi}}_t\}$. From equation (3.3.5),

$$\hat{\underline{\xi}}_t = \underline{S}_{t+1} \left[\underline{y}_{t+1} - \underline{C}_{t+1} (\underline{A}_t \hat{\underline{x}}_{t|t} + \underline{B}_t \underline{u}_t) \right] \qquad (4.3.6)$$

where $\underline{S}_{t+1}$ is the Kalman gain** given by

$$\underline{S}_{t+1} = (\underline{A}_t \underline{\Sigma}_{t|t} \underline{A}'_t + \underline{\Xi}_t)$$

$$\times \underline{C}'_{t+1} \left[\underline{C}_{t+1} (\underline{A}_t \underline{\Sigma}_{t|t} \underline{A}_t + \underline{\Xi}_t) \underline{C}'_{t+1} + \underline{D}_{t+1} \underline{\Theta}_{t+1} \underline{D}'_{t+1} \right]^{-1} \qquad (4.3.7)$$

---

*As in the previous section, the dependence of $\underline{C}_{t+1}$ and $\underline{D}_{t+1}$ on $\underline{v}_{t+1}$ have been supressed in the notation to simplify.

**The correct expression for $\underline{S}_0$ is

$$S_0 = \underline{\Sigma}_0 \underline{C}'_0 \left[\underline{C}_0 \underline{\Sigma}_0 \underline{C}'_0 + \underline{D}_0 \underline{\Theta}_0 \underline{D}'_0 \right]^{-1}$$

It has already been observed in Chapter III that the innovations process $\{\hat{\underline{\xi}}_t\}$ is zero mean and white, with covariance matrix given by

$$\text{cov}(\hat{\underline{\xi}}_t, \hat{\underline{\xi}}_t) = E\{\hat{\underline{\xi}}_t \ \hat{\underline{\xi}}'_t\}$$

$$= \underline{S}_{t+1}\left[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}'_t + \underline{\Xi}_t)\,\underline{C}_{t+1} + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1}\right]\underline{S}_{t+1} \quad (4.3.8)$$

To simplify notation in this section, the $t|t$ subscript denoting a quantity at time t conditioned on data up to time t will be supressed during the analysis and only the subscript t will be used. This should generally cause no confusion, because no "offset" conditioning of the type $t|t-1$ will be needed. One point, however, may not be clear in the simplified notation: which initial mean and which initial covariance are being indicated by a given symbol.

Since $\hat{\underline{x}}_0$ and $\underline{\Sigma}_0$ have already been used to denote the a-priori mean and covariance, respectively, it would be confusing to abbreviate the notation for $\hat{\underline{x}}_{0|0}$ and $\underline{\Sigma}_{0|0}$ given in equations (4.3.2) and (4.3.5). Accordingly the $0|0$ subscript will be retained for the initial values.

The system equations thus become:

$$\hat{\underline{x}}_{t+1} = \underline{A}_t\hat{\underline{x}}_t + \underline{B}_t\underline{u}_t + \hat{\underline{\xi}}_t; \quad \text{initial value} = \hat{\underline{x}}_{0|0} \quad (4.3.9)$$

$$\underline{\Sigma}_{t+1} = \underline{A}_t\underline{\Sigma}_t\underline{A}'_t + \underline{\Xi}_t$$
$$- (\underline{A}_t\underline{\Sigma}_t\underline{A}'_t + \underline{\Xi}_t)\,\underline{C}'_{t+1}\left[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_t\underline{A}'_t + \underline{\Xi}_t)\,\underline{C}'_{t+1} + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1}\right]^{-1}\underline{C}_{t+1}$$
$$(\underline{A}_t\underline{\Sigma}_t\underline{A}'_t + \underline{\Xi}_t); \quad \text{initial value} = \underline{\Sigma}_{0|0} \quad (4.3.10)$$

To apply the discrete Minimum Principle below, it is necessary to have the state equations in the form $\underline{z}_{t+1} - \underline{z}_t = \underline{f}(\underline{z}_t, \underline{u}_t)$ where $\underline{z}_t$ is the state. To obtain this form, $\hat{\underline{x}}_t$ and $\underline{\Sigma}_t$ will be subtracted from both sides of (4.3.9) and (4.3.10) respectively to give the final equations

$$\hat{\underline{x}}_{t+1} - \hat{\underline{x}}_t = (\underline{A}_t - \underline{I})\hat{\underline{x}}_t + \underline{B}_t\underline{u}_t + \hat{\underline{\xi}}_t \tag{4.3.11}$$

$$\begin{aligned}\underline{\Sigma}_{t+1} - \underline{\Sigma}_t = {} & \underline{A}_t\underline{\Sigma}_t\underline{A}_t' - \underline{\Sigma}_t + \underline{\Xi}_t \\ & - (\underline{A}_t\underline{\Sigma}_t\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}'\left[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_t\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}'\right]^{-1} \\ & \bullet\ \underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_t\underline{A}_t' + \underline{\Xi}_t)\end{aligned} \tag{4.3.12}$$

which are of the form

$$\hat{\underline{x}}_{t+1} - \hat{\underline{x}}_t = \underline{f}_t(\hat{\underline{x}}_t, \underline{u}_t, \hat{\underline{\xi}}_t) \tag{4.3.13}$$

$$\underline{\Sigma}_{t+1} - \underline{\Sigma}_t = \underline{F}_t(\underline{\Sigma}_t, \underline{v}_{t+1}) \tag{4.3.14}$$

Let equations (4.3.11) - (4.3.14) define the $\underline{f}$ and $\underline{F}$ functions.

To convert the original cost function to the $\hat{\underline{x}}_t$- $\underline{\Sigma}_t$ formulation, note that if $\underline{x}$ is a random vector with mean $\hat{\underline{x}}$ and covariance $\underline{\Sigma}$, then

$$E\{\underline{x}'\underline{Q}\underline{x}\} = \hat{\underline{x}}'\underline{Q}\hat{\underline{x}} + \mathrm{tr}\left[\underline{Q}\underline{\Sigma}\right] \tag{4.3.15}$$

where tr denotes matrix trace. Thus, according to the methods of Chapter III, the cost to be associated with system (4.3.9) - (4.3.14), which is obtained by transforming (4.1.3) is*

$$\begin{aligned}\hat{J} = \frac{1}{2}\Big\{ & \hat{\underline{x}}_T' \underline{Q}_T \hat{\underline{x}}_T + \mathrm{tr}\left[\underline{Q}_T\underline{\Sigma}_T\right] + \ell_T(\underline{v}_T) \\ & + \sum_{t=0}^{T-1} \hat{\underline{x}}_t'\underline{Q}_t\hat{\underline{x}}_t + \mathrm{tr}\left[\underline{Q}_t\underline{\Sigma}_t\right] + \underline{u}_t'\underline{R}_t\underline{u}_t + \ell_t(\underline{v}_t)\Big\}\end{aligned} \tag{4.3.16}$$

---

*The factor of $\frac{1}{2}$ is for algebraic convenience and does not change the problem in any way.

Remark: For the moment, let the cost $\hat{J}$ be expressed as follows

$$\hat{J} = \hat{J}_u + \hat{J}_v \tag{4.3.17}$$

where

$$\hat{J}_u \triangleq \frac{1}{2}\left\{ \hat{\underline{x}}_T' \, \underline{Q}_T \hat{\underline{x}}_T + \sum_{t=0}^{T-1} \hat{\underline{x}}_t' \, \underline{Q}_t \hat{\underline{x}}_t + \underline{u}_t' \, \underline{R}_t \underline{u}_t \right\} \tag{4.3.18}$$

$$J_v \triangleq \frac{1}{2}\left\{ tr\left[\underline{Q}_T \underline{\Sigma}_T\right] + \ell_T(\underline{v}_T) + \sum_{t=0}^{T-1} tr\left[\underline{Q}_t \underline{\Sigma}_t\right] + \ell_t(\underline{v}_t) \right\} \tag{4.3.19}$$

One might suppose at this point that the separation of the $\{\underline{u}_t\}$ problem from the $\{\underline{v}_t\}$ problem has been established, since for fixed $\{\hat{\underline{\xi}}_t\}$, (4.3.11) and (4.3.18) are independent of $\{\underline{v}_t\}$ and are controlled only by $\{\underline{u}_t\}$, while (4.3.12) and (4.3.19) are independent of $\{\underline{u}_t\}$ and are controlled only by $\{\underline{v}_t\}$. *The situation is not that simple, however.*

Recall Remark 3 associated with Theorem 3.3.2. There were no measurement control considerations in that theorem. The covariance equation (4.3.12) was not even considered in the proof, since it was not needed. The remark pointed out that the measurement control aspects of a problem would enter into the solution when two expected values were taken after the deterministic feedback control law was found: the expected value of the deterministic feedback control law over its $\hat{\underline{\xi}}_t$ arguments and the expected value of the cost resulting from the use of that average control. Since each of these two expected values is taken over the sequence of $\hat{\underline{\xi}}_t$ drives in equation (4.3.11), and since the covariance of $\hat{\underline{\xi}}_t$ is controlled by $\underline{v}_{t+1}$ (see equation (4.3.8)), the true and complete nature of the influence which $\underline{v}_{t+1}$ has on the overall cost will not be exposed until those two expected values are taken. It will be seen that the expected value of $\hat{J}_u$ given in (4.3.18) evaluated when the expected value of the deterministic feedback control law is applied contains terms involving $\underline{\Sigma}_t$. Thus $\hat{J}_v$ given in (4.3.19) is not the entire cost due to $\{\underline{v}_t\}$ and the complete two-way separation has not yet been established.

4.3.2 Application of the Minimum Principle

The analysis of system (4.3.11) - (4.3.14) with cost (4.3.16) will now proceed using Theorem 3.3.2. Accordingly, suppose that the innovations process $\{\hat{\underline{\xi}}_t\}$ is fixed and known. The Minimum Principle will be applied to the deterministic optimization problem defined by equations (4.3.11), (4.3.12), and (4.3.16). Let $\{\underline{p}_t\}$ be a sequence of n-vectors, to be called the mean costates. Let $\{\underline{P}_t\}$ be a sequence of $n \times n$ matrices, to be called the covariance costates.

Define the Hamiltonian Function as follows:

$$\begin{aligned} &H(\hat{\underline{x}}_t, \underline{\Sigma}_t, \underline{p}_{t+1}, \underline{P}_{t+1}, \underline{u}_t, \underline{v}_{t+1}) \\ &\quad = \frac{1}{2}\left[\underline{x}_t' \underline{Q}_t \underline{x}_t + \underline{u}_t' \underline{R}_t \underline{u}_t + \mathrm{tr}\, \underline{Q}_t \underline{\Sigma}_t + \ell_{t+1}(\underline{v}_{t+1})\right] + \underline{p}_{t+1}' \underline{f}_t(\underline{x}_t, \underline{u}_t, \underline{\xi}_t) \\ &\qquad + \mathrm{tr}\, \underline{P}_t' \underline{F}_t(\underline{\Sigma}_t, \underline{v}_{t+1}) \end{aligned} \tag{4.3.20}$$

where $\underline{f}_t$ and $\underline{F}_t$ are defined by (4.3.11) - (4.3.14). Then according to the discrete Minimum Principle if $\{\underline{u}_t^*\}$ is the sequence of dynamic controls and $\{\underline{v}_t^*\}$ is the sequence of measurement controls which together minimize (4.3.16) subject to (4.3.11) and (4.3.12), and if $\{\hat{\underline{x}}_t^*\}$, $\{\underline{p}_t^*\}$, $\{\underline{\Sigma}_t^*\}$, and $\{\underline{P}_t^*\}$ are the resulting optimal trajectories of the states and costates, then the following conditions hold:

I. Canonical Equations:

$$\underline{\hat{x}}^*_{t+1} - \underline{\hat{x}}^*_t = \frac{\partial H}{\partial \underline{p}_{t+1}} = \underline{f}_t(\underline{\hat{x}}^*_t, \underline{u}^*_t, \underline{\hat{\xi}}_t) \qquad (4.3.21)$$

$$\underline{\Sigma}^*_{t+1} - \underline{\Sigma}^*_t = \frac{\partial H}{\partial \underline{P}_{t+1}} = \underline{F}_t(\underline{\Sigma}^*_t, \underline{v}^*_{t+1}) \qquad (4.3.22)$$

$$\underline{p}^*_{t+1} - \underline{p}^*_t = - \frac{\partial H}{\partial \underline{\hat{x}}_t} \qquad (4.3.23)$$

$$\underline{P}^*_{t+1} - \underline{P}^*_t = - \frac{\partial H}{\partial \underline{\Sigma}_t} \qquad (4.3.24)$$

II. Minimization of Hamiltonian

$$H(\underline{\hat{x}}^*_t, \underline{\Sigma}^*_t, \underline{p}^*_{t+1}, \underline{P}_{t+1}, \underline{u}^*_t, \underline{v}^*_{t+1}) \leq H(\underline{\hat{x}}^*_t, \underline{\Sigma}^*_t \underline{p}^*_{t+1}, \underline{P}^*_{t+1}, \underline{u}_t, \underline{v}_{t+1}) \qquad (4.3.25)$$

III. Transversality Condition

$$\underline{p}^*_T = \underline{Q}_T \underline{\hat{x}}^*_T \qquad (4.3.26)$$

$$P^*_T = \underline{Q}_T \qquad (4.3.27)$$

These conditions will now be considered more carefully.

The canonical equations for $\underline{\hat{x}}^*_t$ and $\underline{p}^*_t$ take the following form when the indicated derivatives in (4.3.21) and (4.3.23) are taken:

$$\underline{\hat{x}}^*_{t+1} - \underline{\hat{x}}^*_t = (\underline{A}_t - \underline{I})\, \underline{\hat{x}}^*_t + \underline{B}_t \underline{u}^*_t + \underline{\hat{\xi}}_t \qquad (4.3.28)$$

$$\underline{p}^*_{t+1} - \underline{p}^*_t = - \underline{Q}_t\, \underline{\hat{x}}^*_t - (\underline{A}_t - \underline{I})'\, \underline{p}^*_{t+1} \qquad (4.3.29)$$

The detailed equations for $\underline{\Sigma}^*$ and $\underline{P}^*$ will not be needed, but note that the $\underline{\Sigma}^*$ equation is precisely (4.3.12) with asterisks added, and it involves only $\underline{\Sigma}_t^*$ and $\underline{v}_{t+1}^*$. Similarly the $\underline{P}^*$ equation involves only $\underline{P}_{t+1}^*$, $\underline{\Sigma}_t^*$, and $\underline{v}_{t+1}^*$.

Since H is a differentiable function of $\underline{u}_t$ and $\underline{v}_{t+1}$, the minimization condition means that

$$\frac{\partial H}{\partial \underline{u}_t} = \underline{0} \; ; \quad \frac{\partial H}{\partial \underline{v}_{t+1}} = \underline{0} \tag{4.3.30}$$

along optimal trajectories. Evaluating the $\underline{u}$-derivative gives

$$\underline{0} = \underline{R}_t \underline{u}_t^* + \underline{B}'_t \underline{p}_{t+1}^* \tag{4.3.31}$$

$$\underline{u}_t^* = -\underline{R}_t^{-1} \underline{B}_t \underline{p}_{t+1}^* \tag{4.3.32}$$

Note that the matrix of second partials of H with respect to $\underline{u}$ is $\underline{R}_t$, a positive definite matrix, indicating a minimum. Evaluating the $\underline{v}_{t+1}$ derivative gives

$$\text{Expression involving only } \underline{\Sigma}_t^*, \underline{v}_{t+1}^*, \underline{P}_{t+1}^* = \underline{0} \tag{4.3.33}$$

Since the $\ell_t(\underline{v}_t)$ functions are not specifically given, equation (4.3.33) cannot be given in detail. It s form, however, is the important part of the analysis.

Notice that the two-point boundary value problem that results from the application of the Minimum Principle separates into two subproblems: one involving $\hat{\underline{x}}_t^*$, $\underline{p}_t^*$ and $\underline{u}_t^*$ and the other involving $\underline{\Sigma}_t^*$, $\underline{P}_t^*$, and v_{t+1}^*.

This is because equations (4.3.28), (4.3.29) and (4.3.32) relate <u>only</u> $\underline{\hat{x}}_t^*$, $\underline{p}_t^*$ and $\underline{u}_t^*$ while the canonical equations and minimization for $\underline{\Sigma}_t^*$, $\underline{P}_t^*$ and $\underline{v}_{t+1}^*$, equations (4.3.22), (4.3.24), and (4.3.33), are independent of those quantities. As was remarked in the previous subsection, however, <u>this separation of two-point boundary value problems is not sufficient to prove the complete two-way separation theorem of Section 4.2</u>.

4.3.3 Derivation of Optimal Deterministic Dynamic Control Law

The next step in the analysis is to manipulate the conditions given by the Minimum Principle to obtain a feedback control law which expresses the present optimal dynamic control $\underline{u}_t^*$ in terms of $\hat{\underline{x}}_t^*$ and $\{\hat{\underline{\xi}}_t, \hat{\underline{\xi}}_{t+1}, \ldots, \hat{\underline{\xi}}_{T-1}\}$. To accomplish this, first substitute (4.3.32) back into (4.3.28) and (4.3.29) to give the following two-point boundary value problem for the mean and mean costate:

$$\hat{\underline{x}}_{t+1}^* - \hat{\underline{x}}_t^* = (\underline{A}_t - I)\hat{\underline{x}}_t^* - \underline{B}_t\underline{R}_t^{-1}\underline{B}_t'\underline{p}_{t+1}^* + \hat{\underline{\xi}}_t \tag{4.3.34}$$

$$\underline{p}_{t+1}^* - \underline{p}_t^* = -\underline{Q}_t\hat{\underline{x}}_t^* - (\underline{A}_t - I)'\underline{p}_{t+1}^* \tag{4.3.35}$$

$$\hat{\underline{x}}_0 \text{ given}; \quad \underline{p}_T^* = \underline{Q}_T\hat{\underline{x}}_T^* \tag{4.3.36}$$

This system, together with (4.3.32), specifies the optimal sequence $\{\underline{u}_t^*\}$.

The solution for the optimal dynamic control given in equations (4.3.32) and (4.3.34) - (4.3.36) will now be converted to feedback form using well established techniques. See, for example, Kleinman and Athans (K.5). One assumes that $\underline{p}_t^*$ is related to $\hat{\underline{x}}_t^*$ by an affine transformation:

$$\underline{p}_t^* = \underline{M}_t\hat{\underline{x}}_t^* + \underline{h}_t \tag{4.3.37}$$

The term $\underline{h}_t$ is needed because of the drive $\hat{\underline{\xi}}_t$ in (4.3.34). Using (4.3.34)-(4.3.37) and considerable algebra, one can derive equations for $\underline{M}_t$ and $\underline{h}_t$ in a straightforward manner. The details are given in Appendix A, and the results are:

$$\underline{M}_t = \underline{Q}_t + \underline{A}_t'[\underline{M}_{t+1} - \underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t\underline{M}_{t+1}]\underline{A}_t \tag{4.3.38}$$

$$M_T = Q_T \qquad (4.3.39)$$

$$h_t = (I + M_{t+1} B_t R_t^{-1} B_t')^{-1} h_{t+1} + A_t' [M_{t+1} - M_{t+1} B_t (R_t + B_t' M_{t+1} B_t)^{-1} B_t' M_{t+1}] \hat{\xi}_t \qquad (4.3.40)$$

$$h_T = 0 \qquad (4.3.41)$$

Note at once that the matrix equations (4.3.38) and (4.3.39) are identical to (4.2.52); therefore, the M_t matrices are identical. If one now plugs the results of (4.3.34) - (4.3.41) into the expression for the optimal control (4.3.32), and if some algebra is done (see Appendix A), the result is

$$u_t^* = -(R_t + B_t' M_{t+1} B_t')^{-1} B_t' [M_{t+1} (A_t \hat{x}_t^* + \hat{\xi}_t) + h_{t+1}] \qquad (4.3.42)$$

4.3.4 Calculation of Expectation of Deterministic Optimal Dynamic Dynamic Control

The first expected value mentioned in the Remark in Subsection 4.3.1 will now be taken. Let the innovations sequence $\{\hat{\xi}_t\}$ again be considered random. Using the expression (4.3.42) for the optimal feedback dynamic control as a function of the present state and future innovations, take its expected value to give

$$u_t^* = -(R_t + B_t' M_{t+1} B_t)^{-1} B_t' M_{t+1} A_t \hat{x}_t^* \qquad (4.3.43)$$

$$\triangleq -K_t \hat{x}_t^* \qquad (4.3.44)$$

This result is correct because the only terms in (4.3.42) which depend on $\{\hat{\xi}_t\}$ are $\hat{\xi}_t$ itself and h_{t+1}. Since $\hat{\xi}_t$ is zero mean and enters into (4.3.42) linearly, and since h_{t+1} enters linearly and is linearly related to $\{\hat{\xi}_t$,

$\underline{\hat{\xi}}_{t+1}, \ldots, \underline{\hat{\xi}}_{T-1}\}$ (see equations (4.3.40) and (4.3.41)), both of these quantities simply vanish when the expected value of (4.3.42) is taken.

Note that (4.3.43) and (4.3.44) are identical to (4.2.58) and (4.2.53). This means that the optimal feedback dynamic control obtained by assuming the noises fixed and then taking the average at the end is identical to that obtained by finding the optimal by stochastic dynamic programming. This property, as discussed in Remark 5 to Theorem 3.3.2, does not hold in general.

4.3.5 Evaluation of Average Cost-to-Go Using Average of Deterministic Optimal Dynamic Control

The second expected value mentioned in the Remark of Subsection 4.3.1 must now be calculated. One again assumes that the innovations $\{\underline{\hat{\xi}}_t\}$ are fixed and calculates the dynamic cost-to-go (i.e., $\hat{J}_u$ of equation (4.3.18) is evaluated) assuming that control law (4.3.43) is applied. One then considers the innovations random again and averages over them.

Let the value of $\hat{J}_u$ which is realized when the control (4.3.43) is applied and $\{\underline{\hat{\xi}}_t\}$ is fixed be denoted $I_0(\underline{\hat{x}}_{0|0}, \underline{\hat{\xi}}_0, \ldots, \underline{\hat{\xi}}_{T-1})$:

$$I_0(\underline{\hat{x}}_{0|0}, \underline{\hat{\xi}}_0, \ldots, \underline{\hat{\xi}}_{T-1}) = \frac{1}{2}\underline{\hat{x}}_T'\underline{Q}_T\underline{\hat{x}}_T + \frac{1}{2}\sum_{t=0}^{T-1}\underline{\hat{x}}_t'(\underline{Q}_t + \underline{K}_t'\underline{R}_t\underline{K}_t)\underline{\hat{x}}_t \quad (4.3.45)$$

where

$$\underline{\hat{x}}_{t+1} = (\underline{A}_t - \underline{B}_t\underline{K}_t)\underline{\hat{x}}_t \,; \qquad \underline{\hat{x}}_{0|0} \text{ fixed} \quad (4.3.46)$$

$$\underline{K}_t = -(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{A}_t \quad (4.3.47)$$

and where $\{\underline{M}_t\}$ satisfies (4.3.38) - (4.3.39). The evaluation of I_0 will be summarized as a lemma:

Lemma 4.3.1 The value of I_0 given by (4.3.45) is

$$I_0(\hat{\underline{x}}_{0|0}, \hat{\underline{\xi}}_0, \ldots, \hat{\underline{\xi}}_{T-1}) = \frac{1}{2}\{\hat{\underline{x}}'_{0|0}\underline{M}_0\hat{\underline{x}}_{0|0} + \hat{\underline{x}}'_{0|0}\underline{m}_0 + c_0\} \quad (4.3.48)$$

where $\hat{\underline{x}}_{0|0}$ is the initial state and where

$$\underline{M}_t = \underline{Q}_t + \underline{A}'_t[\underline{M}_{t+1} - \underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}'_t\underline{M}_{t+1}]\underline{A}_t \quad (4.3.49)$$

$$\underline{M}_T = \underline{Q}_T \quad (4.3.50)$$

$$\underline{m}_t = [\underline{A}_t - \underline{B}_t(\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}'_t\underline{M}_{t+1}\underline{A}_t]'[2\underline{M}_{t+1}\hat{\underline{\xi}}_{t+1} + \underline{m}_{t+1}] \quad (4.3.51)$$

$$\underline{m}_T = \underline{0} \quad (4.3.52)$$

$$c_0 = \sum_{t=0}^{T-1} \hat{\underline{\xi}}'_t\underline{M}_{t+1}\hat{\underline{\xi}}_t + \underline{m}'_{t+1}\hat{\underline{\xi}}_t \quad (4.3.53)$$

Proof: Given in Appendix B.

It will now be convenient to substitute the value of $\hat{\underline{x}}_{0|0}$ in terms of the a-priori estimate $\hat{\underline{x}}_0$, as given in equation (4.3.3), into the expression for I_0 given in Lemma 4.3.1. The result is

$$I_0 = \frac{1}{2}\{(\hat{\underline{x}}_0 + \hat{\underline{\xi}}_{-1})'\underline{M}_0(\hat{\underline{x}}_0 + \hat{\underline{\xi}}_{-1}) + (\hat{\underline{x}}_0 + \hat{\underline{\xi}}_{-1})'\underline{m}_0 + c_0\} \quad (4.3.54)$$

Having expressed the cost-to-go using the average control law (4.3.43) and assuming the innovations $\{\hat{\underline{\xi}}_t\}$ fixed, the only remaining step is to once more view the innovations sequence as random and take the expected value of I_0 over the $\hat{\underline{\xi}}_t$ vectors. Carrying this out gives

$$\hat{I}_0(\hat{\underline{x}}_0) \triangleq E\{I_0 | \hat{\underline{x}}_0\}$$

$$= E\{\hat{\underline{x}}_0' \underline{M}_0 \hat{\underline{x}}_0 + \hat{\underline{\xi}}_{-1}' \underline{M}_0 \hat{\underline{\xi}}_{-1} + 2\hat{\underline{\xi}}_{-1}' \underline{M}_0 \hat{\underline{x}}_0 + (\hat{\underline{x}}_0 + \hat{\underline{\xi}}_{-1})' \underline{m}_0 + c_0 | \hat{\underline{x}}_0\} \tag{4.3.55}$$

$$= \hat{\underline{x}}_0' \underline{M}_0 \hat{\underline{x}}_0 + E\{\hat{\underline{\xi}}_{-1}' \underline{M}_0 \hat{\underline{\xi}}_{-1} + c_0 | \hat{\underline{x}}_0\} \tag{4.3.56}$$

where the expected value of the cross term in $\hat{\underline{x}}_0$ and $\hat{\underline{\xi}}_{-1}$ is zero by the zero mean of $\hat{\underline{\xi}}_{-1}$, the cross term in $\hat{\underline{x}}_0$ and $\underline{m}_0$ is zero because $\underline{m}_0$ is zero on the average because of the zero average of $\{\hat{\underline{\xi}}_t\}$ (see (4.5.51) and (4.5.52)), and the cross term in $\hat{\underline{\xi}}_{-1}$ and $\underline{m}_0$ is zero because they are independent and zero mean (again see (4.5.51)). Continuing from (4.3.56), note that

$$E\{\hat{\underline{\xi}}_{-1} \underline{M}_0 \hat{\underline{\xi}}_{-1} + c_0 | \hat{\underline{x}}_0\} = E\{\sum_{t=0}^{T} \hat{\underline{\xi}}_{t-1}' \underline{M}_t \hat{\underline{\xi}}_{t-1} + \underline{m}_{t+1}' \hat{\underline{\xi}}_t | \hat{\underline{x}}_0\} \tag{4.3.57}$$

$$= \sum_{t=0}^{T} tr[\underline{M}_t \cdot cov(\hat{\underline{\xi}}_{t-1}, \hat{\underline{\xi}}_{t-1} | \hat{\underline{x}}_0)] \tag{4.3.58}$$

where the $\underline{m}_{t+1}$ terms again are zero since $\underline{m}_{t+1}$ is independent of $\hat{\underline{\xi}}_t$.

Now substituting the value of $cov(\hat{\xi}_t, \hat{\xi}_t)$ from equation (4.3.8) and using the fact that the whiteness of the noise makes the conditioning on $\hat{\underline{x}}_0$ superfluous, one obtains

$$\hat{I}_0(\hat{\underline{x}}_0) = \frac{1}{2} \hat{\underline{x}}_0' \underline{M}_0 \hat{\underline{x}}_0 + \frac{1}{2} \sum_{t=-1}^{T-1} tr\{\underline{M}_{t+1} \underline{S}_{t+1} [\underline{C}_{t+1} (\underline{A}_t \underline{\Sigma}_t \underline{A}_t' + \underline{\Xi}_t) \underline{C}_{t+1}'$$

$$+ \underline{D}_{t+1} \underline{\Theta}_{t+1} \underline{D}_{t+1}'] \underline{S}_{t+1}'\} \tag{4.3.59}$$

Note the lower limit of the summation is minus one.

The form of this expression is important. Notice that there is one term (the first) which depends on the initial state and is independent of the sequence $\{\underline{v}_t\}$ and one term (the summation) which is independent of the plant state and dynamic control, but which is influenced by the measurement control sequence $\{\underline{v}_t\}$. In other words, in the expected value of the dynamic cost resulting from using the average of the optimal deterministic control law (equation (4.3.43)) there are terms which are influenced by the measurement control. Thus equation (4.3.19) does not represent the entire contribution of the measurement control to the overall cost; one must consider the sum of equation (4.3.19) and the summation term in (4.3.59).

4.3.6 Consideration of Measurement Control Optimization

What has been done up to now in this section? In subsection 4.3.1, the stochastic linear system with quadratic cost was transformed to the Kalman filter system. In subsection 4.3.2, the innovations were considered fixed and the Minimum Principle was applied. In doing this, the measurement optimization aspects of the problem were supressed. In subsections 4.3.3 and 4.3.4, the optimal deterministic feedback dynamic control law was derived and its expected value was taken after re-introducing the random character of the innovations. This did not re-introduce the influence of the measurement control. The measurement control was finally re-introduced in subsection 4.3.5, in which the average dynamic cost-to-go using the expectation of the deterministic optimal dynamic control law was evaluated. The collection of terms in the total stochastic cost-to-go which are influenced by the measurement control were shown to be the sum of $\hat{J}_v$ given in (4.3.19) and certain terms in (4.3.18) when the expected value is taken. Let this total cost influenced by the measurement control sequence $\{\underline{v}_t\}$ be denoted by $\hat{J}_{v,total}$:

$$\hat{J}_{v,total} = \hat{J}_v + \frac{1}{2} \sum_{t=-1}^{T-1} tr\{\underline{M}_{t+1}\underline{S}_{t+1}[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_t\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']\underline{S}_{t+1}'\} \tag{4.3.60}$$

Consideration of equation (4.3.60) and the analysis of the previous five subsections finally shows that the complete two-way separation property holds:

1. the calculation of the optimal dynamic control law (equation (4.3.43)) is independent of the measurement control, and
2. the covariance equations, and the minimization of $J_{v,total}$ in (4.3.60), subject to the covariance equation (4.3.10), is independent of the evolution of the conditional mean $\hat{\underline{x}}_t$ and can be done a-priori.

There remains only one task to complete in order to demonstrate the consistency of the analysis in this section using the results of Chapter III and the analysis of the previous section using the results of Chapter II: One must show that the subproblem which must be solved to determine the optimal measurement control sequence is the same regardless of the analysis which led to that subproblem. More specifically, one must show that subject to the equations given here (re-introducing the $t|t$ notation)

$$\underline{\Sigma}_{t+1|t} = \underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t \; ; \quad \underline{\Sigma}_{0|-1} = \underline{\Sigma}_0 \tag{4.3.61}$$

$$\underline{\Sigma}_{t+1|t+1} = \underline{\Sigma}_{t+1|t} - \underline{S}_{t+1}[\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']\underline{S}_{t+1}' \tag{4.3.62}$$

$$\underline{S}_{t+1} = \underline{\Sigma}_{t+1|t}\underline{C}_{t+1}'[\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']^{-1} \tag{4.3.63}$$

$$\underline{K}_t = -(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{A}_t \tag{4.3.64}$$

$$\underline{M}_t = \underline{Q}_t + \underline{A}_t'\underline{M}_{t+1}\underline{A}_t - \underline{A}_t'\underline{M}_{t+1}\underline{B}_t\underline{K}_t ; \quad \underline{M}_T = \underline{Q}_T \tag{4.3.65}$$

The following two quantities are equal: the cost influenced by $\{\underline{v}_t\}$ as derived by dynamic programming and the cost influenced by $\{\underline{v}_t\}$ as derived by deterministic methods. These quantities are*

$$J_v = \mathrm{tr}\{\underline{M}_0\underline{\Sigma}_0 + \sum_{t=0}^{T-1} \underline{M}_{t+1}(\underline{\Xi}_t + \underline{B}_t\underline{K}_{t+1}\underline{\Sigma}_{t|t}\underline{A}_t')\} + \sum_{t=0}^{T} \ell_t(\underline{v}_t) \tag{4.3.66}$$

and

$$\hat{J}_{v,\,total} = \sum_{t=0}^{T} \ell_t(\underline{v}_t) + \mathrm{tr}\{\underline{Q}_T\underline{\Sigma}_{T|T} + \sum_{t=0}^{T-1} \underline{Q}_t\underline{\Sigma}_{t|t} + \sum_{t=-1}^{T-1} + M_{t+1}S_{t+1}$$

$$\times [\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t)\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']\underline{S}_{t+1}'\} \tag{4.3.67}$$

where the quantity obtained by dynamic programming, J_v, is taken from Theorem 4.2.2, equation (4.2.54) and the quantity obtained by deterministic methods, $\hat{J}_{v,\,total}$, is taken from equation (4.3.60).

---

* The factor of $\frac{1}{2}$ has been deleted from $\hat{J}_{v,\,total}$. This changes nothing.

It is not obvious that J_v and $\hat{J}_{v,\,total}$ are equal, but a certain amount of algebra shows that indeed they are. The details are given in Appendix C.

This completes the analysis of the linear-Gaussian-quadratic problem with measurement control capability using the Minimum Principle. The results have been shown to be identical to those obtained by dynamic programming.

4.4 Results Using the Open Loop Feedback Optimal Strategy

In the two previous sections, the linear-Gaussian-quadratic problem with measurement control capability was analyzed and the optimal cost realized using a true stochastic optimal control was derived. In this section, the open loop feedback optimal (O. L. F. O.) control strategy introduced in subsection 2.4.3 will be applied to the same problem. This will serve the purpose of clearly illustrating exactly what one gains, in terms of a smaller cost, as a result of taking measurements.

As was pointed out in subsection 2.4.3, the O. L. F. O. approach is not new. This section is included, however, to provide a useful comparison to the results of the previous two.

Theorem 4.4.1: If the linear-Gaussian-quadratic system (4.1.1) - (4.1.3) is analyzed using the O. L. F. O. control strategy of subsection 2.4.3, the following properties hold:

(1) If j is a time point in the time index set $\mathfrak{J} = \{0, 1, \ldots, T\}$, and if the O. L. F. O. strategy is applied from time j to the terminal time T using the assumption

that no further measurements will be added to the current measurement set Y_j, then for all $t \geq j$, the O. L. F. O. dynamic control, denoted by $\underline{\tilde{u}}_t$, is given by $\underline{\tilde{u}}_t = -\underline{K}_t \underline{\hat{x}}_{t|j}$ (4.4.1)

where $\underline{K}_t$ is a gain matrix identical to that which appears in the true stochastic solution and where $\underline{\hat{x}}_{t|j}$ is the conditional mean of $\underline{x}_t$ given Y_j, U_{j-1}, and V_j.

(2) The expected cost-to-go $\tilde{\mathcal{J}}_j$ from step j to Step T using the O. L. F. O. control (4.4.1) is given by

$$\tilde{\mathcal{J}}_j = \hat{x}_{j|j}' \underline{M}_j \underline{\hat{x}}_{j|j} + \mathrm{tr}\{\underline{M}_j \underline{\Sigma}_{j|j} + \sum_{t=j}^{T-1} \underline{M}_{t+1} [\Xi_t + \underline{B}_t \underline{K}_t \Sigma_{t|j} A_t'] \} \tag{4.4.2}$$

where:

$$\underline{\hat{x}}_{j|j} = \text{estimate of } \underline{x}_j \text{ given the data } Y_j, U_{j-1}, V_j \tag{4.4.3}$$

$$\underline{\Sigma}_{j|j} = \text{covariance of } \underline{x}_j \text{ given the data } Y_j, U_{j-1}, V_j. \tag{4.4.4}$$

$$\underline{\Sigma}_{t+1|j} = \underline{A}_t \underline{\Sigma}_{t|j} \underline{A}_t' + \underline{\Xi}_t \tag{4.4.5}$$

and where $\{\underline{M}_t\}$ satisfies

$$\underline{M}_t = \underline{Q}_t + \underline{A}_t'[\underline{M}_{t+1} - \underline{M}_{t+1}\underline{B}_t(R_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}]\underline{A}_t \tag{4.4.6}$$

$$\underline{M}_T = \underline{Q}_T \tag{4.4.7}$$

Remark 1: As a result of property (1) above and the analysis of the previous two sections, one can say that for a linear-Gaussian-quadratic problem, the optimal dynamic control at time t is always given by the gain matrix $\underline{K}_t$ times the best estimate of the plant state $\underline{x}_t$ at time t, whether or not measurements have been made up to time t.

Remark 2: Comparison of equations (4.4.2) - (4.4.5) with the results of Theorem 4.2.2 shows that the form of the terms in the cost-to-go which involve the covariance matrix (assuming the use of the optimal dynamic control) is the same whether measurements are made or not. The difference is entirely in the propagation equation for the covariance (compare (4.4.5) and (4.2.55), (4.2.56)). Since the covariance equation is nonlinear if measurements are made, one cannot conveniently derive an expression in closed form for the benefit derived from measuring.

Proof of Theorem 4.4.1: Let j be a fixed integer in the time index set. Suppose the measurement set Y_j is available along with all the control values that have been applied up to time j so that $\underline{x}_j$ is a Gaussian random vector. Let $\hat{\underline{x}}_{j|j}$ represent the conditional mean of $\underline{x}_j$ given $\tilde{Y}_j$ and $\underline{\Sigma}_{j|j}$ the conditional covariance. Assume that no measurements will be made beyond time j. Then the cost-to-go from time j to the terminal time T, on the average, is given by*

$$\tilde{\mathcal{J}}_j = E\{\underline{x}_T'\underline{Q}_T\underline{x}_T + \sum_{t=j}^{T-1} \underline{x}_t'\underline{Q}_t\underline{x}_t + \underline{u}_t'\underline{R}_t\underline{u}_t | \tilde{Y}_j\} \tag{4.4.7}$$

Let $\hat{\underline{x}}_{t|j}$ represent the conditional mean of $\underline{x}_t$ given $\tilde{Y}_j$ and $\underline{\Sigma}_{t|j}$ represent the conditional covariance. Then using (4.3.15) and considering $\underline{u}_t$ to be

---

* There are no $\ell_t(\underline{v}_t)$ terms in the cost because there are no measurements.

deterministic* for all t greater than or equal to j in equation (4.4.7) allows one to evaluate $\mathcal{J}_j$:

$$\tilde{\mathcal{J}}_j = \hat{\underline{x}}'_{T|j}\underline{Q}_T\hat{\underline{x}}_{T|j} + \sum_{t=j}^{T-1} \hat{\underline{x}}'_{t|j}\underline{Q}_t\hat{\underline{x}}_{t|j} + \underline{u}'_t\underline{R}_t\underline{u}_t$$

$$+ \mathrm{tr}\{\underline{Q}_T\underline{\Sigma}_{T|j} + \sum_{t=j}^{T-1} \underline{Q}_t\underline{\Sigma}_{t|j}\} \qquad (4.4.8)$$

The conditional mean $\hat{\underline{x}}_{t|j}$ propagates according to

$$\hat{\underline{x}}_{t+1|j} = \underline{A}_t\hat{\underline{x}}_{t|j} + \underline{B}_t\underline{u}_t \qquad (4.4.9)$$

and the conditional covariance $\underline{\Sigma}_{t|j}$ according to

$$\underline{\Sigma}_{t+1|j} = \underline{A}_t\underline{\Sigma}_{t|j}\underline{A}'_t + \underline{\Xi}_t \qquad (4.4.10)$$

Equations (4.4.9) and (4.4.10) may be viewed as the Kalman Filter equations in the case that $\underline{C}_t \equiv \underline{0}$ (no measurements).

At this point, it is clear that in order to pick the O.L.F.O. dynamic control, it is only necessary to minimize those terms in (4.4.8) which $\underline{u}_t$ influences, namely

$$\tilde{\mathcal{J}}_{j,\,u\ \text{terms}} = \hat{\underline{x}}'_{T|j}\underline{Q}_T\hat{\underline{x}}_{T|j} + \sum_{t=j}^{T-1} \hat{\underline{x}}'_{t|j}\underline{Q}_t\hat{\underline{x}}_{t|j} + \underline{u}'_t\underline{R}_t\underline{u}_t \qquad (4.4.11)$$

---

* $\underline{u}_t$ must be deterministic because Y_j has been observed and is no longer random, and no other measurements will be made, by assumption.

This is to be minimized subject to (4.4.9). The combination of (4.4.9) and (4.4.11) is nothing more or less than the usual linear-quadratic optimal regulator, the solution of which is widely known (see Bryson and Ho (B.4)). The O.L.F.O. control is given by

$$u_t = -\underline{K}_t \hat{\underline{x}}_{t|j} \tag{4.4.12}$$

where

$$\underline{K}_t = (\underline{R}_t - \underline{B}_t' \underline{M}_{t+1} \underline{B}_t')^{-1} \underline{B}_t \underline{M}_{t+1} \underline{A}_t \tag{4.4.13}$$

$$\underline{M}_t = \underline{Q}_t + \underline{A}_t' \underline{M}_{t+1} \underline{A}_t - \underline{A}_t' \underline{M}_{t+1} \underline{B}_t \underline{K}_t \tag{4.4.14}$$

$$\underline{M}_T = \underline{Q}_T \tag{4.4.15}$$

In addition, the value of $\tilde{\mathcal{J}}_{j,\, u\ \text{terms}}$ which is achieved as a result of control (4.4.12) is

$$\tilde{\mathcal{J}}_j = \hat{\underline{x}}_{j|j}' \underline{M}_j \hat{\underline{x}}_{j|j} \tag{4.4.16}$$

Thus the O.L.F.O. value of $\tilde{\mathcal{J}}_j$ given in (4.4.8) is

$$\tilde{\mathcal{J}}_j = \hat{\underline{x}}_{j|j} \underline{M}_j \hat{\underline{x}}_{j|j} + \mathrm{tr}\{\underline{Q}_T \underline{\Sigma}_{T|j} + \sum_{t=j}^{T-1} \underline{Q}_t \underline{\Sigma}_{t|j}\} \tag{4.4.17}$$

Using (4.4.14) to express $\underline{Q}_t$ in terms of $\underline{M}_t$ and $\underline{M}_{t+1}$, and using (4.4.15) results in

$$\tilde{\mathcal{J}}_j = \hat{\underline{x}}_{j|j}' \underline{M}_j \hat{\underline{x}}_{j|j} + \mathrm{tr}\{\underline{M}_T \underline{\Sigma}_{T|j} + \sum_{t=j}^{T-1} [\underline{M}_t - \underline{A}_t' \underline{M}_{t+1} \underline{A}_t$$

$$+ \underline{A}_t' \underline{M}_{t+1} \underline{B}_t \underline{K}_t] \underline{\Sigma}_{t|j}\} \tag{4.4.18}$$

Using (4.4.10) this can be expressed

$$\tilde{\mathcal{J}}_j = \hat{\underline{x}}'_{j|j}\underline{M}_j\hat{\underline{x}}_{j|j} + \mathrm{tr}\{\underline{M}_T(\underline{A}_{T-1}\underline{\Sigma}_{T-1|j}\underline{A}'_{T-1} + \underline{\Xi}_{T-1})$$

$$+ \underline{M}_j\underline{\Sigma}_{j|j} + \sum_{t=j+1}^{T-1} \underline{M}_t(\underline{A}_{t-1}\underline{\Sigma}_{t-1|j}\underline{A}'_{t-1} + \underline{\Xi}_{t-1})$$

$$- \sum_{t=j}^{T-1} \underline{A}'_t\underline{M}_{t+1}\underline{A}_t\underline{\Sigma}_{t|j} + \sum_{t=j}^{T-1} \underline{A}'_t\underline{M}_{t+1}\underline{B}_t\underline{K}_t\underline{\Sigma}_{t|j}\} \qquad (4.4.19)$$

which simplifies to

$$\tilde{\mathcal{J}}_j = \hat{\underline{x}}'_{j|j}\underline{M}_j\hat{\underline{x}}_{j|j} + \mathrm{tr}\{\underline{M}_j\underline{\Sigma}_{j|j} + \sum_{t=j}^{T-1} \underline{M}_{t+1}(\underline{B}_t\underline{K}_t\underline{\Sigma}_{t|j}\underline{A}'_t + \underline{\Xi}_t)\}$$

(4.4.20)

This is exactly the expression that was hypothesized in equation (4.4.2)

Q.E.D.

As mentioned above in Remark 2, the nonlinearity of the covariance equation when measurements are taken prevents one from deriving a closed form expression for the difference between $\tilde{\mathcal{J}}_0$, the O.L.F.O. cost-to-go from time zero, and the true stochastic optimal. Note, however, that if the measurement control $\underline{v}_t$ operates in such a that no measurements are made if $\underline{v}_t = \underline{0}$, then the true stochastic optimal solution of the previous two sections includes O.L.F.O. as a special case: in carrying out the a-priori determination of $\{\underline{v}_t^*\}$ as described in the second part of Theorem 4.2.2, one may find that open loop operation is optimal over some interval. In fact, the proof of Theorem 4.4.1 above is in a sense redudant since the expression for J_v given in Theorem 4.2.2, which turns out to be $\tilde{\gamma}_{0,v}$ terms in Theorem 4.4.1, is correct for <u>any</u> measurement policy, including no measurements.

4.5 Summary

The results of Chapters II and III have been applied to linear-Gaussian-quadratic stochastic optimal control problems in this chapter. In particular, the optimal controls were found using dynamic programming (the Chapter II method) in Section 4.2 and using the Minimum Principle (the Chapter III method) in Section 4.3. These analyses demonstrated that:

(1) The optimal measurement control sequence can be computed off-line and is simply applied open-loop to the system as time evolves.

(2) The optimal dynamic control sequence is computed on-line by multiplying the plant state estimate by a pre-computed gain matrix.

These structural properties of the optimal control laws were illustrated in Figure 4.2. In showing this structure to be optimal, the so-called <u>a-priori-v</u> technique for determining the measurement control sequence (introduced in Section 1.4) was shown to be optimal. This was shown both by dynamic programming and by use of the Minimum Principle.

Finally, the open loop feedback optimal control technique was applied to linear-Gaussian-quadratic systems, and it was shown that the O.L.F.O. dynamic control at each time t could be computed by multiplying the same gain matrix described in (2) above by the estimate of the plant state at time t given whatever measurements have been made.

CHAPTER V

STRUCTURE OF DYNAMIC AND MEASUREMENT CONTROL SYSTEMS FOR LINEAR-GAUSSIAN PROBLEMS WITH GENERAL COSTS

5.1 General Perspective

Having considered linear-Gaussian-quadratic problems in the previous chapter, the natural extension is to next consider problems in which the linearity of the system and Gaussian statistics of the noise are maintained, but in which the cost functional is non-quadratic. By keeping the linear system and Gaussian noise assumptions, one preserves the ability to generate state estimates by use of the Kalman filter. In other words, the estimation problem is still relatively simple. The control part of the problem becomes more difficult, however.

In the previous chapter, it was shown that for a linear-Gaussian-quadratic problem, two separation properties hold: The optimal measurement control sequence $\{\underline{v}_t^*\}$ may be determined off-line by solving a certain deterministic control problem, and the optimal dynamic control sequence $\{\underline{u}_t^*\}$ may be determined on-line by multiplying the best estimate of the state by a pre-computable gain matrix obtained by solving a second deterministic problem. One might conjecture that the same structure describes the solution to a linear-Gaussian-non-quadratic problem: a reasonable argument would be that the nature of the cost functional does not influence the optimal estimation of the state so the measurement program should be insensitive to the form of the cost on state and dynamic control. It will be demonstrated by counterexample that the analysis above is false, and a more complete discussion will then be given. The analysis will show that it is difficult to specify conditions which guarantee the double separation property in a useful form, i.e., in terms of the nature of the given cost functional. Sufficient conditions for double separation will be discussed, but the conditions given turn out not to be especially useful in applications to practical problems.

With regard to "ordinary" stochastic control problems, that is, those with no measurement control capability, the standard result is that a one-way separation theorem does hold: the optimal control is obtained by first estimating the state using a Kalman filter and then plugging the estimate into the feedback solution of a deterministic problem related to the given stochastic one. See Wonham (W.2), where this property is discussed with regard to continuous-time systems. The separation is one-way because the filter equations influence the nature of the control law. Thus the related problem that one must solve to determine the dynamic control is not the one obtained by simply discarding the noises in the original problem. The nature of the related problem that one must solve in this context will be clarified below as a byproduct of the discussion of the optimal measurement problem.

It should be stressed that for linear-Gaussian-non-quadratic optimal control problems, either with or without measurement control capability, the following structural property holds: the stochastic parameters (e.g., the covariances of the noises) influence the control laws, both dynamic and measurement, if any, while the control problem parameters (i.e., the functions appearing in the non-quadratic cost functional) do not affect the structure or parameters of the plant state estimator.

4.2 Two Numerical Examples

Consider the following linear system, in which all quantities are scalars.

$$x_0 = N(\hat{x}_0, \Sigma_0) \tag{5.2.1}$$

$$x_1 = x_0 + u_0 + \xi_0 \quad ; \quad \xi_0 = N(0,1) \tag{5.2.2}$$

$$y_1 = v_1 x_1 + \theta_1 \quad ; \quad \theta_1 = N(0,1) \tag{5.2.3}$$

$$x_2 = x_1 + u_1 + \xi_1 \quad ; \quad \xi_1 = N(0,1) \tag{5.2.4}$$

$$y_2 = v_2 x_2 + \theta_2 \quad ; \quad \theta_2 = N(0,1) \tag{5.2.5}$$

where $z = N(a,b)$ means z is a Gaussian (normal) random variable with mean a and covariance b. Note that no measurement is made at time zero. This is merely for convenience in one of the examples to be considered below: it allows complete solution in that example, which would be impossible if one had to determine an optimal measurement at $t = 0$.

Two cost functionals will be considered in connection with the system (5.2.1) - (5.2.5): a linear cost and an exponential cost. The optimal controls will be determined by stochastic dynamic programming, and it will be seen that the complete two-way separation property with measurement/dynamic control separation* that was found to hold for

---

* Recall Fig. 4.2 of Section 4.1 and the related discussion.

linear-Gaussian-quadratic problems might or might not hold for other problems. In particular, it will be shown that for the linear cost problem, both the optimal dynamic controls and the optimal measurement controls may be determined a-priori, independent of the state estimate generated on line and independent of each other. In contrast to this, for the exponential cost problem, both the optimal dynamic controls and the optimal measurement controls depend on the on-line state estimate and its covariance. Thus for the exponential cost, the optimal measurement program cannot be determined a-priori, nor are the measurement and dynamic controls independent of each other. As in Chapter IV, let $\tilde{Y}_t$ denote the entire data set $Y_t \cup U_{t-1} \cup V_t$.

5.2.1 Linear Cost

Suppose the cost associated with the system (5.2.1) - (5.2.5) is linear in the state:

$$J = x_0 + x_1 + x_2 + u_0^2 + u_1^2 + v_1^2 + v_2^2 \tag{5.2.6}$$

The optimal controls $\{u_0^*, u_1^*\}$ and $\{v_1^*, v_2^*\}$ will be found by dynamic programming applied as described in Chapter II.

Step 2: The terminal cost-to-go is

$$\mathcal{J}_2 = E\{x_2 | \tilde{Y}_2\} \tag{5.2.7}$$

$$= \hat{x}_{2/2} \tag{5.2.8}$$

Since there is nothing left to optimize, this is $\mathcal{J}_2(\breve{Y}_2)$.

Step 1: The cost-to-go is

$$\mathcal{J}_1 = E\{x_1 + u_1^2 + v_2^2 + \mathcal{J}_2(\breve{Y}_2) \mid \breve{Y}_1\} \tag{5.2.9}$$

$$= E\{x_1 + u_1^2 + v_2^2 + \hat{x}_{2/2} \mid \breve{Y}_1\} \tag{5.2.10}$$

The Kalman filter equations are

$$\hat{x}_{2/2} = \hat{x}_{2/1} + S_2(\theta_2 + v_2(x_2 - \hat{x}_{2/1})) \tag{5.2.11}$$

$$\hat{x}_{2/1} = \hat{x}_{1/1} + u_1 \tag{5.2.12}$$

$$S_2 = \frac{v_2 \Sigma_{2/1}}{1 + v_2^2 \Sigma_{2/1}} \tag{5.2.13}$$

$$\Sigma_{2/1} = \Sigma_{1/1} + 1 \tag{5.2.14}$$

Substitution into (5.2.10) and taking the expected value yields

$$\mathcal{J}_1(\breve{Y}_1) = \hat{x}_{1/1} + u_1^2 + v_2^2 + \hat{x}_{1/1} + u_1 + 0 \tag{5.2.15}$$

since $E\{\theta_2 \mid \breve{Y}_1\} = 0$ and $E\{x_2 - \hat{x}_{2/1} \mid \breve{Y}_1\} = 0$.

Minimizing (5.2.15) with respect to v_2 and u_1 by taking derivatives:

$$\frac{\partial \mathcal{J}_1}{\partial u_1} = 0 = 2u_1 + 1 \ ; \quad u_1^* = -\frac{1}{2} \tag{6.2.16}$$

$$\frac{\partial \mathcal{J}_1}{\partial v_2} = 0 = 2v_2 \ ; \quad v_2^* = 0 \tag{5.2.17}$$

Note that the values of both u_1^* and v_2^* are independent of $\hat{x}_{1/1}$ and $\Sigma_{1/1}$. Substituting these back into $\mathcal{J}_1$ gives $\mathcal{S}_1(\breve{Y}_1)$:

$$\mathcal{S}_1(\breve{Y}_1) = 2\hat{x}_{1/1} - \frac{1}{4} \tag{5.2.18}$$

<u>Step 0:</u> The cost-to-go is

$$\mathcal{J}_0 = E\{x_0 + u_0^2 + v_1^2 + \mathcal{S}_1(\breve{Y}_1) \mid \breve{Y}_0\} \tag{5.2.19}$$

$$= E\{x_0 + u_0^2 + v_1^2 + 2\hat{x}_{1/1} - \frac{1}{4} \mid \breve{Y}_0\} \tag{5.2.20}$$

The Kalman filter equations are

$$\hat{x}_{1/1} = \hat{x}_{0/0} + u_0 + S_1(\theta_1 + v_1(x_1 - \hat{x}_{1/0})) \tag{5.2.21}$$

$$S_1 = \frac{v_1 \Sigma_{1/0}}{1 + v_1^2 \Sigma_{1/0}} \tag{5.2.22}$$

$$\Sigma_{1/0} = \Sigma_{0/0} + 1 \tag{5.2.23}$$

Substituting into equation (5.2.20) and taking expected values:

$$\mathcal{J}_1 = \hat{x}_{0/0} + u_0^2 + v_1^2 + 2\hat{x}_{0/0} + 2u_0 - \frac{1}{4} \tag{5.2.24}$$

Minimizing yields

$$v_1^* = 0 \tag{5.2.25}$$

$$u_0^* = -1 \tag{5.2.26}$$

$$\mathcal{S}_0(\breve{Y}_0) = 3\hat{x}_{0/0} - 5/4 \tag{5.2.27}$$

Since y_0 is constrained to be zero, the measurement set Y_0 is null and (5.2.27) is the overall a-priori minimum cost-to-go. Of course, $\hat{x}_{0/0} = \hat{x}_0$ = a-priori mean. The optimal control sequences are $\{u_0^*, u_1^*\} = \{-1, -\frac{1}{2}\}$ and $\{v_1^*, v_2^*\} = \{0, 0\}$. Nothing depends on the conditional mean or covariance. It therefore makes sense to take no measurements, as it turned out anyway. Note that optimal measurement sequence can be determined a-priori. It turns out that the optimal dynamic controls are degenerate functions of the conditional mean and covariance, viz., constant functions, but that is within the framework being considered.

5.2.2 Exponential Cost

Suppose now that the system (5.2.1)-(5.2.5) has associated with it the cost

$$J = e^{x_0} + u_0^2 + e^{x_1} + u_1^2 + e^{x_2} + v_1^2 + v_2^2 \qquad (5.2.28)$$

This measurement/control problem will also be analyzed using dynamic programming as described in Section 2.3 to demonstrate that the optimal sequence $\{v_1^*, v_2^*\}$ cannot be determined a-priori.

The following theorem is needed:

Theorem 5.2.1 If $x = N(a,b)$, then

$$E\{e^x\} = e^{a+\frac{1}{2}b} \qquad (5.2.29)$$

Proof If $x = N(a, b)$, then $p(x) = c \exp\{(-\frac{1}{2}(x-a)^2/b\}$ is the probability density function of x, where c is a constant. Then

$$E\{e^x\} = \int_{-\infty}^{\infty} e^x\, p(x)\, dx$$

$$= c \int_{-\infty}^{\infty} \exp\{x\} \cdot \exp\{-\frac{1}{2}(x-a)^2/b\}\, dx$$

$$= c \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}(-2bx)/b\} \cdot \exp\{-\frac{1}{2}(x-a)^2/b\}\, dx$$

$$= c \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}[-2bx + (x-a)^2]/b\}\, dx$$

$$= c \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}[x^2 - (2ax + 2bx) + a^2]/b\}\, dx$$

Completing the square yields

$$E\{e^x\} = c \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}[(x-a-b)^2 - (a+b)^2 + a^2]/b\}\, dx$$

$$= \exp\{[\frac{1}{2}(a+b)^2 - \frac{1}{2}a^2]/b\} \cdot c \cdot \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}[x-a-b]^2/b\}\, dx$$

Now the constant c times the integral in the previous line is equal to one, so

$$E\{e^x\} = \exp\{[\frac{1}{2}(a^2+2ab+b^2) - \frac{1}{2}a^2]/b\}$$

$$= e^{a+\frac{1}{2}b}$$

Q.E.D.

Now consider the application of the dynamic programming algorithm.

Step 2: Assume all of the following quantities are fixed: $u_0, u_1, v_1, v_2, y_1, y_2$. Then the cost-to-go is

$$\mathcal{J}_2 = E\{e^{x_2} | \tilde{Y}_2\} \quad (5.2.30)$$

$$= e^{\hat{x}_{2/2} + \frac{1}{2}\Sigma_{2/2}} \quad (5.2.31)$$

since x_2 is a Gaussian random variable with mean $\hat{x}_{2/2}$ and covariance $\Sigma_{2/2}$ given the data up to $t = 2$. Since there is nothing left to optimize, $\mathcal{J}_2 = \mathcal{V}_2$.

Step 1: Assume u_0, v_1, y_1 are fixed. Then the cost-to-go is

$$\mathcal{J}_1 = E\{e^{x_1} + u_1^2 + v_2^2 + \mathcal{V}_2(\hat{x}_{2/2}, \Sigma_{2/2}) | \tilde{Y}_1\} \quad (5.2.32)$$

$$= e^{\hat{x}_{1/1} + \frac{1}{2}\Sigma_{1/1}} + u_1^2 + v_2^2 + E\{e^{\hat{x}_{2/2} + \frac{1}{2}\Sigma_{2/2}} | \tilde{Y}_1\} \quad (5.2.33)$$

One must choose u_1 and v_2 to minimize this expression.

With regard to the optimal choice of v_2, one can prove the following general result:

<u>Lemma 5.2.1</u> For an arbitrary measurement control/dynamic control optimization problem of the type formulated in Section 1.3, if the cost term

$$E\{ L_{T-1}(\underline{x}_{T-1}, \underline{u}_{T-1}, \underline{v}_T) \mid \breve{Y}_{T-1}\} \tag{5.2.34}$$

is minimized with respect to $\underline{v}_T$ by the choice $\underline{v}_T = 0$, regardless of $\underline{x}_{T-1}$ and $\underline{u}_{T-1}$, then the optimal value of $\underline{v}_T$ is zero regardless of the function $L_T(\underline{x}_T)$.

<u>Proof:</u> The optimal cost-to-go from step T is

$$\mathcal{J}_T = E\{L_T(\underline{x}_T) \mid \breve{Y}_T\} \tag{5.2.35}$$

The cost-to-go from step T-1 is

$$\mathcal{J}_{T-1} = E\{L_{T-1}(\underline{x}_{T-1}, \underline{u}_{T-1}, \underline{v}_T) + \mathcal{J}_T \mid \breve{Y}_{T-1}\} \tag{5.2.36}$$

$$= E\{L_{T-1}(\underline{x}_{T-1}, \underline{u}_{T-1}, \underline{v}_T) \mid \breve{Y}_{T-1}\}$$

$$+ E\{E\{L_T(\underline{x}_T) \mid \breve{Y}_T\} \mid \breve{Y}_{T-1}\} \tag{5.2.37}$$

or, applying Theorem 2.3.2 to the last term in (5.2.37)

$$\mathcal{J}_{T-1} = E\{L_{T-1}(\underline{x}_{T-1}, \underline{u}_{T-1}, \underline{v}_T) \mid \breve{Y}_{T-1}\}$$

$$+ E\{L_T(\underline{x}_T) \mid \breve{Y}_{T-1}\} \tag{5.2.38}$$

The only term in (5.2.38) which depends on $\underline{v}_T$ is the first, which by hypothesis is minimized if $\underline{v}_T = 0$. Q.E.D.

Remark: This result makes intuitive sense. Since nothing is "done" with the last measurement, (i.e., no $\underline{u}_T$ is generated using it) one should not make a last measurement if it "costs something" to do so.

Applying this result now to the exponential cost example means that $v_2 = 0$, and $\hat{x}_{2/2}$ and $\Sigma_{2/2}$ may be obtained from $\hat{x}_{1/1}$ and $\Sigma_{1/1}$ by prediction:

$$\hat{x}_{2/2} = \hat{x}_{1/1} + u_1 \tag{5.2.39}$$

$$\Sigma_{2/2} = \Sigma_{1/1} + 1 \tag{5.2.40}$$

Substituting these into (5.2.33) gives:

$$\mathcal{J}_1 = e^{\hat{x}_{1/1} + \frac{1}{2}\Sigma_{1/1}} + u_1^2 + e^{\hat{x}_{1/1} + u_1 + \frac{1}{2}\Sigma_{1/1} + \frac{1}{2}} \tag{5.2.41}$$

To find the optimal u_1, one sets the derivative of $\mathcal{J}_1$ with respect to u_1 equal to zero giving

$$0 = 2u_1 + e^{\hat{x}_{1/1} + u_1 + \frac{1}{2}\Sigma_{1/1} + \frac{1}{2}} \tag{5.2.42}$$

$$u_1^* = -\frac{1}{2} e^{u_1^* + \hat{x}_{1/1} + \frac{1}{2}\Sigma_{1/1} + \frac{1}{2}} \tag{5.2.43}$$

Equation (5.2.43) expresses u_1^* (implicitly) as a function of $\hat{x}_{1/1}$ and $\Sigma_{1/1}$. Examination of the second derivative of $\mathcal{J}_1$ shows that (5.2.43) does indeed represent a minimizing value of u_1. Unfortunately, (5.2.43) cannot be solved in closed form. If it could be, one could substitute the result into (5.2.41) to obtain

$$\mathcal{J}_1 = e^{\hat{x}_{1/1} + \frac{1}{2}\Sigma_{1/1}} + [u_1^*(\hat{x}_{1/1}, \Sigma_{1/1})]^2 - 2u_1^*(\hat{x}_{1/1}, \Sigma_{1/1}) \tag{5.2.44}$$

<u>Step 0</u> It only remains to determine u_0 and v_1.

The cost-to-go is

$$\mathcal{J}_0 = E\{e^{x_0} + u_0^2 + v_1^2 + \mathcal{J}_1(\hat{x}_{1/1}, \Sigma_{1/1}) \,|\, \tilde{Y}_0\} \tag{5.2.45}$$

$$= e^{\hat{x}_0 + \frac{1}{2}\Sigma_0} + u_0^2 + v_1^2 + E\{e^{\hat{x}_{1/1} + \frac{1}{2}\Sigma_{1/1}} + [u_1^*(\hat{x}_{1/1}, \Sigma_{1/1})]^2 - 2u_1^*(\hat{x}_{1/1}, \Sigma_{1/1}) \,|\, \tilde{Y}_0\} \tag{5.2.46}$$

where the propagation equations are

$$\hat{x}_{1/1} = \hat{x}_0 + u_0 + S_1[y_1 - v_1(\hat{x}_0 + u_0)] \tag{5.2.47}$$

$$\Sigma_{1/1} = (1 - S_1 v_1)\Sigma_{1/0} \tag{5.2.48}$$

$$\Sigma_{1/0} = \Sigma_0 + 1 \tag{5.2.49}$$

$$S_1 = \frac{v_1 \Sigma_{1/0}}{1 + v_1^2 \Sigma_{1/0}} \tag{5.2.50}$$

It is impossible to evaluate (5.2.46) in closed form as a function of u_0 and v_1 since one cannot evaluate the expectation. For this reason, a computer program was written to do a numerical evaluation. The procedure was as follows:

1) $\Sigma_{1/1}$ was (analytically) expressed in terms of Σ_0 using (5.2.48) - (5.2.50).

2) A set of possible values for $\hat{x}_{1/1}$ was generated by taking equation (5.2.47) and allowing the innovation $S_1[y_1 - v_1(\hat{x}_0 + u_0)]$ to vary plus and minus five standard deviations from its mean of zero. While doing this, u_0 and v_1 were held at fixed values.

3) For each value of $\hat{x}_{1/1}$ determined in step 2, equation (5.2.43) was solved numerically for $u_1^*(\hat{x}_{1/1}, \Sigma_{1/1})$. The technique used to do this was a logarithmic search: the value of u_1^* was bracketed into successive intervals whose length decreased by a factor of 1/2 at each iteration until u_1^* was determined to within an additive error of plus or minus 0.0005 .

4) Using the results of steps 2 and 3, $\mathcal{J}_1$ given in (5.2.44) was determined.

5) Finally, the expected value $\mathcal{J}_0$ indicated in (5.2.45) was numerically evaluated by a rectangular integration against the Gaussian probability density function. The

variable of integration was the innovation used in step 2 to generate the set of possible $\hat{x}_{1/1}$ values.

A listing of the programs used is given in Appendix D. The APL language was used.

Sources of error in the integration were two: the true expectation is an integral which has limits of $\pm\infty$ while the numerical integral was taken between plus and minus 5 standard deviations from the mean, and the numerical integration was done with a step size of one-tenth standard deviation using a rectangular integration algorithm. A few experiments with step size and limits, however, indicated that the essential nature of the results is not affected by numerical inaccuracy.

The results of the numerical evaluation of $\mathcal{J}_0$ are shown in the figures below. Figure 5.1 shows a plot of values of $\mathcal{J}_0(u_0, v_1)$ at various points in the $u_0 - v_1$ plane when $\hat{x}_0$ is -5 and Σ_0 is 1. The plot is symmetric about the u_0 axis and the values increase to the right of the v_1 axis. Figures 5.2, 5.3, and 5.4 give plots of $\mathcal{J}_0(u_0, v_1)$ when $\hat{x}_0$ is 1, 5, and 10 respectively. In all cases Σ_0 is held at one. Notice that the optimal pairs (u_0^*, v_1^*) are (0,0), (-1.5,0), (-4,0.5), and (-8.5,0.75) as $\hat{x}_0$ increases from -5 to 10. This shows that $v_1^*(\hat{x}_0, \Sigma_0)$ is <u>not</u> independent of $\hat{x}_0$, which in turn shows that the optimal measurement control cannot in general be determined <u>a-priori</u> as was the case for quadratic cost problems.*

---

* To be more precise, although v_1^* depends on $\hat{x}_0$ which is known <u>a-priori</u> in this two-time-step problem, the analysis shows that in an N step problem, v_{N-1}^* depends on $\hat{x}_{N-2/N-2}$, which is not known a-priori.

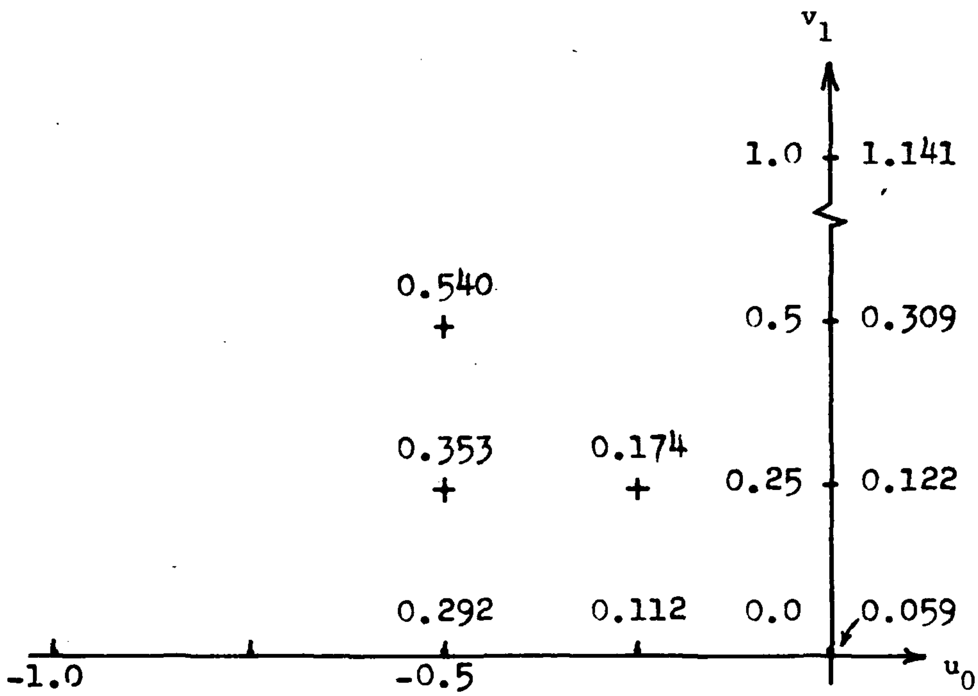

Figure 5.1

Plot of values of $\mathcal{J}_0(u_0,v_1)$ for $\Sigma_0 = 1$, $\hat{x}_0 = -5$

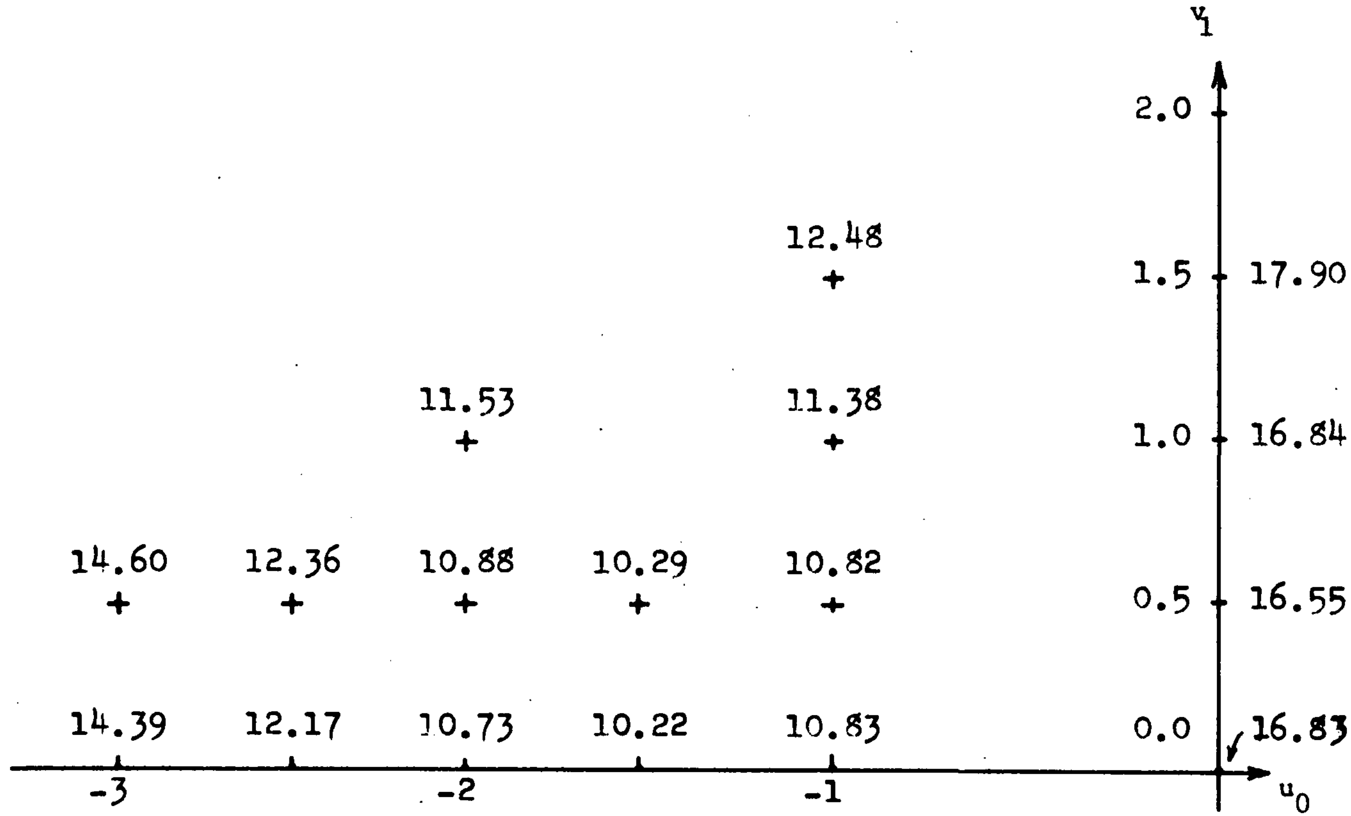

Figure 5.2

Plot of values of $\mathcal{J}_o(u_0,v_1)$ for $\Sigma_0 = 1$, $\hat{x}_0=1$

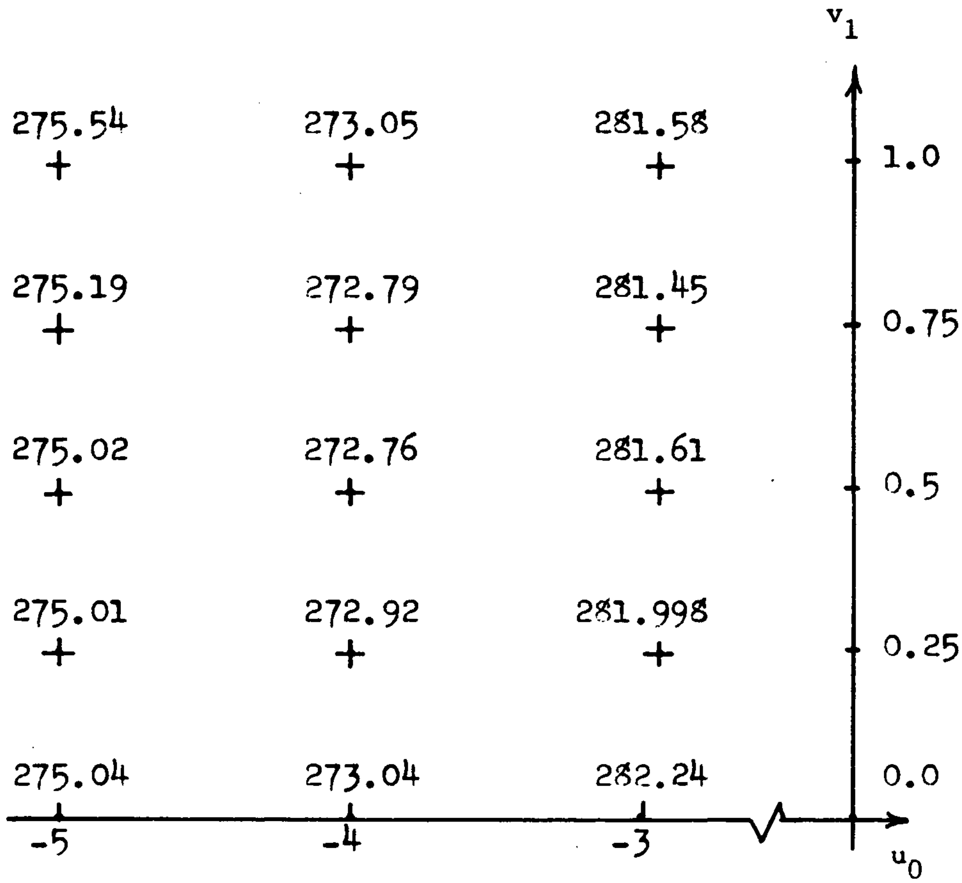

Figure 5.3

Plot of values of $\mathcal{J}_0(u_0,v_1)$ for $\Sigma_0 = 1$, $\hat{x}_0 = 5$

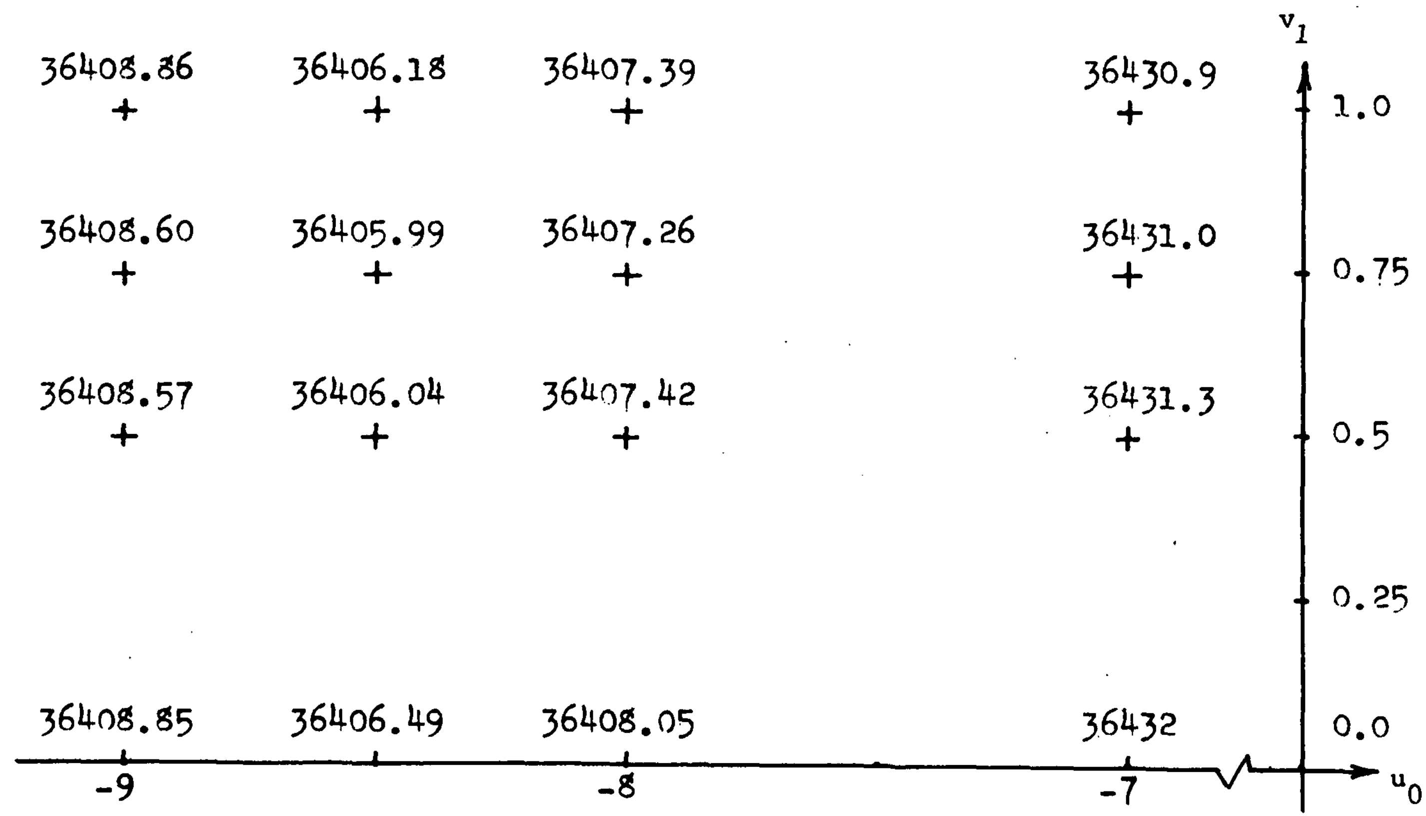

Figure 5.4

Plot of values of $\mathcal{J}_0(u_0, v_1)$ for $\Sigma_0 = 1$, $\hat{x}_0 = 10$

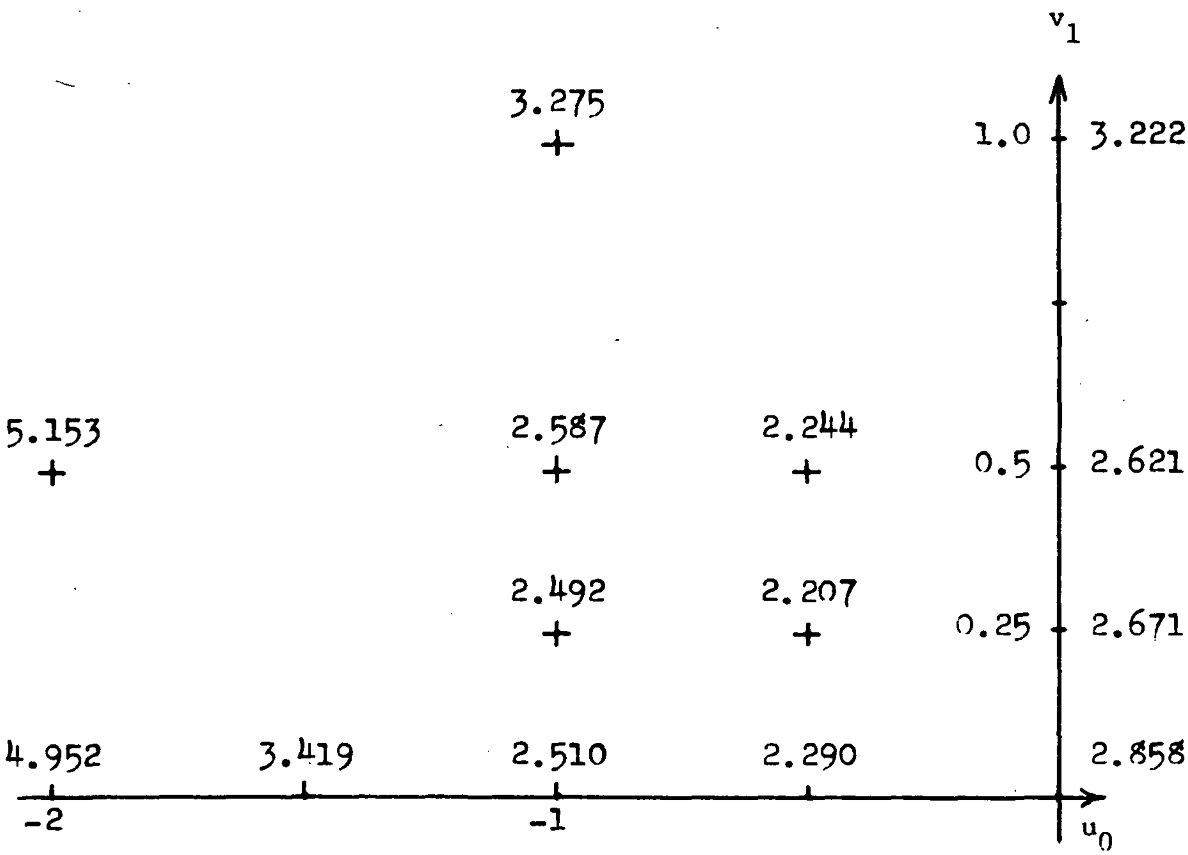

Figure 5.5

Plot of values of $\mathcal{J}_0(u_0,v_1)$ for $\Sigma_0 = 5$, $\hat{x}_0 = -3$

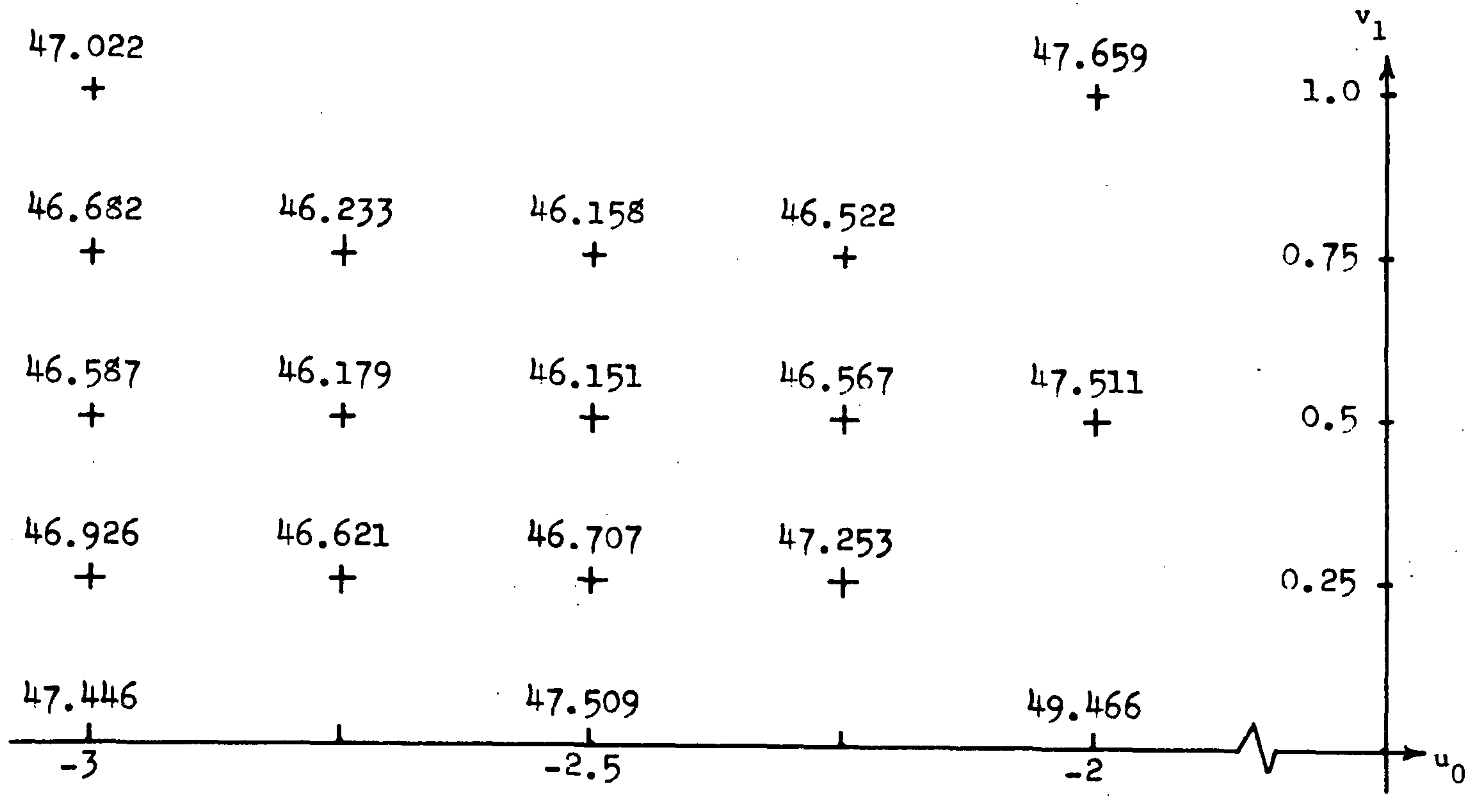

Figure 5.6

Plot of values of $\mathcal{J}_o(u_0,v_1)$ for $\Sigma_0 = 5$, $\hat{x}_0 = 1$

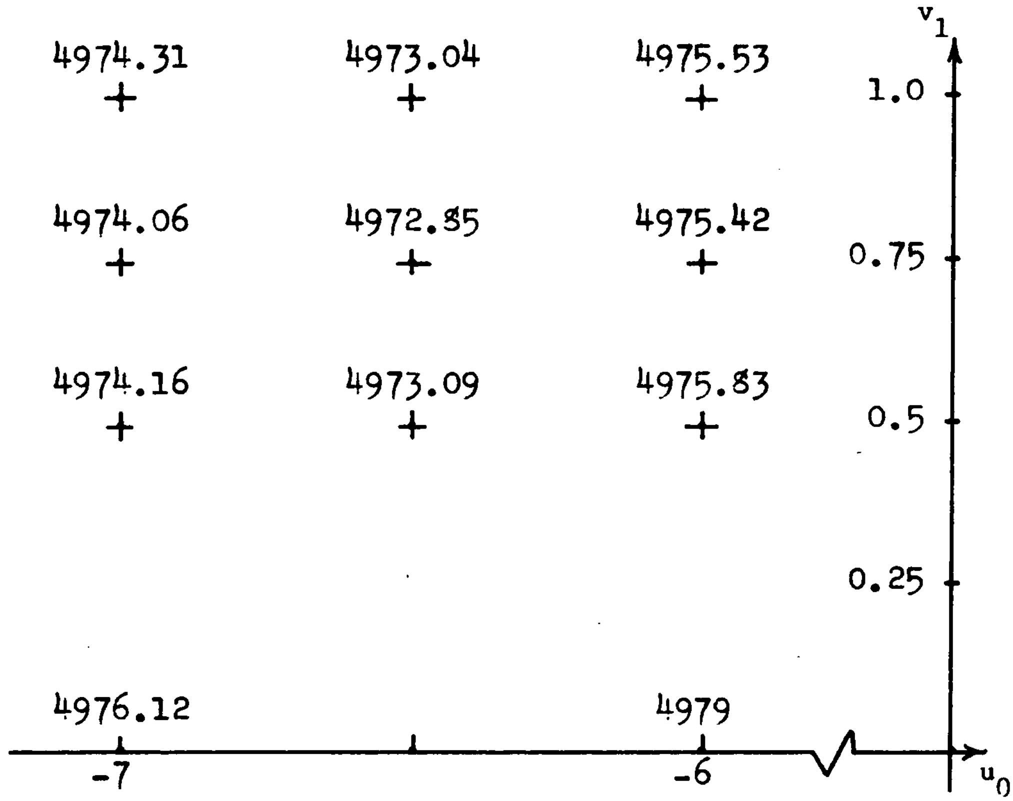

Figure 5.7

Plot of values of $\mathcal{J}_0(u_0,v_1)$ for $\Sigma_0 = 5$, $\hat{x}_0 = 6$

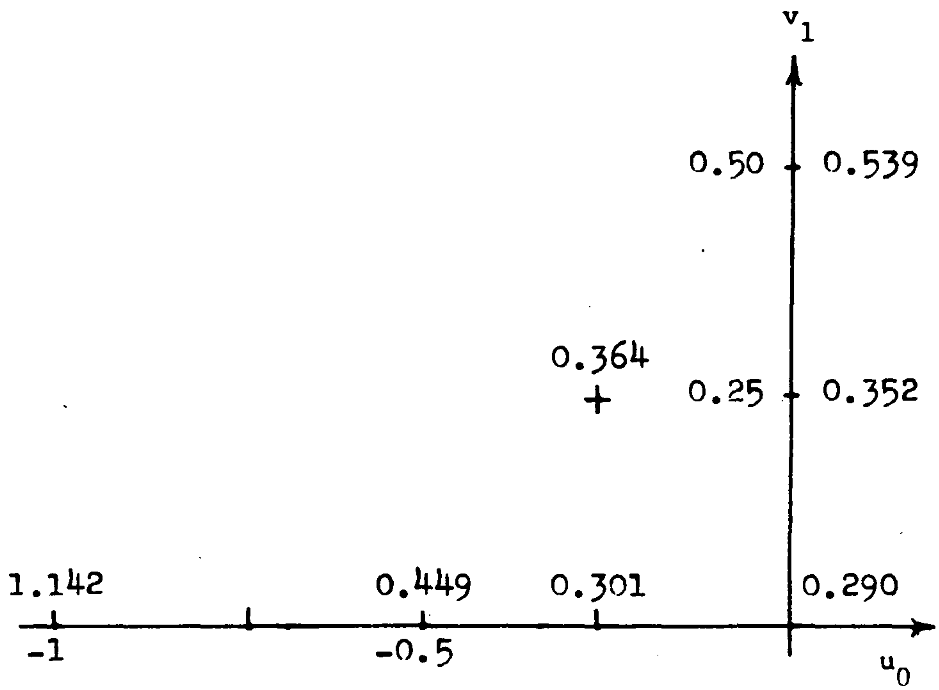

Figure 5.8

Plot of values of $\mathcal{J}_0(u_0,v_1)$ for $\Sigma_0 = 0.2$, $\hat{x}_0 = -3$

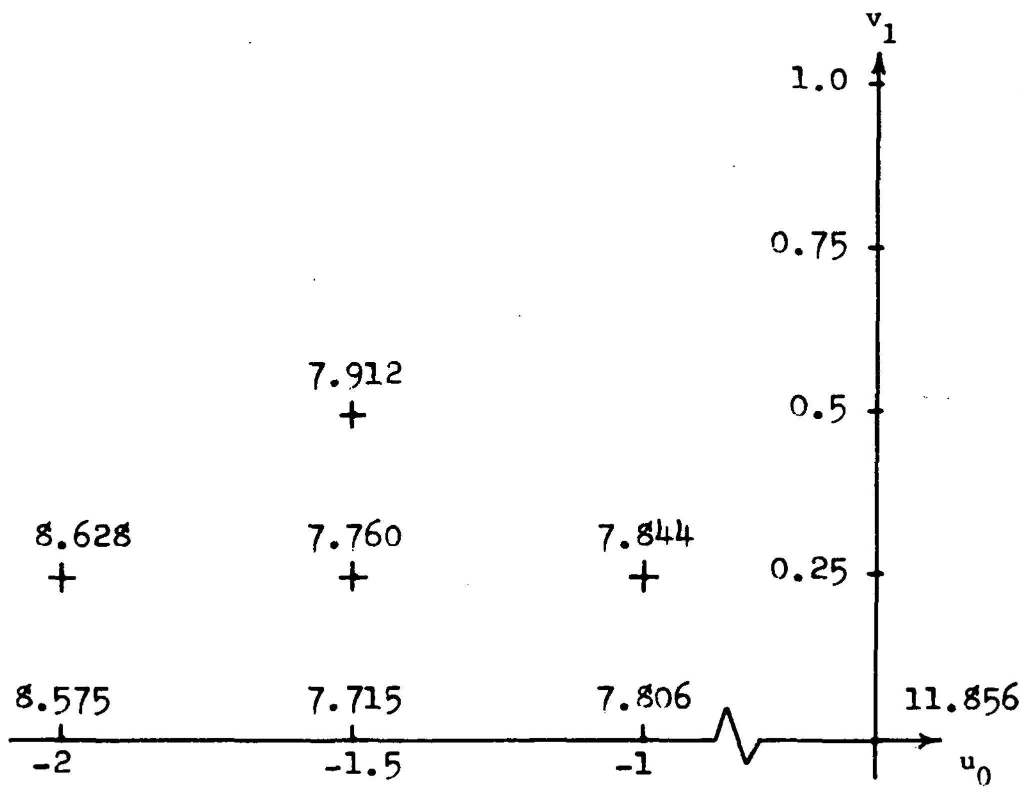

Figure 5.9

Plot of values of $\mathcal{J}_0(u_0,v_1)$ for $\Sigma_0 = 0.2$, $\hat{x}_0 = 1$

PRECEDING PAGE BLANK NOT FILMED

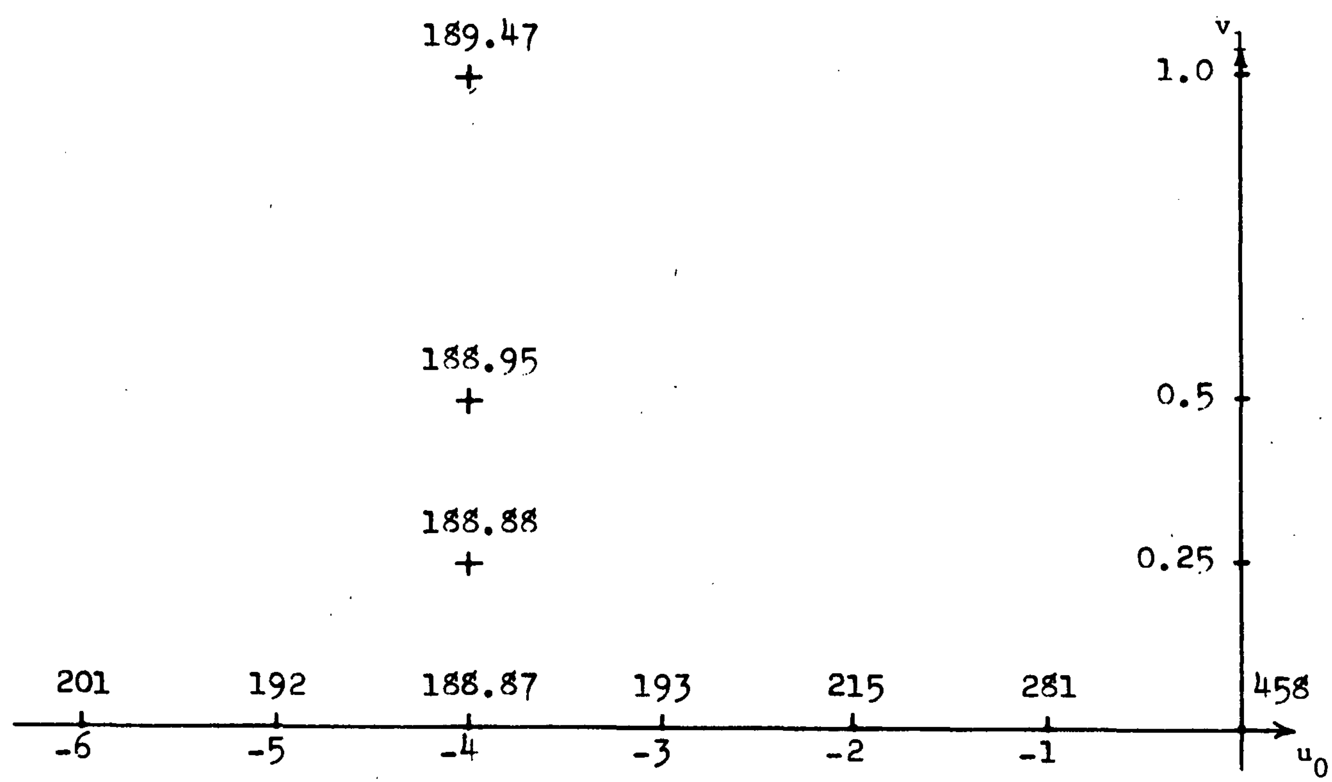

Figure 5.10

Plot of values of $\mathcal{J}_0(u_0,v_1)$ for $\Sigma_0 = 0.2$, $\hat{x}_0 = 5$

Figures 5.5 - 5.7 and 5.8 - 5.10 show plots of $\mathcal{J}_0(u_0, v_1)$ when $\Sigma_0 = 5$ and $\Sigma_0 = 0.2$ for various values of $\hat{x}_0$. Note that for $\Sigma_0 = 5$, v_1^* seems to depend on $\hat{x}_0$ and for $\Sigma_0 = 0.2$, it does not, over the range of $\hat{x}_0$ tested.

The two examples presented in this section have demonstrated the following fact: stochastic optimal control problems with Gaussian noise, non-quadratic costs, and measurement control capability might or might not have a solution with the property that the optimal measurement control sequence can be completely precomputed. The next two sections will consist of analytic consideration of necessary and sufficient conditions which specify when this property does hold.

5.3 An Analysis Using Dynamic Programming

A general linear-Gaussian-non-quadratic stochastic optimal control problem can be formulated around the following equations:

$$\underline{x}_{t+1} = \underline{A}_t \underline{x}_t + \underline{B}_t \underline{u}_t + \underline{\xi}_t \tag{5.3.1}$$

$$\underline{y}_t = \underline{C}_t(\underline{v}_t)\underline{x}_t + \underline{D}_t(\underline{v}_t)\underline{\theta}_t \tag{5.3.2}$$

$$J = L_T(\underline{x}_T) + \sum_{t=0}^{T-1} L_t(\underline{x}_t, \underline{u}_t) + \sum_{t=0}^{T} \ell_t(\underline{v}_t) \tag{5.3.3}$$

where the variables in these expressions have the same meaning and properties as the corresponding ones in Section 3.2. The most general cost functional (see (1.3.3)) is not considered for reasons that will be discussed later in this section.

The question to be considered is: under what circumstances does a measurement optimization separation properly hold for the stochastic control problem described by (5.3.1) - (5.3.3)? Equivently, when is it possible to precompute the optimal measurement profile $\{\underline{v}_t^*\}$? The examples in the previous section show that one cannot do so in general for non-quadratic costs. Conditions under which such a procedure is optimal will be considered in this section, although necessary and sufficient conditions in a useful form have not yet been discovered.

Suppose that one carefully examines the proof of Theorem 4.2.1 (the complete two-way separation theorem for linear-Gaussian-quadratic systems) in order to determine precisely what characteristics of that special problem result in the optimality of the a-priori $\underline{v}$ technique. The apparent reason is that the optimal choice of $\{\underline{v}_t\}$ for the quadratic cost problem depends only upon the evolution of the conditional covariance of the plant state, which may be computed a-priori, and not upon the conditional mean, which may only be determined on-line. Looking more closely shows that the reason why this property holds is that the expected value of a quadratic function of a random vector is the sum of a mean term and a covariance term. Specifically, if $\underline{x}$ is a random vector, then

$$E\{\underline{x}'\underline{Q}\,\underline{x}\} = \hat{\underline{x}}'\underline{Q}\,\hat{\underline{x}} + \mathrm{tr}\left[\underline{Q}\,\underline{\Sigma}\right] \qquad (5.3.4)$$

where $\hat{\underline{x}}$ is the mean of $\underline{x}$ and $\underline{\Sigma}$ is its covariance. Since both the "instantaneous" cost terms $L_t(\underline{x}_t, \underline{u}_t)$ and the optimal cost-to-go from time t to the terminal time T are quadratic in the plant state for the problem analyzed in Chapter IV, at each step in the dynamic programming, the cost-to-go is expressed as the sum of terms involving $\hat{\underline{x}}_{t/t}$ and $\underline{u}_t$ and terms involving $\underline{\Sigma}_{t/t}$ and $\underline{v}_{t+1}$. Thus the optimization of $\underline{v}_{t+1}$ depends only on $\underline{\Sigma}_{t/t}$ and the optimization of $\underline{u}_t$ depends only on $\hat{\underline{x}}_{t/t}$.

For the linear-Gaussian-non-quadratic problem, the conditional covariance still propagates according to a deterministic equation, in fact, precisely the same equation as that used in the quadratic cost problem. One is therefore tempted to make the following conjecture regarding the circumstances under which the a-priori $\underline{v}$ technique would be optimal for a non-quadratic cost problem:

> Conjecture 5.3.1: A necessary and sufficient condition for the a-priori $\underline{v}$ technique to be optimal for the optimal measurement/control problem specified by equations (5.3.1) - (5.3.3) is that under the assumption that $\underline{x}_t$ is $N(\hat{\underline{x}}_{t/t}, \underline{\Sigma}_{t/t})$, one may express

$$E\{L_t(\underline{x}_t, \underline{u}_t)\} = \mathcal{L}_t^1(\hat{\underline{x}}_{t/t}, \underline{u}_t) + \mathcal{L}_t^2(\underline{\Sigma}_{t/t}) \qquad (5.3.5)$$

The analysis below will show that this conjecture is false.

Remark 1: If the two examples considered in Section 5.2 are examined, one sees that the linear cost example, for which the a-priori $\underline{v}$ technique is optimal, indeed satisfies the hypotheses of Conjecture 5.3.1, while the exponential cost example, for which a-priori $\underline{v}$ is not optimal, does not. Of course, this coincidence is not conclusive.

Remark 2: It is not clear that there are any cost functions besides quadratic (and linear, which may be viewed as degenerate quadratics) which satisfy the hypothesis of Conjecture 5.3.1. Power series cost functions of degree higher than two do not.

Examining the hypothesis of Conjecture 5.3.1 and comparing again to the quadratic cost case shows that not all of the features of the quadratic cost problem are included. One is led to the following more restrictive hypothesis:

Conjecture 5.3.2: Necessary and sufficient conditions for the a-priori $\underline{v}$ technique to be optimal for the problem specified by (5.3.1) - (5.3.3) are that the following two conditions be satisfied:

(1) Under the assumption that $\underline{x}_t = N(\hat{\underline{x}}_{t/t}, \underline{\Sigma}_{t/t})$,

$$E\{L_t(\underline{x}_t, \underline{u}_t)\} = \mathcal{L}_t^1(\hat{\underline{x}}_{t/t}, \underline{u}_t) + \mathcal{L}_t^2(\underline{\Sigma}_{t/t})$$

and

(2) For each t, the optimal cost-to-go from the next step $t+1$, $\mathcal{V}_{t+1}(\breve{Y}_{t+1})$, satisfies

$$E\{\mathcal{J}_{t+1}(\tilde{Y}_{t+1}) | \underline{x}_t = N(\hat{\underline{x}}_{t/t}, \Sigma_{t/t})\} = \hat{\mathcal{J}}^1_{t+1}(\hat{\underline{x}}_{t/t}, \underline{u}_t) + \hat{\mathcal{J}}^2_{t+1}(\Sigma_{t/t}, \underline{v}_{t+1}) \tag{5.3.6}$$

The hypotheses of Conjecture 5.3.2 are quite strong and may indeed rule out all but quadratic cost problems. On the other hand, it is only possible at present to prove that the conditions of Conjecture 5.3.2 are sufficient, but not necessary. In addition, it is nearly impossible to test the hypothesis of Conjecture 5.3.2 for any practical problem other than one with quadratic cost. Thus the questions surrounding the linear-Gaussian-non-quadratic problem are not yet fully resolved.

The hypotheses of Conjectures 5.3.1 and 5.3.2 will now be analyzed to illustrate the difficulties one faces in finding true necessary and sufficient conditions. First, the non-sufficiency of the hypothesis of Conjecture 5.3.1 will be shown.

Non-Sufficiency of Conjecture 5.3.1:

It will be shown that even if the hypothesis of Conjecture 5.3.1 holds, it does not follow in general that the optimal measurement control is independent of the on-line plant state estimate. An inductive argument will be made.

Step T: The cost-to-go may be written as

$$\mathcal{J}_T = E\{L_T(\underline{x}_T) | \tilde{Y}_T\} \tag{5.3.7}$$

which by the hypothesis may be expressed in the form

$$\mathcal{J}_T = \mathcal{L}_T^1(\hat{\underline{x}}_{T/T}) + \mathcal{L}_T^2(\underline{\Sigma}_{T/T}) \qquad (5.3.8)$$

Since there is nothing to optimize, this is $\mathcal{V}_T(\breve{Y}_T)$.

<u>Step T-1:</u> The cost-to-go is given by

$$\mathcal{J}_{T-1} = E\{L_{T-1}(\underline{x}_{T-1}, \underline{u}_{T-1}) + \ell_T(\underline{v}_T) + \mathcal{L}_T^1(\hat{\underline{x}}_{T/T}) + \mathcal{L}_T^2(\underline{\Sigma}_{T/T}) | \breve{Y}_{T-1}\} \qquad (5.3.9)$$

Using Lemma 5.2.1 and assuming $\ell_T(\underline{v}_T)$ is positive-definite implies that $\underline{v}_T^* = 0$, and $\mathcal{J}_{T-1}$ may be expressed

$$\mathcal{J}_{T-1} = \mathcal{L}_{T-1}^1(\hat{\underline{x}}_{T-1/T-1}, \underline{u}_{T-1}) + \mathcal{L}_{T-1}^2(\underline{\Sigma}_{T-1/T-1}) + E\{\mathcal{L}_T^1(\hat{\underline{x}}_{T/T}) + \mathcal{L}_T^2(\underline{\Sigma}_{T/T}) | \breve{Y}_{T-1}\} \qquad (5.3.10)$$

Since no measurement is made at time $t = T$, $\hat{\underline{x}}_{T/T}$ and $\underline{\Sigma}_{T/T}$ are are obtained from $\hat{\underline{x}}_{T-1/T-1}$ and $\underline{\Sigma}_{T-1/T-1}$ by prediction. Furthermore,

$$E\{\mathcal{L}_T^1(\hat{\underline{x}}_{T/T}) + \mathcal{L}_T^2(\underline{\Sigma}_{T/T}) \mid \check{Y}_{T-1}\}$$

$$= E\{E\{L_T(\underline{x}_T) \mid \check{Y}_T\} \mid \check{Y}_{T-1}\} \qquad (5.3.11)$$

$$= E\{L_T(\underline{x}_T) \mid \check{Y}_{T-1}\} \qquad (5.3.12)$$

$$= \mathcal{L}_T^1(\underline{x}_{T/T-1}) + \mathcal{L}_T^2(\underline{\Sigma}_{T/T-1}) \qquad (5.3.13)$$

where $\mathcal{L}_T^1$ and $\mathcal{L}_T^2$ in (5.3.13) have precisely the same form as in (5.3.8). Thus

$$\mathcal{J}_{T-1} = \mathcal{L}_{T-1}^1(\hat{\underline{x}}_{T-1/T-1}, \underline{u}_{T-1}) + \mathcal{L}_{T-1}^2(\underline{\Sigma}_{T-1/T-1})$$

$$+ \mathcal{L}_T^1(\underline{A}_{T-1}\hat{\underline{x}}_{T-1/T-1} + \underline{B}_{T-1}\underline{u}_{T-1})$$

$$+ \mathcal{L}_T^2(\underline{A}_{T-1}\underline{\Sigma}_{T-1/T-1}\underline{A}'_{T-1} + \underline{\Xi}_{T-1}) \qquad (5.3.14)$$

Suppose the feedback control function which minimizes $\mathcal{J}_{T-1}$ is

$$\underline{u}^*_{T-1} = \underline{\phi}^*_{T-1}(\hat{\underline{x}}_{T-1/T-1}) \qquad (5.3.15)$$

It is clear that the optimal value of $\underline{u}_{T-1}$ depends only on $\hat{\underline{x}}_{T-1/T-1}$ since those terms which are influenced by $\underline{u}_{T-1}$ in (5.3.14) are independent of $\underline{\Sigma}_{T-1/T-1}$.

When (5.3.15) is substituted into (5.3.14) to give the optimal return function from time $T-1$, one obtains

$$\mathcal{J}_{T-1} = \mathcal{J}^{1}_{T-1}(\underline{\hat{x}}_{T-1/T-1}) + \mathcal{J}^{2}_{T-1}(\underline{\Sigma}_{T-1/T-1}) \tag{5.3.16}$$

where

$$\mathcal{J}^{1}_{T-1}(\underline{\hat{x}}_{T-1/T-1}) = \mathcal{L}^{1}_{T-1}(\underline{\hat{x}}_{T-1/T-1}, \underline{\phi}^{*}_{T-1}(\underline{\hat{x}}_{T-1/T-1}))$$
$$+ \mathcal{L}^{1}_{T}(\underline{A}_{T-1}\underline{\hat{x}}_{T-1/T-1} + \underline{B}_{T-1}\underline{\phi}^{*}_{T-1}(\underline{\hat{x}}_{T-1/T-1})) \tag{5.3.17}$$

$$\mathcal{J}^{2}_{T-1}(\underline{\Sigma}_{T-1/T-1}) = \mathcal{L}^{2}_{T-1}(\underline{\Sigma}_{T-1/T-1}) + \mathcal{L}^{2}_{T}(\underline{A}_{T-1}\underline{\Sigma}_{T-1/T-1}\underline{A}'_{T-1} + \underline{\Xi}_{T-1}) \tag{5.3.18}$$

<u>Step T-2:</u> The cost-to-go from step T-2 is

$$\mathcal{J}_{T-2} = E\{L_{T-2}(\underline{x}_{T-2}, \underline{u}_{T-2}) + \ell_{T-1}(\underline{v}_{T-1}) + \mathcal{J}_{T-1}(\breve{Y}_{T-1}) \,|\, \breve{Y}_{T-2}\} \tag{5.3.19}$$

$$= \mathcal{L}^{1}_{T-2}(\underline{\hat{x}}_{T-2/T-2}, \underline{u}_{T-2}) + \ell_{T-1}(\underline{v}_{T-1}) + \mathcal{L}^{2}_{T-2}(\underline{\Sigma}_{T-2/T-2})$$
$$+ E\{\mathcal{J}^{1}_{T-1}(\underline{\hat{x}}_{T-1/T-1}) + \mathcal{J}^{2}_{T-1}(\underline{\Sigma}_{T-1/T-1}) \,|\, \breve{Y}_{T-2}\} \tag{5.3.20}$$

Now $\underline{\hat{x}}_{T-1/T-1}$ and $\underline{\Sigma}_{T-1/T-1}$ are generated from $\underline{\hat{x}}_{T-2/T-2}$ and $\underline{\Sigma}_{T-2/T-2}$ by the usual Kalman filter equations:

$$\underline{\hat{x}}_{T-1/T-2} = \underline{A}_{T-2}\underline{\hat{x}}_{T-2/T-2} + \underline{B}_{T-2}\underline{u}_{T-2} \tag{5.3.21}$$

$$\hat{\underline{x}}_{T-1/T-1} = \hat{\underline{x}}_{T-1/T-2} + S_{T-1}[\underline{y}_{T-1} - \underline{C}_{T-1}\hat{\underline{x}}_{T-1/T-2}] \quad (5.3.22)$$

$$\underline{\Sigma}_{T-1/T-2} = \underline{A}_{T-2}\underline{\Sigma}_{T-2/T-2}\underline{A}'_{T-2} + \underline{\Xi}_{T-2} \quad (5.3.23)$$

$$\underline{\Sigma}_{T-1/T-1} = \underline{\Sigma}_{T-1/T-2} - \underline{S}_{T-1}[\underline{C}_{T-1}\underline{\Sigma}_{T-1/T-2}\underline{C}'_{T-1} + \underline{D}_{T-1}\underline{\Theta}_{T-1}D'_{T-1}]\underline{S}'_{T-1} \quad (5.3.24)$$

$$\underline{S}_{T-1} = \underline{\Sigma}_{T-1/T-2}\underline{C}'_{T-1}[\underline{C}_{T-1}\underline{\Sigma}_{T-1/T-2}\underline{C}'_{T-1} + \underline{D}_{T-1}\underline{\Theta}_{T-1}\underline{D}'_{T-1}]^{-1} \quad (5.3.25)$$

Recall that $\underline{C}_{T-1}$ and $\underline{D}_{T-1}$ depend on $\underline{v}_{T-1}$. Thus the term $E\{\mathcal{J}^2_{T-1}(\underline{\Sigma}_{T-1/T-1}) | \tilde{Y}_{T-2}\}$ in (5.3.20) may be expressed in terms of $\underline{\Sigma}_{T-2/T-2}$ and $\underline{v}_{T-1}$ only, but the term $E\{\mathcal{J}^1_{T-1}(\hat{\underline{x}}_{T-1/T-1}) | \tilde{Y}_{T-2}\}$ <u>in general depends on</u> $\hat{\underline{x}}_{T-2/T-2}$, $\underline{u}_{T-2}$, $\underline{v}_{T-2}$ <u>and</u> $\underline{\Sigma}_{T-2/T-2}$. To see exactly how this results, note that the expected value in (5.3.20) is over the random variables $\underline{\theta}_{T-1}$ and $(\underline{x}_{T-1} - \hat{\underline{x}}_{T-1/T-2})$, as can be seen from the expression

$$E\{\mathcal{J}^1_{T-1}(\hat{\underline{x}}_{T-1/T-1}) | \tilde{Y}_{T-2}\}$$
$$= E\{\mathcal{J}^1_{T-1}(\underline{A}_{T-2}\hat{\underline{x}}_{T-2/T-2} + \underline{B}_{T-2}\underline{u}_{T-2} + \underline{S}_{T-1}[\underline{D}_{T-1}\underline{\theta}_{T-1} + \underline{C}_{T-1}(\underline{x}_{T-1} - \hat{\underline{x}}_{T-1/T-2})]) | \tilde{Y}_{T-2}\} \quad (5.3.26)$$

The random variables $\underline{\theta}_{T-1}$ and $(\underline{x}_{T-1} - \hat{\underline{x}}_{T-1/T-2})$ are both zero mean (given $\breve{Y}_{T-2}$) with covariance matrices $\underline{\Theta}_{T-1}$ and $\underline{\Sigma}_{T-1/T-2}$ respectively. Thus, these statistics do not depend on $\underline{v}_{T-1}$ or $\underline{u}_{T-2}$, but because the general nature of $\underline{\phi}^*_{T-1}$ produces a general nature for $\mathcal{J}^1_{T-1}$ (see (5.3.17)), and since $\underline{C}_{T-1}$, $\underline{D}_{T-1}$ and $\underline{S}_{T-1}$ depend on $\underline{v}_{T-1}$, the form of (5.3.26) is

$$E\{\mathcal{J}^1_{T-1}(\hat{\underline{x}}_{T-1/T-1}) | \breve{Y}_{T-2}\} \triangleq \hat{\mathcal{J}}^1_{T-1}(\hat{\underline{x}}_{T-2/T-2}, \underline{\Sigma}_{T-2/T-2}, \underline{u}_{T-2}, \underline{v}_{T-1}) \tag{5.3.27}$$

If one defines*

$$E\{\mathcal{J}^2_{T-1}(\underline{\Sigma}_{T-1/T-1}) | \breve{Y}_{T-2}\} \triangleq \hat{\mathcal{J}}^2_{T-1}(\underline{\Sigma}_{T-2/T-2}, \underline{v}_{T-1}) \tag{5.3.28}$$

one can write, using (5.3.20)

$$\begin{aligned} \mathcal{J}_{T-2} = {} & \mathcal{L}^1_{T-2}(\underline{x}_{T-2/T-2}, \underline{u}_{T-2}) + \ell_{T-1}(\underline{v}_{T-1}) + \mathcal{L}^2_{T-2}(\underline{\Sigma}_{T-2/T-2}) \\ & + \hat{\mathcal{J}}^1_{T-1}(\underline{x}_{T-2/T-2}, \underline{\Sigma}_{T-2/T-2}, \underline{u}_{T-2}, \underline{v}_{T-1}) \\ & + \hat{\mathcal{J}}^2_{T-1}(\underline{\Sigma}_{T-2/T-2}, \underline{v}_{T-1}) \end{aligned} \tag{5.3.29}$$

---

* $\mathcal{J}^2_{T-1}(\Sigma_{T-1/T-1})$ is actually a deterministic quantity but its expected value is the form which appears in $\mathcal{J}_{T-2}$ in (5.3.20).

Because of the coupling of $\hat{\underline{x}}_{T-2/T-2}$, $\underline{\Sigma}_{T-2/T-2}$, $\underline{u}_{T-2}$, and $\underline{v}_{T-1}$ in the fourth term of (5.3.29), it follows that the general optimal feedback functions must take the form

$$\underline{u}^*_{T-2} = \underline{\phi}^*_{T-2}(\hat{\underline{x}}_{T-2/T-2}, \underline{\Sigma}_{T-2/T-2}) \tag{5.3.30}$$

$$\underline{v}^*_{T-1} = \underline{\psi}^*_{T-1}(\hat{\underline{x}}_{T-2/T-2}, \underline{\Sigma}_{T-2/T-2}) \tag{5.3.31}$$

Equation (5.3.31) <u>disproves</u> the conjecture. The fact that (5.3.5) holds does <u>not</u> guarantee (i.e., is not sufficient for) the optimality of the <u>a-priori v technique</u>.

<u>Remark 3:</u> Note that the difficulty in showing that the optimal $\{\underline{v}_t\}$ depends only on $\{\underline{\Sigma}_{t/t}\}$ under the hypothesis of Conjecture 5.3.1 results from the general nature of the <u>dynamic control feedback function</u> $\underline{\phi}^*_{T-1}$. Since the structure of this function is general, so is $\mathcal{J}^1_{T-1}$ (see (5.3.17)), and this prevents the separation of $\underline{v}_{T-1}$ from $\hat{\underline{x}}_{T-2/T-2}$ at step $t = T-2$. This situation will be further clarified in Section 5.4, in which the transformation approach and the Minimum Principle will be applied to the study of Conjecture 5.3.1.

<u>Proof of Sufficiency Part of Conjecture 5.3.1</u>

The hypotheses of Conjecture 5.3.1 are so strong that the sufficiency part of the statement is nearly self-evident, while it is not totally clear that these restrictive conditions are necessary for the conclusion to hold. To see the sufficiency, note that at step t of a dynamic programming solution for the optimal controls, the cost-to-go is

$$\mathcal{J}_t = E\{L_t(\underline{x}_t, \underline{u}_t) + \ell_{t+1}(\underline{v}_{t+1}) + \mathcal{V}_{t+1} | \tilde{Y}_t\} \tag{5.3.32}$$

which by the conjecture hypothesis may be written as follows:

$$\mathcal{J}_t = \mathcal{L}^1_t(\hat{\underline{x}}_{t/t}, \underline{u}_t) + \mathcal{L}^2_t(\underline{\Sigma}_{t/t}) + \ell_{t+1}(\underline{v}_{t+1}) + \hat{\mathcal{V}}^1_{t+1}(\underline{x}_{t/t}, \underline{u}_t) + \hat{\mathcal{V}}^2_{t+1}(\underline{\Sigma}_{t/t}, \underline{v}_{t+1}) \tag{5.3.33}$$

Since the terms in $\mathcal{J}_t$ which depend upon $\underline{u}_t$ are the first and fourth, it is clear that the optimal $\underline{u}_t$ may be expressed

$$\underline{u}^*_t = \underline{\phi}^*_t(\hat{\underline{x}}_{t/t}) \tag{5.3.34}$$

Since the terms in $\mathcal{J}_t$ which depend on $\underline{v}_{t+1}$ are the third and the last in (5.3.33), the optimal $\underline{v}_{t+1}$ may be expressed

$$\underline{v}^*_{t+1} = \underline{\psi}^*_{t+1}(\underline{\Sigma}_{t/t}) \tag{5.3.35}$$

Since the covariance $\underline{\Sigma}_{t/t}$ satisfies a deterministic equation of motion and may be calculated a-priori, the optimal measurement sequence $\{\underline{v}^*_t\}$ may be calculated a-priori. This shows the sufficiency of the hypotheses of Conjecture 5.3.2.

Remark 4: Although the hypothesis that (5.3.6) holds for all t makes it unnecessary to consider what occurs when (5.3.34) and (3.3.35) are substituted back to $\mathcal{J}_t$, consider what happens if this is done. The value of $\mathcal{V}_t$ is then

$$\mathcal{J}_t = \mathcal{L}_t^1(\hat{\underline{x}}_{t/t}, \underline{\phi}_t^*(\hat{x}_{t/t})) + \mathcal{L}_t^2(\underline{\Sigma}_{t/t})$$
$$+ \ell_{t+1}(\underline{\psi}_{t+1}^*(\underline{\Sigma}_{t/t})) + \hat{\mathcal{J}}_{t+1}^1(\hat{\underline{x}}_{t/t}, \underline{\phi}_t^*(\hat{\underline{x}}_{t/t}))$$
$$+ \mathcal{J}_{t+1}^2(\underline{\Sigma}_{t/t}, \underline{\psi}_{t+1}^*(\underline{\Sigma}_{t/t})) \qquad (5.3.36)$$

As in the investigation of the sufficiency part of Conjecture 5.3.1, the terms that "cause trouble" at time step t-1 are

$$\text{terms} = E\{\mathcal{L}_t^1(\hat{\underline{x}}_{t/t}, \underline{\phi}_t^*(\underline{x}_{t/t})) + \hat{\mathcal{J}}_{t+1}^1(\hat{\underline{x}}_{t/t}, \underline{\phi}_t^*(\hat{\underline{x}}_{t/t})) | \tilde{Y}_{t-1}\} \qquad (5.3.37)$$

By hypothesis in Conjecture 5.3.2, these terms are not troublesome because they evaluate to a mean part plus a covariance part. One can see, however, that whether or not the hypothesis (2) of Conjecture 5.3.2 is satisfied at time t-1 given that it is satisfied at time t depends only on the nature of the $\underline{\phi}_t^*(\cdot)$, the $\mathcal{L}_t^1(\cdot,\cdot)$ and the $\hat{\mathcal{J}}_{t+1}^1(\cdot,\cdot)$ functions. In particular, the $\underline{\psi}_{t+1}^*(\cdot)$ function specifying the optimal measurement control plays no role, nor do the $\mathcal{L}_t^2(\cdot)$ or $\hat{\mathcal{J}}_{t+1}^2(\cdot)$ functions, which display the influence of the covariance on the problem. This situation will be studied from a different point of view in the next section, using deterministic techniques, as was mentioned in Remark 3 above.

Remark 5: Although the hypotheses of Conjecture 5.3.2 are so strong as to perhaps rule out any cost functionals but quadratic,* a slightly different formulation suggests itself. Up to now, the $\mathcal{L}_t^1$ and $\mathcal{L}_t^2$ functions have been specified by taking the expected value of $L_t(\underline{x}_t, \underline{u}_t)$ terms in costs. (See (5.3.5)). Thus $\mathcal{L}_t^1$ and $\mathcal{L}_t^2$ are not completely arbitrary. What would happen if the original problem formulation started not with a cost of the type given in (5.3.2), but with a cost made up of general $\mathcal{L}_t^1$ and $\mathcal{L}_t^2$ terms depending on $\hat{\underline{x}}_{t/t}$, $\underline{u}_t$, and $\underline{\Sigma}_{t/t}$? It is not clear what the effect on overall system behavior would be, or what analytic benefits would result. This is a good area for future work.

Remark 6: It was mentioned after equation (5.3.3) that the most general cost-term-form $L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1})$ would not be considered, but rather the more restrictive form $L_t(\underline{x}_t, \underline{u}_t) + \ell_{t+1}(\underline{v}_{t+1})$. It is now obvious that this was done because the more general form would typically prevent the $\underline{v}_{t+1}$-optimization from being independent of $\hat{\underline{x}}_{t/t}$, which is the result that was being sought.

Remark 7: It has been shown that the conditions of Conjecture 5.3.1 are not sufficient to guarantee that the a-priori $\underline{v}$ technique is optimal, while those of Conjecture 5.3.2 are sufficient. Nothing has been said

---

* No others have yet been found, although no proof that none exist is yet available.

about necessity. It is obvious, however, from the analysis of the sufficiency conditions that the separation of $E\{L_t(\underline{x}_t, \underline{u}_t) | \breve{Y}_t\}$ into two terms $\mathcal{L}_t^1(\hat{\underline{x}}_{t/t}, \underline{u}_t) + \mathcal{L}_t^2(\underline{\Sigma}_{t/t})$, while itself not sufficient, is indeed necessary. If this separation does not occur at time t, then it is clear that the optimal $\underline{v}_t$ determined at step $t-1$ will in general depend on $\hat{\underline{x}}_{t-1|t-1}$.

5.4 An Analysis Using the Transformation Approach

In this section, the effect of the hypothesis of Conjecture 5.2.1 as applied to a linear-Gaussian-non-quadratic control problem will be studied using the deterministic approach of Chapter III. This will serve to clarify the reasons why that conjecture was not true, and as a subsidiary result, the nature of linear-Gaussian-non-quadratic control problems without measurement control capability will be clarified as well. The structure of this section is very similar to that of Section 4.3.

5.4.1 Transformation of the System Equations

Since the form of the cost functional has no effect on the filter which is used to estimate the state, the equations of motion for the transformed system with non-quadratic cost are identical to those of Section 4.3. As in that section, the t/t subscript will be suppressed to give the following system:

$$\hat{\underline{x}}_{t+1} - \hat{\underline{x}}_t = (\underline{A}_t - \underline{I})\hat{\underline{x}}_t + \underline{B}_t\underline{u}_t + \hat{\underline{\xi}}_t \qquad (5.4.1)$$

$$\underline{\Sigma}_{t+1} - \underline{\Sigma}_t = \underline{A}_t\underline{\Sigma}_t\underline{A}'_t - \underline{\Sigma}_t + \underline{\Xi}_t$$

$$-(\underline{A}_t\underline{\Sigma}_t\underline{A}'_t + \underline{\Xi}_t)\underline{C}'_{t+1}[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_t\underline{A}'_t + \underline{\Xi}_t)\underline{C}'_{t+1} + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1}]^{-1}\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_t\underline{A}'_t + \underline{\Xi}_t)$$

(5.4.2)

The initial values are

$$\hat{\underline{x}}_{0/0} = \text{Initial value of } \hat{\underline{x}}_t \qquad (5.4.3)$$

$$= \hat{\underline{x}}_0 + \underline{S}_0[\underline{y}_0 - \underline{C}_0 v_0 \hat{\underline{x}}_0] \qquad (5.4.4)$$

$$= \hat{\underline{x}}_0 + \hat{\underline{\xi}}_{-1} \qquad (5.4.5)$$

$$\underline{\Sigma}_{0/0} = \text{Initial value of } \underline{\Sigma}_t$$

$$= \underline{\Sigma}_0 - \underline{\Sigma}_0\underline{C}'_0(\underline{C}_0\underline{\Sigma}_0\underline{C}'_0 + \underline{D}_0\underline{\Theta}_0\underline{D}'_0)\underline{C}_0\underline{\Sigma}'_0 \qquad (5.4.6)$$

where, as in Section 4.3, the 0/0 subscript is retained for the initial values and a single zero subscript refers to the a-priori mean and covariance. Recall that $\hat{\underline{\xi}}_t$ is given by (see equations (4.3.6)-(4.3.8)

$$\hat{\underline{\xi}}_t = \underline{S}_{t+1}[\underline{y}_{t+1} - \underline{C}_{t+1}(\underline{A}_t\hat{\underline{x}}_t + \underline{B}_t\underline{u}_t)] \qquad (5.4.7)$$

where $\{\hat{\underline{\xi}}_t\}$ is a white noise process with covariance

$$\text{cov}\ (\hat{\underline{\xi}}_t, \hat{\underline{\xi}}_t) =$$

$$\underline{S}_{t+1}[\underline{C}_{t+1}(\underline{A}_t\underline{\Sigma}_t\underline{A}'_t + \underline{\Xi}'_t)\underline{C}'_{t+1} + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}'_{t+1}]\underline{S}'_{t+1} \qquad (5.4.8)$$

with

$$\underline{S}_{t+1} = (\underline{A}_t\underline{\Sigma}_t\underline{A}_t' + \underline{\Xi}_t)\underline{C}_t'[\underline{C}_t(\underline{A}_t\underline{\Sigma}_t\underline{A}_t' + \underline{\Xi}_t)\underline{C}_t' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']^{-1} \quad (5.4.9)$$

Equations (5.4.1) - (5.4.9) come directly from Section 4.3. As in that section, it will sometimes be convenient to abbreviate the equations of motion as

$$\hat{\underline{x}}_{t+1} - \hat{\underline{x}}_t = \underline{f}_t(\hat{\underline{x}}_t, \underline{u}_t, \underline{\xi}_t) \quad (5.4.10)$$

$$\underline{\Sigma}_{t+1} - \underline{\Sigma}_t = \underline{F}_t(\underline{\Sigma}_t, \underline{v}_{t+1}) \quad (5.4.11)$$

where $\underline{f}_t$ and $\underline{F}_t$ are defined as the right-hand-sides of (5.4.1) and (5.4.2) respectively.

To transform the cost functional (5.3.3), let

$$\mathcal{L}_t(\hat{\underline{x}}_t, \underline{\Sigma}_t, \underline{u}_t) = E\{L_t(\underline{x}_t, \underline{u}_t) \mid \underline{x}_t = N(\hat{\underline{x}}_t, \underline{\Sigma}_t)\} \quad (5.4.12)$$

Then according to Theorem 3.3.1, the cost to be associated with the transformed system is

$$J = \mathcal{L}_T(\hat{\underline{x}}_T, \underline{\Sigma}_T) + \sum_{t=0}^{T=1} \mathcal{L}_t(\hat{\underline{x}}_t, \underline{\Sigma}_t, \underline{u}_t) + \sum_{t=0}^{T} \ell_t(\underline{v}_t) \quad (5.4.13)$$

Since this section is investigating Conjecture 5.3.1, the following assumption will also be made:

$$\mathcal{L}_t(\hat{\underline{x}}_t, \underline{\Sigma}_t, \underline{u}_t) = \mathcal{L}_t^1(\hat{\underline{x}}_t, \underline{u}_t) + \mathcal{L}_t^2(\underline{\Sigma}_t) \quad (5.4.14)$$

5.4.2 Application of the Minimum Principle

Let the innovations sequence $\{\underline{\hat{\xi}}_t\}$ be considered fixed. Let the Hamiltonian function be defined as follows:

$$H(\underline{\hat{x}}_t, \underline{\Sigma}_t, \underline{p}_{t+1}, \underline{P}_{t+1}, \underline{u}_t, \underline{v}_{t+1}) =$$

$$\mathcal{L}^1_t(\underline{\hat{x}}_t, \underline{u}_t) + \mathcal{L}^2_t(\underline{\Sigma}_t) + \ell_{t+1}(\underline{v}_{t+1})$$

$$+ \underline{p}'_{t+1}[(\underline{A}_t - \underline{I})\underline{\hat{x}}_t + \underline{B}_t\underline{u}_t + \underline{\hat{\xi}}_t]$$

$$+ \operatorname{tr}\left[\underline{P}'_{t+1}\underline{F}_t(\underline{\Sigma}_t, \underline{v}_{t+1})\right] \tag{5.4.15}$$

where, as in Section 4.3, $\underline{p}_{t+1}$ and $\underline{P}_{t+1}$ are an n-vector and an nxn matrix respectively which serve as the costates for the optimal control problem. Then according to the Minimum Principle, if $\{\underline{u}^*_t\}$ and $\{\underline{v}^*_t\}$ are optimal controls and if $\{\underline{\hat{x}}^*_t\}$, $\{\underline{\Sigma}^*_t\}$, $\{\underline{p}^*_t\}$, and $\{\underline{P}^*_t\}$ are the resulting optimal trajectories, then the following conditions hold:

I. Canonical Equations

$$\underline{\hat{x}}^*_{t+1} - \underline{\hat{x}}^*_t = (\underline{A}_t - \underline{I})\underline{\hat{x}}^*_t + \underline{B}_t \underline{u}^*_t + \underline{\hat{\xi}}_t \tag{5.4.16}$$

$$\underline{p}^*_{t+1} - \underline{p}^*_t = -\frac{\partial H}{\partial \underline{\hat{x}}_t} = -(\underline{A}_t - \underline{I})' \underline{p}_{t+1} + \frac{\partial}{\partial \underline{x}_t} \mathcal{L}^1_t(\underline{\hat{x}}^*_t, \underline{u}^*_t) \tag{5.4.17}$$

$$\underline{\Sigma}^*_{t+1} - \underline{\Sigma}^*_t = \underline{F}_t(\underline{\Sigma}^*_t, \underline{v}^*_{t+1}) \tag{5.4.18}$$

$$\underline{P}^*_{t+1} - \underline{P}^*_t = -\frac{\partial H}{\partial \underline{\Sigma}} \tag{5.4.19}$$

II. Minimization of the Hamiltonian

$$H(\underline{\hat{x}}^*_t, \underline{\Sigma}^*_t, \underline{p}^*_{t+1}, \underline{P}_{t+1}, \underline{u}^*_t, \underline{v}^*_{t+1}) \leq H(\underline{\hat{x}}^*_t, \underline{\Sigma}^*_t, \underline{p}^*_{t+1}, \underline{P}^*_{t+1}, \underline{u}_t, \underline{v}_{t+1}) \tag{5.4.20}$$

III. Transversality Condition:

$$\underline{p}_T = \frac{\partial \mathcal{L}^1_T(\underline{\hat{x}}_T)}{\partial \underline{\hat{x}}_T} \tag{5.4.21}$$

$$\underline{P}_T = \frac{\partial \mathcal{L}^2_T(\underline{\Sigma}_T)}{\partial \underline{\Sigma}_T} \tag{5.4.22}$$

4

Assume the Hamiltonian is a differentiable function of $\underline{u}_t$ and $\underline{v}_{t+1}$, so that equation (5.4.20) can be expressed

$$\frac{\partial H}{\partial \underline{u}} = \underline{0} = \frac{\partial \mathcal{L}_t^1(\hat{\underline{x}}_t^*, \underline{u}_t^*)}{\partial \underline{u}_t} + \underline{B}_t' \underline{p}_{t+1}^* \qquad (5.4.23)$$

$$\frac{\partial H}{\partial \underline{v}_{t+1}} = 0 = \text{expression involving } \underline{P}_{t+1}^*, \underline{\Sigma}_t^*, \underline{v}_{t+1}^* \qquad (5.4.24)$$

Assuming that these stationarity conditions may be shown to yield a (unique) pair of minimizing values for the controls, it is clear that the functional forms are

$$\underline{u}_t^* = \underline{\phi}_t^*(\hat{\underline{x}}_t^*, \underline{p}_{t+1}^*) \qquad (5.4.25)$$

$$\underline{v}_{t+1}^* = \underline{\psi}_{t+1}^*(\underline{\Sigma}_t^*, \underline{P}_{t+1}^*) \qquad (5.4.26)$$

When these are substituted into (5.4.16) - (5.4.19), it is clear that two separate two-point boundary value problems will result — one involving $\{\hat{\underline{x}}_t\}$ and $\{\underline{p}_t\}$ and one involving $\{\underline{\Sigma}_t\}$ and $\{\underline{P}_t\}$.

This, however, does not prove that a complete two-way separation holds and that the a-priori $\underline{v}$ technique is optimal. Recall that in Chapter IV, the separation of the two-point boundary value problems was not sufficient. One had to take the two-point boundary value problem in $\hat{\underline{x}}_t$ and $\underline{p}_t$ and convert it to a feedback form to obtain

$$\underline{u}_t^* = \underline{u}_t^*(\hat{\underline{x}}_t, \hat{\underline{\xi}}_t, \ldots, \hat{\underline{\xi}}_{T-1}) \qquad (5.4.27)$$

then take the expected value of this control over the noises to obtain

$$\bar{\underline{u}}_t^* = E\{\underline{u}_t^*(\hat{\underline{x}}_t, \hat{\underline{\xi}}_t, \ldots, \hat{\underline{\xi}}_{T-1}) | \hat{\underline{x}}_t\} \tag{5.4.28}$$

and finally evaluate the cost-to-go from time $t = 0$ using $\{\bar{\underline{u}}_t^*\}$, viz., $\underline{I}_0(\hat{\underline{x}}_t, \hat{\underline{\xi}}_t, \ldots, \hat{\underline{\xi}}_{T-1})$ in order to determine what the stochastic effects were. In taking the expected value of $\underline{I}_0(\hat{\underline{x}}_0, \hat{\underline{\xi}}_0, \ldots, \hat{\underline{\xi}}_{T-1})$ to determine the stochastic cost, one found that two properties were true in the quadratic cost case: $\bar{\underline{u}}_t^*$ did not involve $\{\underline{v}_t^*\}$ or $\{\underline{\Sigma}_t^*\}$ and the expected cost $\underline{I}_0(\hat{\underline{x}}_0) = E\{\underline{I}_0(\hat{\underline{x}}_0, \hat{\underline{\xi}}_0, \ldots, \hat{\underline{\xi}}_{T-1}) | \hat{\underline{x}}_0\}$ had a term in $\hat{\underline{x}}_0$ adding to terms in $\{\underline{\Sigma}_t^*\}$. These facts showed the separation of the $\underline{v}$-problem.

Of course, for the general problem being considered here, such a procedure cannot be carried out analytically. One can not make general statements about the nature of $\underline{\phi}_t^*(\cdot, \cdot, \ldots, \cdot)$ in (5.4.27) or about its expected value over the noises or about the cost-to-go using that expected value for a control. As was found to be the case in the previous section, in which the analysis was carried out by dynamic programming, the fact that equation (5.4.14) holds is not, in general, sufficient to guarantee that the a-priori $\underline{v}$ technique is optimal, and the reason why it is not sufficient is the general nature of the optimal dynamic control feedback function.

5.4.3 Application to Problems Without Measurement Control

An important aspect of the deterministic approach to a stochastic problem is that the full stochastic effect is not seen until the two expected values discussed in the Remark of subsection 4.3.1 are taken: the expected value of the deterministic optimal dynamic feedback control law and the expected value of the resulting cost. If the problem under consideration contains linear dynamics and a fixed linear measurement equation (no measurement control), then it is not necessary to go so far as to take both expected values. If one does not need to know the stochastic cost in order to determine the full effect of the measurement control, one can stop with taking the expected value of the deterministic optimal dynamic feedback control. Thus, for a problem with only dynamic control, nearly the entire analysis can be carried out using deterministic methods: everything except an expected value at the beginning (equation (5.4.12) to transform the cost) and an expected value at the end (to average the future innovations out of the deterministic control law).

It has already been pointed out in this thesis* that a separation theorem is known to hold for "ordinary" (i.e., without measurement control) linear-Gaussian-non-quadratic control problems. The result is that the optimal dynamic control is obtained by plugging the state estimate $\hat{\underline{x}}_{t|t}$ into the control law derived as the solution of a different optimal control problem than the original one with the noises set to zero. Using the

---

* See Section 4.1 and Wonham (W.2).

transformation approach, one can see precisely what this different problem is: precisely equations (5.4.1), (5.4.2), and (5.4.13).

Examination of these equations shows that there are basically two reasons why the problem one solves for the optimal dynamic control is different than the original problem with the noises discarded. The original dynamic equation with the noise discarded becomes

$$\underline{x}_{t+1} = \underline{A}_t \underline{x}_t + \underline{B}_t \underline{u}_t \tag{5.4.29}$$

while the equation that one must consider is

$$\hat{\underline{x}}_{t+1} = \underline{A}_t \hat{\underline{x}}_t + \underline{B}_t \underline{u}_t + \hat{\underline{\xi}}_t \tag{5.4.30}$$

The original cost equation is

$$J = L_t(\underline{x}_T) + \sum_{t=0}^{T-1} L_t(\underline{x}_t, \underline{u}_t) \tag{5.4.31}$$

while the transformed cost that one must consider is

$$\hat{J} = \mathcal{L}_T(\underline{x}_T, \underline{\Sigma}_T) + \sum_{t=0}^{T-1} \mathcal{L}_t(\hat{\underline{x}}_t, \underline{\Sigma}_t, \underline{u}_t) \tag{5.4.32}$$

The differences between "the original problem without noise", (5.4.29) and (5.4.31), and the actual problem one must consider using deterministic techniques (5.4.30) and (5.4.32), are the influence of $\hat{\underline{\xi}}_t$ in the equation of motion and the influence of $\underline{\Sigma}_t$ (which, by the way, is now a <u>completely precomputable</u> sequence of matrices satisfying (5.4.2)) in the cost equation. In addition, the covariance $\underline{\Sigma}_t$ will typically influence the final result further when the expected value indicated in (5.4.28) is taken.

What characteristics of the linear-Gaussian-quadratic problem result in its control law being precisely the same as that for the "same problem without the noise"? First, equation (5.3.4), which gives the result that

$$\mathscr{L}_t(\hat{\underline{x}}_t, \underline{\Sigma}_t, \underline{u}_t)$$

$$= E\{L_t(\underline{x}_t, \underline{u}_t) | \underline{x}_t = N(\hat{\underline{x}}_t, \underline{\Sigma}_t)\} \tag{5.4.33}$$

$$= E\{\underline{x}_t' \underline{Q}_t \underline{x}_t + \underline{u}_t \underline{R}_t \underline{u}_t | \underline{x}_t = N(\hat{\underline{x}}_t, \underline{\Sigma}_t)\} \tag{5.4.34}$$

$$= \hat{\underline{x}}_t' \underline{Q}_t \hat{\underline{x}}_t + \underline{u}_t' \underline{R}_t \underline{u}_t + \mathrm{tr}\,[\underline{Q}_t \underline{\Sigma}_t] \tag{5.4.35}$$

$$= L_t(\hat{\underline{x}}_t, \underline{u}_t) + \mathrm{tr}\,[\underline{Q}_t \underline{\Sigma}_t] \tag{5.4.36}$$

Thus, for the quadratic problem, the transformed cost is identical in form to the original cost, with the addition of a term independent of the dynamic control. Second, as was derived in subsection 4.3.3, in the quadratic cost case, the optimal deterministic dynamic control is linear in the future innovations, so that its expected value is independent of future innovations. These two special circumstances do not in general hold for non-quadratic-cost problems.

5.5 Summary

In this chapter, linear systems corrupted by Gaussian noise and penalized by general (i.e., non-quadratic) cost functionals were considered. It was shown both by example and by an analytic argument that the optimal measurement control at time t can be a function of the plant state estimate at time $t-1$. As a result, it may be impossible to compute the optimal measurement control time sequence off-line.

Both dynamic programming and the Minimum Principle were used in the analysis. By comparing the general cost case to the quadratic case, several conjectures were made concerning conditions which would be necessary and sufficient to allow one to precompute the optimal measurement control. Sufficient conditions were found, but necessary and sufficient conditions in a convenient form for applications have not yet been derived.

As a byproduct of the analysis of combined dynamic control/measurement control problems, the nature of "ordinary" dynamic control problems with no measurement control capability was clarified.

CHAPTER VI

A PROBLEM WITH COUPLED MEASUREMENT AND CONTROL

6.1 General Perspective

In this chapter, a particular class of problems will be introduced in which the dynamic control and state measurement functions are inextricably coupled. Although problems of this type have not been previously considered, they are of great practical importance, and this importance will increase as control technology advances. This is because control structures involving a large, powerful central computer which controls many remote dynamic systems fall into the class of problems to be described in this chapter. As control algorithms and techniques become more and more complex, structures of this type will inevitably become common.

Consider a situation in which the dynamic system being controlled is at some distance from the location at which the control commands are generated. Evidently, in order to transmit a control command or make a measurement of the state of the dynamic system, a communications link must be established. Frequently, if the communications facilities are limited or time shared, one might decide to conserve effort by insisting that measurements can only be taken while control commands are being sent, or vice-versa. In other words, communications will only be established when it is advantageous to both measure and to control.

As a practical example, consider a spacecraft being guided by a ground based facility on the basis of tracking information obtained from a radar. When the radar pulse is sent out for tracking, it is shaped in such a way as to also convey guidance information to the spacecraft. Thus a single transmitter is employed for both tracking and guidance. The mathematical description of such a system must include a provision that control commands will be received at some time t if and only if a measurement is made at time t.

What does the system do if no control command is received at a given time? One assumes that there is a relatively simple device on board the spacecraft which receives the guidance signals from the ground and translates them into appropriate signals sent to reaction jets, control surfaces etc. This on-board device remains in control of the spacecraft if no guidance command is received from the ground. Thus, a whole hierarchy of overall control structures can be represented in this form, depending on how sophisticated the on-board guidance system is. For example, if the on-board device is a very simple computer, it might merely store and continue executing the most recent control command until a new one is received. If the on-board guidance computer is more complex, it might be capable of determining the spacecraft state itself and then function as the major control component with only parameter updating ("gain tuning") from the ground.

In addition to these considerations, slightly different problems will result from the adoption of different stochastic structures that may be justified intuitively. Specifically, one might assume that while the ground makes noisy measurements of the spacecraft state, the trans-

mission of control commands from ground to on-board is noise-free.*
Alternatively, these transmissions may be noisy. Either structure might be appropriate, depending on the sophistication of the communications link.

Notivated by such problems, a relatively simple mathematical model will be analyzed below; the method of extension to more complicated situations should be evident.

6.2 Mathematical System Description

Let the dynamic system be represented by

$$\underline{x}_{t+1} = \underline{A}_t \underline{x}_t + \underline{B}_t \underline{w}_t + \underline{\xi}_t \; ; \qquad t = 0, 1, \ldots, T \tag{6.2.1}$$

where all the variables have the same meaning as in Section 3.2, and $\underline{w}_t \in R^m$ is the output of the on-board computer. Suppose that measurements are made according to

$$\underline{y}_t = v_t[\underline{C}_t \underline{x}_t + \underline{\theta}_t] \; ; \qquad v_t \in \{0, 1\} \tag{6.2.2}$$

(identical to (3.2.10)). The on-board computer obeys the difference equation

$$\underline{w}_{t+1} = (1 - v_t)\underline{w}_t + v_t \underline{u}_t \tag{6.2.3}$$

Thus, if no measurement is made at time t ($v_t = 0$), the computer simply remembers its old state ($\underline{w}_{t+1} = \underline{w}_t$) and if a measurement is made

---

* This is not to say that the dynamic system is free of driving noise, but only that the control command is received exactly as sent. Such an assumption can be justified by the fact that the ground-based transmitter can be made very powerful.

($v_t = 1$), the new value of control, $\underline{u}_t$, is "loaded" into the computer ($\underline{w}_{t+1} = \underline{u}_t$). A diagram of this structure is shown in Figure 6.1.

Because of the special structure of this system, the timing considerations are somewhat different than those assumed in previous chapters. In particular, since the control command $\underline{u}_t$ is already being sent simultaneously with the tracking pulse associated with v_t, it is clear that $\underline{u}_t$ cannot depend on $\underline{y}_t$ since $\underline{y}_t$ has not yet been received at the time when $\underline{u}_t$ is being computed. Thus, the type of dynamic control and measurement control dependcies that will be considered in this chapter are

$$\underline{u}_t = \underline{u}_t(Y_{t-1}) \qquad (6.2.4)$$

$$v_t = v_t(Y_{t-1}) \qquad (6.2.5)$$

Compare these to (1.3.5) and (1.3.6).

Examination of equations (6.2.1) and (6.2.3) will show that when a control command $\underline{u}_t$ is sent to the spacecraft, it does not get applied to the dynamics (equation (6.2.1)) until time t+1 since there is a unit delay in processing it through the on-board computer. This is to be expected in a practical system, however. If a control command $\underline{u}_t$ is just being received at time t, it obviously cannot be influencing the dynamics at time t. There is always a delay for reception and decoding of the message.

An appropriate cost functional for this system is

$$J = \underline{x}_T' \underline{Q}_T \underline{x}_T + \ell_T v_T + \sum_{t=0}^{T-1} \underline{x}_t' \underline{Q}_t \underline{x}_t + \underline{u}_t' \underline{R}_t \underline{u}_t + \ell_t \cdot v_t \qquad (6.2.6)$$

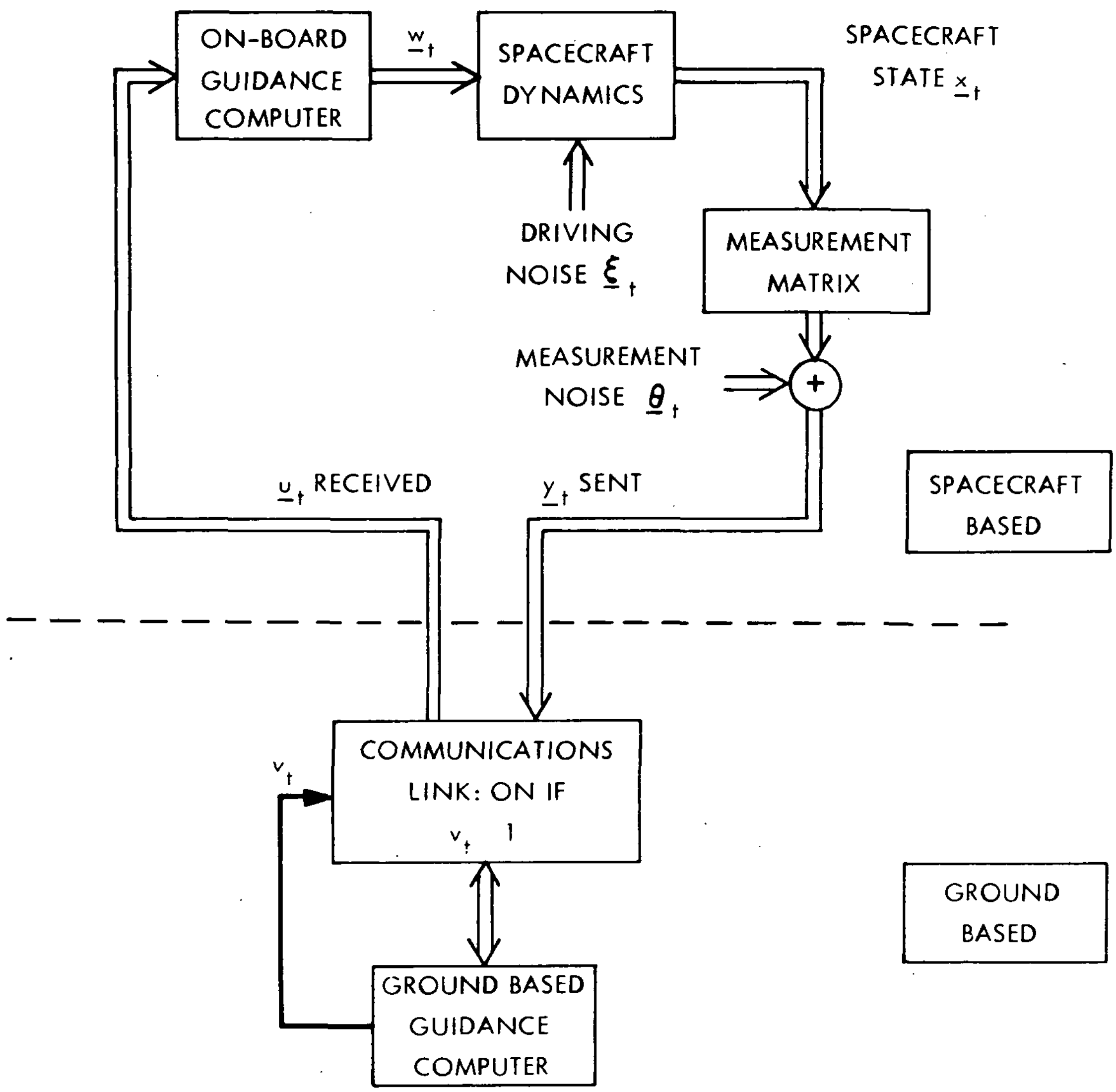

Fig. 6.1 System Structure

The terms are grouped slightly differently than in the analogous equation (4.1.3) in Section 4.1 because of the different timing considerations here. The parameters ℓ_t are constants which multiply v_t, which itself is either zero or one.

Obviously, the system described so far in this section represents only one of many possible structures. For example, should the penalty for dynamic control in the cost functional depend on $\underline{u}_t$ or $\underline{w}_t$? If $\underline{u}_t$ is used, cost is only accumulated once if $\underline{u}_t$ is non-zero once and then zero over an interval. If $\underline{w}_t$ is used in the penalty function, cost will continue to be built up even if no further controls are sent after "loading" a non-zero value into the on-board computer. (see equation (6.2.3)). As another example, there is no noise in the $\underline{w}$-equation (6.2.3), as was discussed in Section 6.1, while in some circumstances noise might occur there.

In this chapter, the specific problem structure introduced in equations (6.2.1) - (6.2.6) will be analyzed. Other equally justifiable structures will be left for future work.

6.3 Solution Techniques

Evidently, the state of the spacecraft at time t is composed of two parts: the "real" state $\underline{x}_t$ and the "computational" state composed of the data $\underline{w}_t$ stored in the computer. If an overall state $\underline{z}_t$ is defined:

$$\underline{z}_t = \begin{bmatrix} \underline{x}_t \\ \cdots \\ \underline{w}_t \end{bmatrix} \qquad (6.3.1)$$

the overall system may be described by an equation of the form

$$\underline{z}_{t+1} = [v_t \underline{A}_t^1 + (1-v_t)\underline{A}_t^0]\underline{z}_t + [v_t \underline{B}_t^1 + (1-v_t)B_t^0]\underline{u}_t + \underline{D}_t \underline{\xi}_t \qquad (6.3.2)$$

where

$$\underline{A}_t^1 \triangleq \left[\begin{array}{c|c} \underline{A}_t & \underline{B}_t \\ \hline \underline{0} & \underline{0} \end{array}\right] \qquad (6.3.3)$$

$$\underline{B}_t^1 \triangleq \left[\begin{array}{c|c} \underline{0} & \underline{0} \\ \hline \underline{0} & \underline{0} \end{array}\right] \qquad (6.3.4)$$

$$\underline{A}_t^0 \triangleq \left[\begin{array}{c|c} \underline{A}_t & \underline{B}_t \\ \hline 0 & \underline{I} \end{array}\right] \qquad (6.3.5)$$

$$\underline{B}_t^0 \triangleq \left[\begin{array}{c|c} \underline{0} & \underline{0} \\ \hline \underline{0} & \underline{0} \end{array}\right] \qquad (6.3.6)$$

$$\underline{D}_t \triangleq \left[\begin{array}{c} \underline{I} \\ \hline \underline{0} \end{array}\right] \qquad (6.3.7)$$

This reflects the fact that different A and B matrices are appropriate depending on v_t. Such a formulation allows more-or-less arbitrary on-board computer actions to be considered so long as they are linear. The problem is reduced to "ordinary" form except that now the matrices in the dynamic equation depend on v_t.

This general structure does not lead to convenient analysis. One does not expect a separation property for $\{v_t\}$ to hold since it directly influences the state. On the other hand, if the particular structure outlined in Section 6.2 is maintained, the essentially different natures of $\underline{x}$ and $\underline{w}$ parts of the state can be exploited to give interesting interpretations to the results.

Thus, the solution of the following optimal control problem will be obtained: Subject to

$$\underline{x}_{t+1} = \underline{A}_t \underline{x}_t + \underline{B}_t \underline{w}_t + \underline{\xi}_t \tag{6.3.8}$$

$$\underline{w}_{t+1} = (1 - v_t)\underline{w}_t + v_t \underline{u}_t \tag{6.3.9}$$

$$\underline{u}_t \in R^m \; ; \quad v_t \in \{0, 1\} \tag{6.3.10}$$

select the dynamic controls $\{\underline{u}_0, \ldots, \underline{u}_{T-1}\}$ and the measurement controls $\{v_0, \ldots, v_T\}$ to minimize

$$J = \underline{x}_T' \underline{Q}_T \underline{x}_T + \ell_T v_T + \sum_{t=0}^{T-1} \underline{x}_t' \underline{Q}_t \underline{x}_t + \underline{u}_t' \underline{R}_t \underline{u}_t + \ell_t v_t \tag{6.3.11}$$

where feedback control laws of the type

$$\underline{u}_t^* = \underline{\phi}_t(Y_{t-1}) \tag{6.3.12}$$

$$v_t^* = \psi_t(Y_{t-1}) \tag{6.3.13}$$

are admissible and where the random variable J is to be minimized in the sense of Chapter II. The solution will be found via dynamic programming. Because of the different timing considerations in this Chapter, the data set $\tilde{Y}_t$ now represents $Y_t \cup U_t \cup V_t$.

<u>Step T</u>: The terminal cost-to-go is

$$\mathcal{J}_T = E\{\underline{x}'_T \underline{Q}_T \underline{x}_T + \ell_T \cdot v_T \,|\, \tilde{Y}_{T-1}\} \tag{6.3.14}$$

$$= \hat{\underline{x}}'_{T|T-1} \underline{Q}_T \hat{\underline{x}}_{T|T-1} + \mathrm{tr}[\underline{Q}_T \underline{\Sigma}_{T|T-1}] + \ell_T v_T \tag{6.3.15}$$

Note the conditioning, which is appropriate for the type of closed-loop control laws represented in (6.3.12) and (6.3.13). Since ℓ_T is presumably positive, the optimal choice of v_T is

$$v_T^* = 0 \tag{6.3.16}$$

which makes sense since a final measurement would be useless from the point-of-view of control. Thus the optimal cost-to-go is

$$\mathcal{V}_T = \hat{\underline{x}}'_{T|T-1} \underline{Q}_T \hat{\underline{x}}_{T|T-1} + \mathrm{tr}\,\underline{Q}_T \underline{\Sigma}_{T|T-1} \tag{6.3.17}$$

Evidently, to continue the dynamic programming, the form of the Kalman filter which generates $\hat{\underline{x}}_{t+1|t}$ from $\hat{\underline{x}}_{t|t-1}$ will be needed. The appropriate equations are (see Rhodes (R.1)):

$$\hat{\underline{x}}_{t+1|t} = \underline{A}_t \hat{\underline{x}}_{t|t-1} + \underline{B}_t \underline{w}_t + \underline{A}_t \underline{\Sigma}_{t|t-1} \underline{C}'_t v_t (\underline{C}_t \underline{\Sigma}_{t|t-1} \underline{C}'_t + \underline{\Theta}_t)^{-1} (\underline{y}_t - \underline{C}_t \hat{\underline{x}}_{t|t-1}) \tag{6.3.18}$$

$$\underline{\Sigma}_{t+1|t} = \underline{\Xi}_t + \underline{A}_t [\underline{\Sigma}_{t|t-1} - \underline{\Sigma}_{t|t-1} \underline{C}'_t v_t (\underline{C}_t \underline{\Sigma}_{t|t-1} \underline{C}'_t + \underline{\Theta}_t)^{-1} \underline{C}_t \underline{\Sigma}_{t|t-1}] \underline{A}'_t \tag{6.3.19}$$

$$\underline{\Sigma}_{0|-1} = \underline{\Sigma}_0 = \text{given}\,; \quad \hat{\underline{x}}_{0|-1} = \hat{\underline{x}}_0 = \text{given} \tag{6.3.20}$$

Note that in (3.5.13), $\underline{w}_t$ is taken to be deterministic. In fact, throughout this derivation it will be assumed that the ground computer knows the state of the on-board computer exactly. This is a reasonable assumption in many circumstances. Also note that the form of the Kalman filter equations given in (6.3.18) - (6.3.20) utilizes the fact that v_t is constrained to be zero or one. Specifically, v_t-factors would appear in several other places in these equations, but for the fact that one times one equals one and zero times zero equals zero.

Step T-1: The cost-to-go is

$$\mathcal{J}_{T-1} = E\{\underline{x}'_{T-1}\underline{Q}_{T-1}\underline{x}_{T-1} + \underline{u}'_{T-1}\underline{R}_{T-1}\underline{u}_{T-1} + \ell_{T-1}\cdot v_{T-1} + \mathcal{J}_T(\hat{\underline{x}}_{T\,|\,T-1}, \underline{\Sigma}_{T|T-1}) \,|\, \breve{Y}_{T-2}\} \tag{6.3.21}$$

$$= \hat{\underline{x}}'_{T-1|T-2}\underline{Q}_{T-1}\hat{\underline{x}}_{T-1|T-2} + \mathrm{tr}[\underline{Q}_{T-1}\underline{\Sigma}_{T-1|T-2}] + \underline{u}'_{T-1}\underline{R}_{T-1}\underline{u}_{T-1} + \ell_{T-1}v_{T-1} + E\{\hat{\underline{x}}'_{T|T-1}\underline{Q}_T\hat{\underline{x}}_{T|T-1} + \mathrm{tr}[\underline{Q}_T\underline{\Sigma}_{T|T-1}] \,|\, \breve{Y}_{T-2}\} \tag{6.3.22}$$

Substituting (6.3.18) and (6.3.19) into the expected value in (6.3.22) and continuing with a considerable amount of algebra results in

$$\mathcal{J}_{T-1} = \hat{\underline{x}}'_{T-1|T-2}\underline{Q}_{T-1}\hat{\underline{x}}_{T-1|T-2} + \mathrm{tr}[\underline{Q}_{T-1}\underline{\Sigma}_{T-1|T-2}] + \underline{u}'_{T-1}\underline{R}_{T-1}\underline{u}_{T-1} + \ell_{T-1}v_{T-1} + (\underline{A}_{T-1}\hat{\underline{x}}_{T-1|T-2} + \underline{B}_{T-1}\underline{w}_{T-1})'\underline{Q}_T(\underline{A}_{T-1}\underline{x}_{T-1|T-2} + \underline{B}_{T-1}\underline{w}_{T-1}) + \mathrm{tr}[\underline{Q}_T\underline{\Xi}_{T-1}] + \mathrm{tr}[\underline{Q}_T\underline{A}_{T-1}\underline{\Sigma}_{T-1|T-2}\underline{A}'_{T-1}] \tag{6.3.23}$$

The optimal choice of v_{T-1} and u_{T-1} are

$$v^*_{T-1} = 0 \qquad (6.3.24)$$

$$u^*_{T-1} = \underline{0} \qquad (6.3.25)$$

which again makes sense: A control command sent at time t = T-1 will not reach the dynamics until t = T, when it would be usless. Similarly it is useless to measure at time T-1 since nothing can be done at t = T. The optimal cost-to-go is given by

$$\begin{aligned} \mathcal{J}_{T-1} = {} & \hat{x}_{T-1|T-2}(Q_{T-1} + \underline{A}'_{T-1}\underline{Q}_T\underline{A}_{T-1})\hat{\underline{x}}_{T-1|T-2} \\ & + \underline{w}'_{T-1}(\underline{B}'_{T-1}\underline{Q}_T\underline{B}_{T-1})\underline{w}_{T-1} \\ & + \hat{\underline{x}}'_{T-1|T-2}(2\underline{A}'_{T-1}\underline{Q}_T\underline{B}_{T-1})\underline{w}_{T-1} \\ & + \mathrm{tr}\left[(\underline{Q}_{T-1} + \underline{A}'_{T-1}\underline{Q}_T\underline{A}_{T-1})\underline{\Sigma}_{T-1|T-2}\right] \\ & + \mathrm{tr}\left[\underline{Q}_T\underline{\Xi}_T\right] \end{aligned} \qquad (6.3.26)$$

This form is suggestive, and one might try to derive a recursive form for the optimal cost-to-go. In fact, one can prove the following theorem:

<u>Theorem 6.3.1</u>: If at step t+1, the optimal cost-to-go for the problem defined by equations (6.3.8) - (6.3.13) is of the form

$$\mathcal{J}_{t+1} = \left[\hat{\underline{x}}'_{t+1|t} \vdots \underline{w}'_{t+1}\right] \underline{M}_{t+1} \begin{bmatrix} \hat{\underline{x}}_{t+1|t} \\ \hline \underline{w}_{t+1} \end{bmatrix} + I_{t+1}(\underline{\Sigma}_{t+1|t}) \qquad (6.3.27)$$

Where $\underline{M}_{t+1}$ is a <u>deterministic</u> matrix which is positive semi-definite and which may be partitioned as follows:

$$\underline{M}_{t+1} = \left[\begin{array}{c|c} \underline{M}^{11}_{t+1} & \underline{M}^{12}_{t+1} \\ \hline \underline{M}^{21}_{t+1} & \underline{M}^{22}_{t+1} \end{array}\right] \tag{6.3.28}$$

with $\underline{M}^{11}_{t+1} = (\underline{M}^{11}_{t+1})'$, $\underline{M}^{22}_{t+1} = (\underline{M}^{22}_{t+1})'$, and $\underline{M}^{21}_{t+1} = (\underline{M}^{12}_{t+1})'$, and where $I(\cdot)$ is a deterministic real value function on the set of nxn matrices, then it follows that:

(1) The optimal measurement control v_t^* may be determined as follows:

(1.1) Calculate $\mathcal{J}_t\big|_{v_t=0}$, given by

$$\begin{aligned} \mathcal{J}_t\big|_{v_t=0} = {} & \underline{\hat{x}}'_{t|t-1}(\underline{Q}_t + \underline{A}'_t\underline{M}^{11}_{t+1}\underline{A}_t)\underline{\hat{x}}_{t|t-1} \\ & + 2\underline{\hat{x}}'_{t|t-1}\underline{A}'_t(\underline{M}^{11}_{t+1}\underline{B}_t + \underline{M}^{12}_{t+1})\underline{w}_t \\ & + \underline{w}'_t(\underline{M}^{22}_{t+1} + \underline{B}'_t\underline{M}^{12}_{t+1} + \underline{M}^{21}_{t+1}\underline{B}_t + \underline{B}'_t\underline{M}^{11}_{t+1}\underline{B}_t)\underline{w}_t \\ & + \underline{I}_{t+1}(\underline{A}_t\underline{\Sigma}_{t|t-1}\underline{A}'_t + \underline{\Xi}_t) + tr[\underline{Q}_t\underline{\Sigma}_{t|t-1}] \end{aligned} \tag{6.3.29}$$

(1.2) Calculate $\mathcal{J}_t\big|_{v_t=1}$ given by

$$\begin{aligned} \mathcal{J}_t\big|_{v_t=1} = {} & \underline{\hat{x}}'_{t|t-1}(\underline{Q}_t + \underline{A}'_t\underline{M}^{11}_{t+1}\underline{A}_t - \underline{A}'_t\underline{P}_t\underline{A}_t)\underline{\hat{x}}_{t|t-1} \\ & + 2\underline{\hat{x}}'_{t|t-1}\underline{A}'_t(\underline{M}^{11}_{t+1} - \underline{P}_t)\underline{B}_t\underline{w}_t \\ & + \underline{w}'_t\underline{B}'_t(\underline{M}^{11}_{t+1} - \underline{P}_t)\underline{B}_t\underline{w}_t \\ & + \ell_t + tr[\underline{Q}_t\underline{\Sigma}_{t|t-1}] \\ & + tr[\underline{M}^{11}_{t+1}\underline{A}_t\underline{\Sigma}_{t|t-1}\underline{C}'_t(\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{C}'_t + \underline{\Theta})^{-1}\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{A}'_t \end{aligned}$$

$$+ I_{t+1}(\Xi_t + A_t(\Sigma_{t|t-1} - \Sigma_{t|t-1}C_t'(C_t\Sigma_{t|t-1}C_t' + \Theta_t)^{-1} C_t\Sigma_{t|t-1})A_t') \quad (6.3.30)$$

$$\text{where } P_t \triangleq M_{t+1}^{12}(R_t + M_{t+1}^{22})^{-1}M_{t+1}^{21} \quad (6.3.31)$$

(1.3) If $\mathcal{J}_t\Big|_{v_t=0} < \mathcal{J}_t\Big|_{v_t=1}$, set $v_t^* = 0$

If $\mathcal{J}_t\Big|_{v_t=1} < \mathcal{J}_t\Big|_{v_t=0}$, set $v_t^* = 1$ (6.3.32)

(2) The optimal dynamic control u_t^* is equal to

$$u_t^* = -v_t^*(R_t + M_{t+1}^{22})^{-1}M_{t+1}^{21}(A_t\hat{x}_{t|t-1} + B_tw_t) \quad (6.3.33)$$

(3) The optimal cost-to-go from step t is given by

$$\mathcal{J}_t = [\hat{x}_{t|t-1}' \vdots w_t']M_t\begin{bmatrix}\hat{x}_{t|t-1} \\ \hline w_t\end{bmatrix} + I_t(\Sigma_{t|t-1}) \quad (6.3.34)$$

where

$$M_t^{11} = Q_t + A_t'M_{t+1}^{11}A_t - v_t^*A_t'P_tA_t \quad (6.3.35)$$

$$M_t^{12} = A_t'(M_{t+1}^{11} - v_t^*P_t)B_t + (1-v_t^*)A_t'M_{t+1}^{22} \quad (6.3.36)$$

$$M_t^{21} = (M_t^{12})' \quad (6.3.37)$$

$$M_t^{22} = B_t'(M_{t+1}^{11} - v_t^*P_t)B_t + (1-v_t^*)(M_{t+1}^{22} + B_t'M_{t+1}^{12} + M_t^{21}B_t) \quad (6.3.38)$$

$$P \triangleq M_t^{12}(R_t + M_t^{22})^{-1}M_{t+1}^{21} \quad (6.3.39)$$

$$I_t = \ell_t v_t^* + \mathrm{tr}[\underline{Q}_t \underline{\Sigma}_{t|t-1}]$$
$$+ v_t^* \cdot \mathrm{tr}[\underline{M}^{11}_{t+1} \underline{A}_t \underline{\Sigma}_{t|t-1} \underline{C}_t' (\underline{C}_t \underline{\Sigma}_{t|t-1} \underline{C}_t' + \underline{\Theta}_t)^{-1} \underline{C}_t \underline{\Sigma}_{t|t-1} \underline{A}_t']$$
$$+ I_{t+1}(\underline{\Xi}_t + \underline{A}_t' \underline{\Sigma}_{t|t-1} \underline{A}_t - v_t^* \underline{A}_t \underline{\Sigma}_{t|t-1} \underline{C}_t'$$
$$(\underline{C}_t \underline{\Sigma}_{t|t-1} \underline{C}_t' + \underline{\Theta}_t)^{-1} \underline{C}_t \underline{\Sigma}_{t|t-1} \underline{A}_t') \qquad (6.3.40)$$

Proof: Given in Appendix E.

Remark 1: If $\mathcal{J}_t \big|_{v_t=0} = \mathcal{J}_t \big|_{v_t=1}$ inequality (6.3.32), the cost resulting from choosing $v_t = 0$ equals the cost resulting from the choice $v_t = 1$, and either alternative may be selected.

Remark 2: Note the particular form of optimal dynamic control $\underline{u}_t^*$ which results if $v_t^* = 1$. It is a gain matrix times the expected value of $\underline{x}_{t+1}$ given the data up to t-1. Since there is a unit delay in processing the control command $\underline{u}_t^*$ through the on-board computer, this means that the current optimal dynamic control is a gain times the value of the state expected at the time when the current control will influence the dynamics. This is consistent with the results obtained in previous chapters.

Remark 3: It is interesting to note the manner in which the state estimate $\hat{\underline{x}}_{t|t-1}$ and the covariance $\underline{\Sigma}_{t|t-1}$ influence the decision rule for v_t given in conclusion (1) of the theorem. One can argue loosely that if the state estimate is "large", the term $-\hat{\underline{x}}_{t|t-1}' \underline{P}_t \hat{\underline{x}}_{t|t-1}$ in the cost resulting from $v_t = 1$ will get large and negative,* and as a result, the optimal

---

* Note that $\underline{P}_t$ is a positive semidefinite matrix.

choice of v_t will be $v_t^* = 1$. This can be interpreted as indicating that if the state estimate is far away from zero, it is advantageous to set $v_t = 1$ in order to send a control command. Similarly, if $\underline{\Sigma}_{t|t-1}$ is "large" one can also argue that the resulting choice of v_t^* will be $v_t^* = 1$ because of the influence of $\underline{\Sigma}_{t|t-1}$ on the argumentof the $I_{t+1}(\cdot)$ function. This can be interpreted as indicating that if the covariance matrix is large indicating a poor estimate of the state, it is advantageous to set $v_t=1$ to obtain a measurement. Of course, these arguments are very heuristic: the other terms in $\mathcal{J}_t\Big|_{v_t=0}$ and $\mathcal{J}_t\Big|_{v_t=1}$ must be considered. The interpretation is, however, consistent with intuition.

Although Theorem 6.3.1 displays the optimal controls and cost-to-to at any time t assuming a certain form for the cost-to-go from step t+1, this does not complete the analysis of the problem under consideration. This is the case even though the terminal cost-to-go (6.3.17) takes the hypothesized form (set $\underline{M}_T^{11} = \underline{Q}_T$, $I_T = tr[\underline{Q}_T\underline{\Sigma}_{T|T-1}]$, $\underline{M}_T^{12} = \underline{0}$, $\underline{M}_T^{22} = \underline{0}$). The reason is that given data Y_{t-1}, the matrix $\underline{M}_{t+1}$ in Theorem 6.3.1 is not deterministic. That assumption was made in Theorem 6.3.1 and it was used extensively in the proof in order to evaluate expected values involving the matrix $\underline{M}_{t+1}$. A similar assumption was made for the $I_{t+1}(\cdot)$ function.

To see why $\underline{M}_{t+1}$, for example, is not deterministic in general given $\tilde{Y}_{t-1}$, consider the expression for $\mathcal{J}_{T-1}$ given in (6.3.26). This is of the form of (6.3.27) if one identifies

$$\underline{M}^{11}_{T-1} = \underline{Q}_{T-1} + \underline{A}'_{T-1}\underline{Q}_T\underline{A}_{T-1} \tag{6.3.41}$$

$$\underline{M}^{22}_{T-1} = \underline{B}'_{T-1}\underline{Q}_{T-1}\underline{B}_{T-1} \tag{6.3.42}$$

$$\underline{M}^{12}_{T-1} = \underline{A}'_{T-1}\underline{Q}_T\underline{B}_{T-1} = (\underline{M}^{21}_{T-1})' \tag{6.3.43}$$

$$I_{T-1} = tr[\underline{Q}_T\underline{\Xi}_T + (\underline{Q}_T + \underline{A}'_{T-1}\underline{Q}_{T-1}\underline{A}_{T-1})\underline{\Sigma}_{T-1|T-2}] \tag{6.3.44}$$

These are deterministic quantities. The optimal value of v_{T-2}, however, depends on $\hat{\underline{x}}_{T-2|T-3}$, as specified by Theorem 6.3.1, and as a result, the values of $\underline{M}^{11}_{T-2}$, $\underline{M}^{22}_{T-2}$, and $\underline{M}^{12}_{T-2}$ and the form of I_{T-2} which result from the evaluation of equations (3.3.34) - (3.3.40) depend on $\hat{\underline{x}}_{T-2|T-3}$. Thus, when the dynamic programming is "stepped backwards" to time T-3, when $\breve{Y}_{T-4}$ is the known data, the conditional mean $\hat{\underline{x}}_{T-2|T-3}$ becomes a random variable and through it the matrix $\underline{M}_{T-2}$ and the function I_{T-2} become random. The overall conclusion, then, is that at any time t less than T-1, the matrix $\underline{M}_t$ is a random variable because the propagation of the matrices $\{\underline{M}_t, \underline{M}_{t+1}, \ldots, \underline{M}_T\}$ backwards from time T to time t depends on the future conditional means, which are random variables at time t. Similarly, the I_t function has random parameters in it.

This does not mean, however, that the results of Theorem 6.3.1 cannot be used to solve the problem at hand. It merely indicates that one must be careful to include the effects of the random character of the variables used. To see how this can be done, the dynamic programming solution that was carried to step t = T-1 (see equation (6.3.26)) will now be continued farther forward in time.

Step T-2: Reference to equations (6.3.26) and (6.3.41) through (6.3.44) shows that the cost-to-go from step T-1 to the end is of the form

$$\mathcal{J}_{T-1} = [\underline{\hat{x}}'_{T-1|T-2} \vdots \underline{w}'_{T-1}] \underline{M}_{T-1} \begin{bmatrix} \underline{\hat{x}}_{T-1|T-2} \\ ------ \\ \underline{w}_{T-1} \end{bmatrix} + I_{T-1}(\Sigma_{T-1|T-2}) \qquad (6.3.45)$$

where given $\tilde{Y}_{T-3}$, $\underline{M}_{T-1}$ and I_{T-1} are deterministic. Thus Theorem 6.3.1 is directly applicable, and it follows that one can evaluate v^*_{T-2} and $\underline{u}^*_{T-2}$ using that theorem:

(1) Calculate $\mathcal{J}_{T-2}\big|_{v_{T-2}=0}$ and $\mathcal{J}_{T-2}\big|_{v_{T-2}=1}$ according to (6.3.29) - (6.3.31)

(2) Choose v^*_{T-2} according to (6.3.32)

(3) Choose $\underline{u}^*_{T-2}$ accordint to (6.3.33)

(4) Calculate $\mathcal{J}_{T-2}$ according to (6.3.34) - (6.3.40)

Step T-3: The cost-to-go is

$$\mathcal{J}_{T-3} = E\{\underline{x}'_{T-3}\underline{Q}_{T-3}\underline{x}_{T-3} + \underline{u}'_{T-3}\underline{R}_{T-3}\underline{u}_{T-3} + \ell_{T-3}\cdot v_{T-3} + \mathcal{J}_{T-2}\,|\,\tilde{Y}_{T-4}\} \qquad (6.3.46)$$

$$= \underline{\hat{x}}'_{T-3|T-4}\underline{Q}_{T-3}\underline{\hat{x}}_{T-3|T-4} + tr[\underline{Q}_{T-3}\underline{\Sigma}_{T-3|T-4}] + \underline{u}'_{T-3}\underline{R}_{T-3}\underline{u}_{T-3} + \ell_{T-3}\cdot v_{T-3} + E\{\mathcal{J}_{T-2}\,|\,\tilde{Y}_{T-4}\} \qquad (6.3.47)$$

The complicated step in the analysis is the evaluation of the last term in (6.3.47):

$$E\{\ell_{T-2} | \tilde{Y}_{T-4}\} = E\{[\underline{\hat{x}}'_{T-2|T-3} \vdots \underline{w}'_{T-2}] \underline{M}_{T-2} \begin{bmatrix} \underline{\hat{x}}_{T-2|T-3} \\ \hline \underline{w}_{T-2} \end{bmatrix}$$

$$+ I_{T-2}(\underline{\Sigma}_{T-2|T-3}) | \tilde{Y}_{T-4}\} \tag{6.3.48}$$

As in the proof of Theorem 6.3.1 which is given in Appendix E, it is convenient to consider separately the two cases $v_{T-3} = 0$ and $v_{T-3} = 1$.

<u>Case 1</u>: If $v_{T-3} = 0$, the propagation equations for $\underline{\hat{x}}_{T-2|T-3}$, $\underline{w}_{T-2}$, and $\underline{\Sigma}_{T-2|T-3}$ are

$$\underline{\hat{x}}_{T-2|T-3} = \underline{A}_{T-3}\underline{\hat{x}}_{T-3|T-4} + \underline{B}_{T-3}\underline{w}_{T-3} \tag{6.3.49}$$

$$\underline{w}_{T-2} = \underline{w}_{T-3} \tag{6.3.50}$$

$$\underline{\Sigma}_{T-2|T-3} = \underline{A}_{T-3}\underline{\Sigma}_{T-2|T-3}\underline{A}'_{T-3} + \underline{\Xi}_{T-3} \tag{6.3.51}$$

If these equations are substituted into (6.3.47), it is clear that only $\underline{M}_{T-2}$ and I_{T-2} are random (check equations (6.3.35) through (6.3.40) which express $\underline{M}_{T-2}$ and I_{T-2} in terms of v^*_{T-2}). Thus, one obtains

$$E\{\ell_{T-2} | \tilde{Y}_{T-4}\} \Big|_{v_{T-3}=0} =$$

$$[\underline{\hat{x}}'_{T-3|T-4}\underline{A}'_{T-3} + \underline{w}'_{T-3}\underline{B}'_{T-3} \vdots \underline{w}'_{T-3}] E\{\underline{M}_{T-2} | \tilde{Y}_{T-4}\} \begin{bmatrix} A_{T-3}\hat{x}_{T-3|T-4} + B_{T-3}w_{T-3} \\ \hline \underline{w}_{T-3} \end{bmatrix}$$

$$+ E\{I_{T-2}(\underline{A}_{T-3}\underline{\Sigma}_{T-3|T-4}\underline{A}'_{T-3} + \underline{\Xi}_{T-3}) | \tilde{Y}_{T-4}\} \tag{6.3.52}$$

Now $\underline{M}_{T-2}$ and I_{T-2} have bi-valued random properties. They take on one pair of values if $v^*_{T-2} = 0$ and a second pair if $v^*_{T-2} = 1$. This means that the expected value of, say, $\underline{M}_{t-2}$ given $\breve{Y}_{T-4}$ may be expressed

$$E\{\underline{M}_{T-2} | \breve{Y}_{T-4}\} = (\underline{M}_{T-2} | v^*_{T-2} = 0) \times p(v^*_{T-2} = 0 | \breve{Y}_{T-4}) + (\underline{M}_{T-2} | v^*_{T-2} = 1) \times p(v^*_{T-2} = 1 | \breve{Y}_{T-4}) \quad (6.3.53)$$

where $p(v^*_{T-2} = z | \breve{Y}_{T-4})$ represents "the probability that v^*_{T-2} equals z given $\breve{Y}_{T-4}$". A similar expression holds for $E\{\underline{I}_{T-2} | \breve{Y}_{T-4}\}$.

It is possible to evaluate the probabilities required to find $E\{I_{T-2} | \breve{Y}_{T-4}\}$ and $E\{\underline{M}_{T-2} | \breve{Y}_{T-4}\}$. Since $\hat{\underline{x}}_{T-2|T-3}$ is a Gaussian random vector given $\breve{Y}_{T-4}$, and since v^*_{T-2} is expressed in terms of $\hat{\underline{x}}_{T-2|T-3}$, (see conclusion (1) of Theorem 6.3.1), the task can be carried out (numerically). Once this is done and the result is substituted into equation (6.3.52), then (6.3.52) into (6.3.47), one has evaluated the cost-to-go $\mathcal{J}_{T-3}$ under the assumption that $\underline{v}_{T-3}$ is zero. Notice that

(1) If v_{T-3} is zero, the only term in $\mathcal{J}_{T-3}$ that depends on $\underline{u}_{T-3}$ in (6.3.45) is the term $\underline{u}'_{T-3}\underline{R}_{T-3}\underline{u}_{T-3}$. Thus the optimal value of u_{T-3}, given $v_{T-3} = 0$, is $u_{T-3} = \underline{0}$.

(2) The resulting cost-to-go is quadratic in the mean $\hat{\underline{x}}_{T-3|T-4}$ and the state $\underline{w}_{T-3}$ with additive terms involving the covariance $\underline{\Sigma}_{T-3|T-4}$.

It is now necessary to consider the case $v_{T-3} = 1$.

Case 2: $v_{T-3} = 1$: If $v_{T-3} = 1$, the propagation equations for $\hat{\underline{x}}_{T-2|T-3}$, $\underline{\Sigma}_{T-2|T-3}$, and $\underline{w}_{T-2}$ are the full Kalman filter equations (6.3.18) and (6.3.19), and the $\underline{w}$-equation becomes

$$\underline{w}_{T-2} = \underline{u}_{T-3} \qquad (6.3.54)$$

the term $E\{\mathcal{J}_{T-2}|\breve{Y}_{T-4}\}$ which appears in (6.3.46) is now somewhat more difficult to evaluate, because $\hat{\underline{x}}_{T-2|T-3}$, $\underline{M}_{T-2}$, and $\underline{I}_{T-2}$ are all random variables. However, one can use the fundamental property of expections (see Loeve (L.1) p. 341) that $E\{f(a,g)\} = E\{E\{f(a,b)|a\}\}$ to write $E\{\mathcal{J}_{T-2}|\breve{Y}_{T-4}\}$ as

$$E\{\mathcal{J}_{T-2}|\breve{Y}_{T-4}\} = E\{E\{\mathcal{J}_{T-2}|\breve{Y}_{T-4}, v_{T-2}\}|\breve{Y}_{T-4}\} \qquad (6.3.55)$$

The inner expectation in (6.3.53) can be evaluated in closed form using the methods of Appendix E. This is becuase given v_{T-2}, only $\hat{\underline{x}}_{T-2|T-3}$ is random in the expression for $\mathcal{J}_{T-2}$; $\underline{M}_{T-2}$ and $\underline{I}_{T-2}$ are deterministic. In particular, $E\{\mathcal{J}_{T-2}|\breve{Y}_{T-4}, v_{T-2}\}$ can be evaluated for $v_{T-2} = 0$ and for $v_{T-2} = 1$, and then the entire expression (6.3.55) can be evaluated by adding the results for $v_{T-2} = 0$ and $v_{T-2} = 1$ weighted by the respective probabilities of these two cases. One can see that again the cost-to-go will be quadratic in $\hat{\underline{x}}_{T-3|T-4}$ and $\underline{w}_{T-3}$ with addative terms in $\underline{\Sigma}_{T-3|T-4}$. The calculations will not be carried out here for two reasons: the evaluation of the inner expectation $E\{\mathcal{J}_{T-2}|\breve{Y}_{T-4}, v_{T-2}\}$ is exactly like the proof of Theorem 6.3.1, which is given in detail in Appendix E, and the evaluation of the probabilities $p(v_{T-2} = 1|\breve{Y}_{T-4})$ and $p(v_{T-2} = 0|\breve{Y}_{T-4})$ cannot be done analytically. This completes the consideration of the case $v_{T-3} = 1$.

Once one has evaluated (6.3.47) for the two cases $v_{T-3}=0$ and $v_{T-3}=1$, the remaining step is simple: For each case choose the optimal value of $\underline{u}_{T-3}$ (easily done since the cost is quadratic in $\underline{u}_{T-3}$) and see whether the resulting cost $\mathcal{J}_{T-3}\big|_{v_{T-3}=0}$ or $\mathcal{J}_{T-3}\big|_{v_{T-3}=1}$ is smaller. Choose the appropriate value of v_{T-3} and $\underline{u}_{T-3}$ on this basis.

This completes the analysis of step T-3 of the dynamic programming. A situation has been reached in which the results cannot be determined analytically, but the procedure to follow at an arbitrary time step in clear: In order to optimize $\mathcal{J}_t$ given by

$$\mathcal{J}_t = E\{\underline{x}_t'\underline{Q}_t\underline{x}_t + \underline{u}_t'\underline{R}_t\underline{u}_t + \ell_t\underline{v}_t + \mathcal{J}_{t+1} \mid \breve{Y}_{t-1}\} \tag{6.3.56}$$

carry out the following steps:

(1) Evaluate the first three terms of $\mathcal{J}_t$ analytically:

$$E\{\underline{x}_t'\underline{Q}_t\underline{x}_t + \underline{u}_t'\underline{R}_t\underline{u}_t + \ell_t v_t \mid \breve{Y}_{t-1}\} = \hat{\underline{x}}_{t|t-1}'\underline{Q}_t\hat{\underline{x}}_{t|t-1} + \mathrm{tr}[\underline{Q}_t\underline{\Sigma}_{t|t-1}] + \underline{u}_t'\underline{R}_t\underline{u}_t + \ell_t v_t \tag{6.3.57}$$

(2) Express $E\{\mathcal{J}_{t+1} \mid \breve{Y}_{t-1}\}$ as follows:

$$E\{\mathcal{J}_{t+1} \mid \breve{Y}_{t-1}\} = E\{E\{\mathcal{J}_{t+1} \mid \breve{Y}_{t-1}, v_{t+1}, \ldots, v_T\} \mid \breve{Y}_{t-1}\} \tag{6.3.58}$$

(3) Evaluate $E\{\mathcal{J}_{t+1} \mid \breve{Y}_{t-1}, v_{t+1}, \ldots, v_t\}$ using Theorem 6.3.1 for every possible future measurement control sequence $\{v_{t+1}, \ldots, v_T\}$. This is possible because the matrices $\underline{M}_{t+1}, \ldots, \underline{M}_T$ and the functions $I_{t+1}, \ldots, I_T$ are deterministic given $\{v_{t+1}, \ldots, v_T\}$.

(4) Evaluate the probability that each future measurement control sequence will occur, given Y_{t-1}.

(5) Evaluate $E\{\vartheta_{t+1} | \tilde{Y}_{t-1}\}$ using (3) and (4).

(6) Finally optimize $\mathcal{J}_t$ over $\underline{u}_t$ and v_t, which can be done once the expected value has been evaluated.

Obviously, the step in this procedure which is difficult is (4). Although it can be carried out in principle, it would be very time consuming to do so on-line in practice. Conceptually, however, the problem originally stated has been solved.

Since the solution to the problem under consideration may not be practically found for actual problems, it would make sense to investigate possible sub-optimal control schemes based on the analysis presented here. This, together with the investigation of the many possible variatious of the problem defined by equations (6.3.8) - (6.3.11), constitutes a potentially fruitful area for future research.

CHAPTER VII

CONCLUSIONS, APPLICATIONS AND AREAS FOR FUTURE RESEARCH

7.1 An Overview

The subject of this thesis has been the optimal control of stochastic systems, and this topic has been treated from several points of view. The purpose of this chapter is to attempt to unify the results presented earlier, to comment on some areas of application, and to suggest possible topics for future research.

The first question to be raised in this thesis concerned the definition of a stochastic optimal control problem. How does one construct a mathematical formulation that is consistent, that is satisfying to physical intuition, and that leads to a well-defined optimization procedure? The issue was discussed in Section 1.3, which presented a problem formulation, and in Chapter II, which dealt with defining stochastic dynamic programming. While it is true that much of the mathematical structure built up in Section 1.3, viz., the measure theoretic foundation of probability theory, was not used explicitly in the sequel, this material was none the less included in the introductory chapter because it was implicitly used.

One of the issues discussed was the difficulty of leaving the stochastic structure of the problem undefined while searching for control laws. The apparent problem could be succinctly phrased: no control law, no random variables; no random variables, no formulation. This issue would not have been clear without reference to axiomatic probability.

In subsection 2.4.1, this difficulty was resolved by noting that when using the dynamic programming formulation, the fact that at time t , one assumes that the future control laws have already been selected and the past control values are fixed but arbitrary assures that all the quantities of interest are well-defined random variables. This fact is not novel; no doubt, those who are well versed in the field of stochastic control are aware of it. On the other hand, this point is often ignored completely, and a discussion of the issue is not to be found in most published works dealing with stochastic control. For this reason, such a discussion was included here.

The nature of dynamic programming as applied to the solution of control problems involving measurement control capability was discussed. One main point of this part of Chapter II was a sort of reverse-causality argument: the optimal measurement control at time t had to be selected in such a way as to optimize the cost-to-go from time t to time T , the terminal time, with no regard to optimizing the estimate of costs that have already been incurred. This allowed the dynamic programming structure to be applied more-or-less in the usual way. The same sort of thing was done in previous works dealing with optimal measurement control, e.g., Meier, et. al (M.3), but no justification or discussion was included. Previous authors have done the correct thing without saying why. This can probably be attributed to the fact that primarily linear-Gaussian-quadratic problems have been considered by other authors, and the optimality of the a-priori-v technique described in Section 1.4 for such problems tends to submerge the issues involved in defining the true stochastic cost. Since the a-priori-$\underline{v}$ technique immediately reduces the given problem to one which has been solved before (the determination of $\{\underline{u}_t^*\}$) and one which can be solved by deterministic techniques without the use of dynamic programming (the determination of $\{\underline{v}_t^*\}$), one need not worry about these deep issues: they have been swept under the rug.

The second main issue considered in the defining of stochastic optimality and dynamic programming was the need to include in each optimization a provision for utilizing the information that optimal controls will be used in the future regardless of present actions. This is, of course, the entire basis of dynamic programming, either stochastic or deterministic. On the other hand, many authors are not precise in their statements of what is being optimized in a particular problem: a frequently made statement is that "the expected value of J given in equation (2.1.5) will be minimized". As was discussed in Section 2.3, there are many types of expectations one can take, and one should be careful to specify what one means to do. If one is minimizing, at each step t, the conditional expected value of the cost-to-go from step t to step T given the data up to time t and given the fact that optimal controls will be applied in the future, one should say precisely that.

After considering the dynamic programming approach to stochastic control, attention was turned to the possibility of using deterministic techniques. In Chapter III it was demonstrated that under certain circumstances, one could suppress the stochastic nature of the problem while optimizing and re-introduce it simply by taking two expected values at the end. This, in the opinion of the author, is a novel approach to stochastic control.

The idea of considering the Kalman filter equations rather than the given system equations as the basic plant is neither new nor suprising. What one is actually doing when adopting this point of view is considering the propagation of the conditional p.d.f. of the plant state $\underline{x}_t$ given the data $\widehat{Y}_t$, which, as was discussed in Chapters II and III, constitutes the true state of the system anyway. The "lucky" thing is that in the linear-Gaussian case, the p.d.f. is finite dimensional, so one can proceed with relatively simple methods once one has built the filter, which itself is simple.

The novel part of Chapter III was Theorem 3.3.2. It is perhaps not surprising that for linear-Gaussian systems, it (loosely) doesn't matter when one takes expected values. After all, so many nice things are true about linear-Gaussian systems, why not this, too. Practically speaking, however, the benefit derived from being able to use this type of analysis, i.e., mainly deterministic, may be considerable in solving actual problems. The most significant benefit will result in the study of linear-Gaussian-non-quadratic problems.

In Chapters IV and V, linear-Gaussian-quadratic and linear-Gaussian-non-quadratic control problems were considered. It is felt that two distinct benefits result from a careful reading of Chapter IV: First, the reasons why the complete-two-way separation Theorem holds (Theorem 4.2.1) become clear, and second, the use of deterministic techniques is clarified.

It was mentioned in Chapter IV that the paper by Meir, et al (M. 3), indeed contains a result that is essentially equivalent to Theorem 4.2.1. However, in their proof, the authors did not clearly indicate why everything "worked" and did not consider the question of whether their dynamic programming was well defined. Hence, these points were cleared up. In addition, the deterministic analysis of Section 4.3 provides additional insight.

Section 4.3 contains many remarks regarding the use of deterministic techniques which will not be repeated here. It must be pointed out, however, that in using the deterministic techniques, one is almost, but not quite, forcing an a-priori-v structure on the solution. This is because during the parts of the analysis in which the innovations process is considered fixed, all information about the influence of the measurement control is suppressed from the problem. This, in turn, is because the measurement control influences primarily the statistics of the innovations, which are not used during the deterministic analysis. However, one does not really force an a-priori-v structure on the problem. The reason is that in finding the optimal feedback dynamic control as a function of the future innovations, one implicitly includes the effect of the measurement control. This effect becomes explicit when the expected value is taken. It may turn out for non-quadratic cost problems that after taking this expected value, one will find that the optimal dynamic control depends on the optimal measurement control, and the optimal measurement control on the mean and covariance of the state estimate given the data. Thus everything eventually requires a minimization in terms of the current conditional mean and covariance.

Chapter V considered the linear-Gaussian-non-quadratic case in detail. It should be noted that all the issues associated with such systems have not been fully resolved; necessary and sufficient conditions for an open-loop measurement program to be optimal have not yet been discovered. On the other hand, it is felt that significant insight into the nature of linear-Gaussian-non-quadratic problems can be gained by examining the analysis of Chapter V. In particular, the influence which the filtering part of the problem exerts on the control part of the problem is clarified by that analysis. Although there are many allusions in the literature to this one-way coupling effect (recall the discussion of Fig. 4.3), essentially nothing concrete has been written about it up to now.

The class of problems defined in Chapter VI represents a fertile field for possible future research, and further comments in this vein will be made in the next section. Let it be noted here, however, that problems of this type, in which there is a fundamental interaction between the concepts of control and communications, are certain to become increasingly important in practical applications. This must be the case as more and more complex control strategies are developed and as more and more ambitious systems are designed. It is inconceivable that as the state of the control art progresses, each dynamic system in the world will have its own computer to control it. Such a situation is absurd economically. It will become increasingly important for big supercomputers to communicate with remotely located dynamic systems in order to control them. Such systems can be formulated in the framework of Chapter VI, as has been previously discussed.

To conclude this section, the following idea should be stressed since this part of the thesis is titled "An Overview": There are many potential areas of application for the results developed here or for straightforward extensions which do not fit into the classical category of control systems. One might consider, for examples, the following possibilities:

1) Communications Systems: Recent work has been done on communications systems with two-way transmission links. This results in a so-called feedback communications system in which the receiver can try to tell the transmitter what it "thinks" it received. See, for example, Schalkwijk and Kailath (S.8), Schalkwijk (S.9), Omura (0.1), and Cruise (C.3). Problems of this type can be formulated as measurement optimization problems since the feedback link from receiver to transmitter is not unlike a controllable measurement equation.

2) Economic and Societal Systems: In this class of systems, the concept of measurement having a disruptive effect on the system is easy to visualize. Do polls influence public opinion, for example? One might try to formulate a measurement optimization with some sort of constraint that a certain minimum amount of information actually be obtained.

3) Physical systems actually subject to the Heisenberg effect: Although the author is not personally acquainted with such systems, he has been advised* that there are actual practical physical systems in which the Heisenberg Uncertainty Principle of physics plays an important role in that the more one measures, the more one disrupts the system. Optical communications systems were mentioned as an example. In designing such systems, one must select the best (in some sense) quantum operator to extract the modulation from a light beam

Evidently, other examples are possible.

7.2 Areas of Future Research

In this section, several topics for future research that are suggested by the work reported will be discussed.

7.2.1 Making the Cost Functional Deterministic

In Chapter II, an important point in the discussion was the requirement that the overall problem formulation result in a situation in which the control sequences $\{\underline{v}_t\}$ and $\{\underline{u}_t\}$ are mapped into the real numbers through the cost functional. The quantity J given in equation (2.1.5), which would ordinarily function as the cost functional in a deterministic control problem, is a random variable in a stochastic control problem, and it was necessary to map J into a real number $\mathcal{J}$

---

* By Prof. A. Baggeroer, one of the readers of this thesis.

in order to have a well-defined optimization. The main concern of Chapter II was determining the correct way in which to take the expected value of J to yield $\mathcal{J}$.

Early in the discussion it was noted that other devices besides taking the expected value of J to produce $\mathcal{J}$ could be used. The work of Sain (S.3) and Sain and Liberty (S.4) was mentioned as an example. These authors consider minimizing the variance of J subject to a constraint on its mean value.

In a problem with measurement control capability, devices other than taking the expected value of J might be very useful. One might heuristically argue that controlling the "quality" of the measurements has a more direct effect on the variance and higher order moments of the p.d.f. of the plant state than on its mean. The true nature of this effect, and means of exploiting the results in order to produce certain types of closed-loop system responses, are potential topics for future study.

7.2.2 Extending the Applicability of Deterministic Techniques

In Chapter III, the assumption was made that the plant under consideration was linear. This hypothesis was used in the proof of Theorem 3.3.2 in order to argue that the map from the dynamic control sequence to the cost was convex, which was in turn used to show that a certain inequality held. Theorem 3.3.2, when proved, justified the use of deterministic techniques by stating that the control derived by deterministic optimization techniques would result in performance which was no worse on the average than that produced by the stochastic optimal control.

In addition to its use in the proof of Theorem 3.3.2, the assumption of a linear system was used, along with the assumption of Gaussian statistics, to allow the propagation of the p.d.f. of the state to be formulated as a finite dimensional system: the Kalman filter. This made the transformation of the system described in subsection 3.3.1 useful in that the finite dimensional plant equations could be replaced by finite dimensional filter equations.

It would be very useful if these ideas could be extended to nonlinear systems. Doing so would involve two areas of study: extending Theorem 3.3.1 and extending Theorem 3.3.2. These two extension tasks seem to be independent at present, but they might turn out to be related.

There is a certain amount of theoretical justification for hoping that Theorems 3.3.1 and 3.3.2 can indeed be extended to include nonlinear systems if due care is exercised: primarily in the work of Kailath and Frost that stresses the innovations point of view in nonlinear filtering (F.1), (K.8). Their results indicate that even in the nonlinear case, the filter may be formulated in such a way that it is driven by a white innovations process, and this could perhaps be treated in a manner analogous to that of Chapter III; the innovations could be considered fixed until after a deterministic feedback control law was derived. Exactly how to deal with the infinite dimensionality of the filter, however, remains a challenging question.

The whole area of extension of the philosophy employed in Chapter III to nonlinear systems is, perhaps, the most pressing area for future work. This topic is discussed further from a different point of view in subsection 7.2.6 below.

7.2.3 Costs Explicitly Depending on Stochastic Parameters

In Remark 5 of Section 5.3, (following equation (5.3.37)) the fact was noted that the transformed cost obtained prior to applying deterministic techniques was not of a completely arbitrary nature. In particular, for linear-Gaussian systems, the types of transformed cost terms $\mathcal{L}_t(\hat{\underline{x}}_{t|t}, \underline{\Sigma}_{t|t}, \underline{u}_t, \underline{v}_{t+1})$ that one can obtain are specified as all possible forms obtainable by taking the expected value of terms of the form $L_t(\underline{x}_t, \underline{u}_t, \underline{v}_{t+1})$. The following question was raised: What types of system performance would result if the basic problem formulation involved a cost functional depending on $\hat{\underline{x}}_{t|t}$ and $\underline{\Sigma}_{t|t}$ (or the entire p.d.f. in non-Gaussian cases)? In other words, what happens if the p.d.f. which constitutes the true state of the stochastic problem is explicitly penalized in the cost functional in a way not directly obtainable as an expected value? This area is completely open for future work.

7.2.4 Linear-Gaussian-Non-Quadratic Problems

The status of linear-Gaussian-non-quadratic problems is not yet quite satisfactory even though more information than previously available has been supplied by the analysis of Chapter IV. It would be helpful if easily verifiable necessary and sufficient conditions could be obtained which would catagorize circumstances under which the optimal measurement control sequence could be computed off-line. It might be useful if several special cases of non-quadratic costs were completely analyzed, for example, "fuel optimal" type costs involving the absolute value of $\underline{u}_t$. (See Athans and Falb (A.7) for background.) This area would tie in nicely with the questions raised in subsection 7.2.3 concerning different types of cost functionals.

7.2.5 Extensions of Chapter VI

In several places in Chapter VI, it was pointed out that the particular problem being studied was just one special case of a whole class of potentially interesting problems which would be left for future study. It was mentioned that problems of this type would become increasingly important as control systems built around large central computers sending commands to remotely located dynamic systems become common. This structure is already being seen, for example, in the control of industrial processes: central computer installations are being used to control remotely located unit processes. Such a structure is also being used in aerospace and guided missile applications.

The types of variations of the problem studied in Chapter VI that might be considered include the following: What are the effects of varying degrees of sophistication of the "on-board" controller? What are the effects of noise corrupting the transmissions assumed noise-free in Chapter VI? What are the effects of various cost functionals which penalize quantities other than those appearing in J given in equation (6.2.6)? Do different mathematical problem formulations make sense? Are there more efficient procedures for calculating optimal controls than those described in Chapter VI? Are there any convenient suboptimal control schemes which perform well in practice?

7.2.6 Applications to Nonlinear Systems

The type of problem that this subsection recommends for future study is the general formulation introduced in Section 1.3, but used very little in this thesis beyond the general discussion of dynamic programming given in Chapter II: systems with generally nonlinear dynamics and nonlinear measurement equations.

There has, of course, been a great deal of work done on nonlinear filtering, that is, on the study of equations of the type of (2.3.1) and (2.3.2) which propagate the conditional p.d.f. of the plant state, and on the continuous time analog of these equations, which is typically given in (partial) differential form. See Jazwinski (J.1). It is unfortunately the case that the true equations for the propagation of the conditional p.d.f. of the plant state are quite difficult if not impossible to solve in practical cases of interest (See Jazwinski (J.1), Kushner (K.1), (K.2), and Wong (W.3) as particular examples of many works dealing with the problem. See also the monograph "Stochastic Problems in Control" published by the American Automatic Control Conference and issued at the 1968 Joint Automatic Control Conference.)

Two approaches to simplify the propagation of the p.d.f. are in common use. One may approximate the p.d.f. by series expansion with the retention of only "low order" terms (see Wishner, Tabaczynski, and Athans (W.4) and Athans, Wishner, and Bertolini (A.9) as specific examples, and see Jazwinski (J.1) for an extensive discussion with many more references). Alternatively, one can approximate the propagation equations (see Bucy and Senne (B.6)). The former technique has been used for some time (see the many references cited by Jazwinski (J.1)), while the latter technique is relatively new. With regard to these techniques, the following question arises concerning the issues addressed in this thesis: What is the effect of measurement control capability and how can optimal measurement controls be determined?

Consider a typical "extended Kalman filter" used to approximate the propagation of the conditional p.d.f. of the plant state of a nonlinear system corrupted by additive white Gaussian noise. (See Wishner, et al, (W.4)). To use such a device, one first specifies a nominal plant state trajectory and a nominal dynamic control trajectory. One then expands the plant and measurement equations in a Taylor series about these nominals and keeps only linear terms.* One then expands the cost functional about the nominal, keeping terms up to quadratic. The result is an approximation to the original system which has linear dynamics and quadratic cost.

---

* Higher order terms can also be kept, resulting in more complex filters. See Wishner, et al (W.4).

Consider now the effect of a measurement control. Suppose that the following scheme is suggested: One does not assume a nominal measurement control; this parameter remains in the linearized equations in its original form. The result of linearizing about the plant state and the dynamic control, however, is to produce a set of approximating equations which fall into the special case considered in Chapter IV. One can thus determine an optimal open loop measurement control for the linearized equations. This open-loop measurement sequence can be used until a new linearization is made.

The following sub-optimal control technique thus suggests itself for use with nonlinear systems:

(1) Select a nominal dynamic control sequence denoted $\{\breve{\underline{u}}_t\}$. Using this control, obtain the nominal plant state trajectory $\{\breve{\underline{x}}_t\}$ which is the solution of the dynamic equations with noises discarded:

$$\breve{\underline{x}}_{t+1} = \underline{f}_t(\breve{\underline{x}}_t, \breve{\underline{u}}_t) \tag{7.2.1}$$

(Compare to (1.3.1). It is assumed that the driving noise sequence $\{\underline{\xi}_t\}$ has zero mean and simply adds to the dynamic equation.)

(2) Expand the state equation, the measurement equation, and the cost functional about these nominals. Thus, if $\underline{x}_t$ and $\underline{u}_t$ are the true plant state and dynamic controls,

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{u}_t) + \underline{\xi}_t \tag{7.2.2}$$

$$= \underline{f}_t(\hat{\underline{x}}_t, \hat{\underline{u}}_t) +$$

$$\left.\frac{\partial \underline{f}_t}{\partial \underline{x}_t}\right|_{\hat{\underline{x}}_t, \hat{\underline{u}}_t} (\underline{x}_t - \hat{\underline{x}}_t) + \left.\frac{\partial \underline{f}_t}{\partial \underline{u}_t}\right|_{\hat{\underline{x}}_t, \hat{\underline{u}}_t} (\underline{u}_t - \hat{\underline{u}}_t)$$

$$+ \underline{\xi}_t + \text{higher order terms} \tag{7.2.3}$$

$$\underline{y}_t = \underline{g}_t(\underline{x}_t, \underline{v}_t) + \underline{\theta}_t \tag{7.2.4}$$

$$= \underline{g}_t(\hat{\underline{x}}_t, \underline{v}_t) + \left.\frac{\partial \underline{g}_t}{\partial \underline{x}_t}\right|_{\hat{\underline{x}}_t} (\underline{x}_t - \hat{x}_t)$$

$$+ \underline{\theta}_t + \text{higher order terms} \tag{7.2.5}$$

$$J = J(\hat{\underline{x}}, \hat{\underline{u}}) + \left.\frac{\partial J}{\partial \underline{x}}\right|_{\hat{x}, \hat{u}} (\underline{x} - \hat{\underline{x}}) + \left.\frac{\partial J}{\partial \underline{u}}\right|_{\hat{x}, \hat{u}} (\underline{u} - \hat{\underline{u}})$$

$$+ 2(\underline{x} - \hat{\underline{x}})' \frac{\partial^2 J}{\partial \underline{x} \partial \underline{u}} (\underline{u} - \hat{\underline{u}})$$

$$+ (\underline{x} - \hat{\underline{x}})' \frac{\partial^2 J}{\partial \underline{x}^2} (\underline{x} - \hat{\underline{x}})$$

$$+ (\underline{u} - \hat{\underline{u}}) \frac{\partial^2 J}{\partial \underline{u}^2} (u - \hat{u})$$

$$+ \text{higher order terms} \tag{7.2.6}$$

where $\underline{x} \stackrel{\Delta}{=} \{\underline{x}_t\}$, and $\underline{u} \stackrel{\Delta}{=} \{\underline{u}_t\}$

If one defines $\underline{\delta x}_t \triangleq \underline{x}_t - \tilde{\underline{x}}_t$ and $\underline{\delta u}_t \triangleq \underline{u}_t - \tilde{\underline{u}}_t$ to be the deviations of plant state and dynamic control from nominal, one can retain linear terms in (7.2.3) and (7.2.5) and quadratic terms in (7.2.6) to obtain a linear-quadratic problem in $\underline{\delta x}_t$ and $\underline{\delta u}_t$:

$$\underline{\delta x}_{t+1} = \underline{A}_t \underline{\delta x}_t + \underline{B}_t \underline{\delta u}_t + \underline{\xi}_t \tag{7.2.7}$$

$$\underline{\delta y}_t = \underline{C}_t(\underline{v}_t)\, \underline{\delta x}_t + \underline{\theta}_t \tag{7.2.8}$$

$$J = \text{quadratic} \tag{7.2.9}$$

where

$$\underline{A}_t \triangleq \left. \frac{\partial \underline{f}_t}{\partial \underline{x}_t} \right|_{\tilde{x}_t,\, \tilde{u}_t} \tag{7.2.10}$$

$$\underline{B}_t \triangleq \left. \frac{\partial \underline{f}_t}{\partial \underline{u}_t} \right|_{\tilde{x}_t,\, \tilde{u}_t} \tag{7.2.11}$$

$$\underline{C}_t(\underline{v}_t) = \left. \frac{\partial g_t(\underline{x}_t, \underline{v}_t)}{\partial \underline{x}_t} \right|_{\tilde{x}_t} \tag{7.2.12}$$

Note that <u>no</u> nominal measurement control has been hypothesized. If the noises are Gaussian, equations (7.2.7) - (7.2.9) represent precisely the class of problems analyzed in Chapter IV.

(3) Use the results of Chapter IV to generate an optimal measurement control sequence and an optimal dynamic control deviation-from-nominal for the linear-quadratic system.

(4) When appropriate, update the nominal and return to Step 1.

Of course, an important issue is when to update the nominal. This, together with specific results for different types of linearizations, represents a prime area to be considered in extending the results of this thesis to nonlinear systems. It is also important to consider techniques that cannot be formulated according to the structure given above, for example, approximations such as those suggested by Bucy and Senne (B.6).

APPENDIX A

VERIFICATION OF EQUATIONS (4.3.38) - (4.3.42)

Two matrix identities will be useful for the purposes of this appendix. The first is

$$(\underline{I}_n + \underline{X}\,\underline{Y}')^{-1} = \underline{I}_n - \underline{X}(\underline{I}_r + \underline{Y}'\underline{X})^{-1}\underline{Y}' \tag{A.1}$$

where $\underline{X}$ and $\underline{Y}$ are n x r matrices, $\underline{I}_n$ is the n x n identity, and $\underline{I}_r$ is the r x r identity. The other identity is

$$\underline{Y}(\underline{Y} + \underline{X})^{-1} = \underline{I} - \underline{X}(\underline{Y} + \underline{X})^{-1} \tag{A.2}$$

These identities are both taken from Kleinman and Athans (K.5) and may be verified by algebra.

It will now be shown that (4.3.38) - (4.3.42) hold. Starting with the two point boundary value problem

$$\hat{\underline{x}}_{t+1} - \hat{\underline{x}}_t = (\underline{A}_t - \underline{I})\hat{\underline{x}}_t - \underline{B}_t R_t^{-1}\underline{B}_t'\underline{p}_{t+1} + \hat{\underline{\xi}}_t \tag{A.3}$$

$$\underline{p}_{t+1} - \underline{p}_t = -\underline{Q}_t\hat{\underline{x}}_t - (\underline{A}_t - \underline{I})'\underline{p}_{t+1} \tag{A.4}$$

$$\underline{p}_T = \underline{Q}_T\hat{\underline{x}}_T; \qquad \hat{\underline{x}}_0 \text{ given} \tag{A.5}$$

suppose

$$\underline{p}_t = \underline{M}_t\hat{\underline{x}}_t + \underline{h}_t \tag{A.6}$$

Putting (A.6) into (A.3) and (A.4) gives

$$\hat{\underline{x}}_{t+1} - \hat{\underline{x}}_t = (\underline{A}_k - \underline{I})\hat{\underline{x}}_t - \underline{B}_t\underline{R}_t^{-1}\underline{B}_t'\underline{M}_{t+1}\hat{\underline{x}}_{t+1} - \underline{B}_t\underline{R}_t^{-1}\underline{B}_t'\underline{h}_{t+1} + \hat{\underline{\xi}}_t \tag{A.7}$$

$$\underline{M}_{t+1}\hat{\underline{x}}_{t+1} + \underline{h}_{t+1} - \underline{M}_t\hat{\underline{x}}_t - \underline{h}_t = -\underline{Q}_t\hat{\underline{x}}_t - (\underline{A}_t - \underline{I})'\underline{M}_{t+1}\hat{\underline{x}}_{t+1} - (A_t - I)'\underline{h}_{t+1} \tag{A.8}$$

Collecting terms gives

$$(\underline{I} + \underline{B}_t\underline{R}_t^{-1}\underline{B}_t'\underline{M}_{t+1})\hat{\underline{x}}_{t+1} = \underline{A}_t\hat{\underline{x}}_t - \underline{B}_t\underline{R}_t^{-1}\underline{B}_t'\underline{h}_{t+1} + \hat{\underline{\xi}}_t \qquad (A.9)$$

$$\underline{A}_t'\underline{M}_{t+1}\hat{\underline{x}}_{t+1} + \underline{Q}_t\hat{\underline{x}}_t + \underline{A}_t'\underline{h}_{t+1} - \underline{M}_t\hat{\underline{x}}_t - \underline{h}_t = \underline{0} \qquad (A.10)$$

Assumming the indicated inverse exists, (A.9) gives

$$\hat{\underline{x}}_{t+1} = (\underline{I} + \underline{B}_t\underline{R}_t\underline{B}_t'\underline{M}_{t+1})^{-1}[\underline{A}_t\hat{\underline{x}}_t - \underline{B}_t\underline{R}_t^{-1}\underline{B}_t'\underline{h}_{t+1} + \hat{\underline{\xi}}_t] \qquad (A.11)$$

Now using (A.1) and identifying $\underline{X} = \underline{B}_t$, $\underline{Y}' = \underline{R}_t^{-1}\underline{B}_t'\underline{M}_{t+1}$

$$(\underline{I} + \underline{B}_t\underline{R}_t\underline{B}_t'\underline{M}_{t+1})^{-1} = \underline{I} - \underline{B}_t(\underline{I} + \underline{R}_t^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{R}_t^{-1}\underline{B}_t'\underline{M}_{t+1} \qquad (A.12)$$

$$= \underline{I} - \underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1} \qquad (A.13)$$

Substituting (A.13) into (A.11) gives

$$\hat{\underline{x}}_{t+1} = [\underline{I} - \underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}][\underline{A}_t\hat{\underline{x}}_t - \underline{B}_t\underline{R}_t^{-1}\underline{B}_t'\underline{h}_{t+1} + \hat{\underline{\xi}}_t] \qquad (A.14)$$

Substituting (A.14) into (A.10) gives

$$\underline{A}_t'\underline{M}_{t+1}[\underline{I} - \underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}][\underline{A}_t\hat{\underline{x}}_t - \underline{B}_t\underline{R}_t^{-1}\underline{B}_t'\underline{h}_{t+1} + \hat{\underline{\xi}}_t]$$
$$+ \underline{Q}_t\hat{\underline{x}}_t + \underline{A}_t'\underline{h}_{t+1} - \underline{M}_t\hat{\underline{x}}_t - \underline{h}_t = \underline{0} \qquad (A.15)$$

Collecting terms:

$$\{-\underline{M}_t + \underline{A}_t'[\underline{M}_{t+1} - \underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}]\underline{A}_t + \underline{Q}_t\}\hat{\underline{x}}_t$$
$$+ \underline{A}_t'[\underline{M}_{t+1} - \underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}](\hat{\underline{\xi}}_t - \underline{B}_t\underline{R}_t^{-1}\underline{B}_t'\underline{h}_{t+1})$$
$$+ \underline{A}_t'\underline{h}_{t+1} - \underline{h}_t = \underline{0} \qquad (A.16)$$

One now makes the usual argument (see Kleinman and Athans (K.5)) that (A.16) must hold for all initial states $\hat{\underline{x}}_o$ and for all $\{\hat{\underline{\xi}}_t\}$, so that the coefficient of $\hat{\underline{x}}_t$ must itself be zero as well as the sum of the remaining terms. Setting the coefficient of $\hat{\underline{x}}_t$ equal to zero in (A.16) gives

$$\underline{M}_t = \underline{A}_t'[\underline{M}_{t+1} - \underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}]\underline{A}_t + \underline{Q}_t \tag{A.17}$$

which proves (4.3.38). The terminal values $\underline{M}_T = \underline{Q}_T$ and $\underline{h}_T = 0$ follow from (A.5) and (A.6). The equation for $\underline{h}_t$ is obtained by taking the non-$\hat{\underline{x}}_t$ terms in (A.16) and setting them equal to zero, giving

$$\begin{aligned} A_t'[\underline{M}_{t+1} - \underline{M}_{t+1}\underline{B}_t(R_t^{-1} + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}]\hat{\underline{\xi}}_t \\ = \underline{h}_t - \underline{A}_t'\underline{h}_{t+1} + \underline{A}_t'[\underline{M}_{t+1} - \underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}] \\ \underline{B}_t\underline{R}_t^{-1}\underline{B}_t' \quad \underline{h}_{t+1} \end{aligned} \tag{A.18}$$

$$\begin{aligned} = \underline{h}_t - \underline{A}_t'\underline{h}_{t+1} + \underline{A}_t'\underline{M}_{t+1}\underline{B}_t R_t^{-1}\underline{B}_t'\underline{h}_{t+1} \\ - \underline{A}_t'\underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{B}_t R_t^{-1}\underline{B}_t'\underline{h}_{t+1} \end{aligned} \tag{A.19}$$

Examining the last two terms in (A.19) gives

$$\begin{aligned} \underline{A}_t'\underline{M}_{t+1}\underline{B}_t R_t^{-1}\underline{B}_t' - \underline{A}_t\underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{B}_t\underline{R}_t^{-1}\underline{B}_t' \\ = \underline{A}_t'\underline{M}_{t+1}\underline{B}_t\underline{R}_t^{-1}\underline{B}_t' - \underline{A}_t\underline{M}_{t+1}\underline{B}_t(\underline{I} + \underline{R}_t^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{R}_t^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{B}_t\underline{R}_t^{-1}\underline{B}_t' \end{aligned} \tag{A.20}$$

$$= \underline{A}_t'\underline{M}_{t+1}\underline{B}_t[\underline{I} - (\underline{I} + \underline{R}_t^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{R}_t^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{B}_t] R_t^{-1}\underline{B}_t' \tag{A.21}$$

$$= \underline{A}_t'[\underline{I} - \underline{M}_{t+1}\underline{B}_t(\underline{I} + \underline{R}_t^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{R}_t^{-1}\underline{B}_t']\underline{M}_{t+1}\underline{B}_t\underline{R}_t^{-1}\underline{B}_t' \tag{A.22}$$

Now identifying $X = M_{t+1}B_t$ and $Y' = R_t^{-1}B_t'$ and using (A.1), continuing from (A.22) gives

$$\text{left hand side} = A_t'[I - M_{t+1}B_tR_t^{-1}B_t']^{-1}M_{t+1}B_tR_t^{-1}B_t' \qquad (A.23)$$

Identifying $I = X$, $Y = M_{t+1}B_tR_t^{-1}B_t'$ and using (A.2) gives

$$\text{left hand side} = A_t'[I - (I + M_{t+1}B_tR_t^{-1}B_t')^{-1}] \qquad (A.24)$$

Substituting (A.24) into (A.19) and collecting terms gives

$$h_t = (I + M_{t+1}B_tR_t^{-1}B_t')^{-1}h_{t+1} + A_t'[M_{t+1} - M_{t+1}B_t (R_t + B_t'M_{t+1}B_t)^{-1}B_t'M_{t+1}]\hat{\xi}_t \qquad (A.25)$$

which proves (4.3.40).

It only remains to show that (4.3.42) is correct. Starting from (4.3.32) one obtains

$$u_t^* = -R_t^{-1}B_t'p_{t+1} \qquad (A.26)$$

$$= -R_t^{-1}B_t'[M_{t+1}\hat{x}_{t+1} + h_{t+1}] \qquad (A.27)$$

$$= -R_t^{-1}B_t'h_{t+1} - R_t^{-1}B_t'M_{t+1}[I - B_t(R_t + B_t'M_{t+1}B_t)^{-1}B_t'M_{t+1}] [A_t\hat{x}_t - B_tR_t^{-1}B_t'h_{t+1} + \hat{\xi}_t] \qquad (A.28)$$

using (A.14). Continuing

$$\begin{aligned} u_t^* = &-R_t^{-1}[I - B_t'M_{t+1}B_t(R_t + B_t'M_{t+1}B_t)^{-1}]B_t'M_{t+1}A_t\hat{x}_t \\ &- R_t^{-1}[I - B_t'M_{t+1}B_t(R_t + B_t'M_{t+1}B_t)^{-1}]B_t'M_{t+1}\hat{\xi}_t \\ &+ R_t^{-1}[I - B_t'M_{t+1}B_t(R_t + B_tM_{t+1}B_t)^{-1}]B_t'M_{t+1}B_tR_t^{-1}B_t'h_{t+1} \\ &- R_t^{-1}B_t'h_{t+1} \end{aligned} \qquad (A.29)$$

Using (A.2) gives

$$\begin{aligned}
\underline{u}_t^* &= -(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{A}_t\hat{\underline{x}}_t \\
&\quad -(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\hat{\underline{\xi}}_t \\
&\quad -[\underline{I} - (\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{B}_t]\,\underline{R}_t^{-1}\underline{B}_t'\underline{h}_{t+1} \qquad (A.30) \\
&= -(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}(\underline{A}_t\hat{\underline{x}}_t + \hat{\underline{\xi}}_t) \\
&\quad -(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{h}_{t+1} \qquad (A.31)
\end{aligned}$$

having used (A.2). This proves (4.3.42).

APPENDIX B

PROOF OF LEMMA 4.3.1

It will be shown that the value of I_0 given by

$$I_0 = \frac{1}{2}[\underline{\hat{x}}'_T \underline{Q}_T \underline{\hat{x}}_T + \sum_{t=0}^{T-1} \underline{\hat{x}}_t(\underline{Q}_t + \underline{K}'_t\underline{R}_t\underline{K}_t)\underline{\hat{x}}_t] \tag{B.1}$$

subject to

$$\underline{\hat{x}}_{t+1} = (\underline{A}_t - \underline{B}_t\underline{K}_t)\underline{\hat{x}}_t + \underline{\hat{\xi}}_t \tag{B.2}$$

$$\underline{K}_t = (\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}'_t\underline{M}_{t+1}\underline{A}_t \tag{B.3}$$

$$\underline{M}_t = \underline{Q}_t + \underline{A}'_t[\underline{M}_{t+1} - \underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}'_t\underline{M}_{t+1}]\underline{A}_t \tag{B.4}$$

$$\underline{M}_T = \underline{Q}_T \tag{B.5}$$

is

$$I_0 = \frac{1}{2}[\underline{\hat{x}}'_0\underline{M}_0\underline{\hat{x}}_0 + \underline{m}'_0\underline{\hat{x}}_0 + \sum_{t=0}^{T-1}\underline{\hat{\xi}}_t\underline{M}_{t+1}\underline{\hat{\xi}}_t + \underline{m}'_{t+1}\underline{\hat{\xi}}_t] \tag{B.6}$$

where

$$\underline{m}_t = [\underline{A}_t - \underline{B}_t(\underline{R}_t + \underline{B}'_t\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}'_t\underline{M}_{t+1}\underline{A}_t]'[2\underline{M}_{t+1}\underline{\hat{\xi}}_t + \underline{m}_{t+1}] \tag{B.7}$$

$$\underline{m}_T = \underline{0} \tag{B.8}$$

The proof is inductive: a general form for the cost-to-go from time t is derived and this result is then specialized to give the cost to go from time t = 0.

The general fact that will be proved is that the cost-to-go from time t is of the form

$$I_t = \frac{1}{2}[\underline{\hat{x}}_t' \underline{W}_t \underline{\hat{x}}_t + \underline{w}_t' \underline{\hat{x}}_t + c_t] \tag{B.9}$$

where $\underline{W}_t = \underline{W}_t'$.

<u>Step T</u>: The cost-to-go from the terminal step T is given by

$$I_T = \frac{1}{2} \underline{\hat{x}}_T' \underline{Q}_T \underline{\hat{x}}_T \tag{B.10}$$

which is of the form given in (B.9) if one sets

$$\underline{W}_T = \underline{Q}_T \tag{B.11}$$

$$\underline{w}_T = \underline{0} \tag{B.12}$$

$$c_T = 0 \tag{B.13}$$

<u>Induction</u>: Assume that (B.9) is the correct form at time t+1, i.e., that

$$I_{t+1} = \frac{1}{2}[\underline{\hat{x}}_{t+1}' \underline{W}_{t+1} \underline{\hat{x}}_{t+1} + \underline{w}_{t+1}' \underline{\hat{x}}_{t+1} + c_{t+1}] \tag{B.14}$$

Then the cost-to-go from step t is

$$I_t = \frac{1}{2}[\underline{\hat{x}}_t' (\underline{Q}_t + \underline{K}_t' \underline{R}_t \underline{K}_t) \underline{\hat{x}}_t + I_{t+1}(\underline{\hat{x}}_{t+1})] \tag{B.15}$$

Substituting (B.14) and (B.2) into (B.15) gives

$$I_t = \frac{1}{2}\{\underline{\hat{x}}_t' (\underline{Q}_t + \underline{K}_t' \underline{R}_t \underline{K}_t) \underline{\hat{x}}_t + [(\underline{A}_t - \underline{B}_t \underline{K}_t) \underline{\hat{x}}_t + \underline{\hat{\xi}}_t]' \underline{W}_{t+1} [(\underline{A}_t - \underline{B}_t \underline{K}_t) \underline{\hat{x}}_t + \underline{\hat{\xi}}_t]$$
$$+ \underline{w}_{t+1}' [(\underline{A}_t - \underline{B}_t \underline{K}_t) \underline{\hat{x}}_t + \underline{\hat{\xi}}_t] + c_{t+1}\} \tag{B.16}$$

Collecting terms gives

$$I_t = \frac{1}{2}\{\underline{\hat{x}}_t [\underline{Q}_t + \underline{K}_t' \underline{R}_t \underline{K}_t + (\underline{A}_t - \underline{B}_t \underline{K}_t)' \underline{W}_{t+1} (\underline{A}_t - \underline{B}_t \underline{K}_t)] \underline{\hat{x}}_t$$
$$+ [2\underline{\hat{\xi}}_t' \underline{W}_{t+1} (\underline{A}_t - \underline{B}_t \underline{K}_t) + \underline{w}_{t+1}' (\underline{A}_t - \underline{B}_t \underline{K}_t)] \underline{\hat{x}}_t$$
$$+ \underline{\hat{\xi}}_t' \underline{W}_{t+1} \underline{\hat{\xi}}_t + \underline{w}_{t+1}' \underline{\hat{\xi}}_t + c_{t+1}\} \tag{B.17}$$

which is of the form of (B.9) if the following identifications are made:

$$\underline{W}_t = \underline{Q}_t + \underline{K}_t'\underline{R}_t\underline{K}_t + (\underline{A}_t - \underline{B}_t\underline{K}_t)'\underline{W}_{t+1}(\underline{A}_t - \underline{B}_t\underline{K}_t) \tag{B.18}$$

$$\underline{w}_t = [\underline{A}_t - \underline{B}_t\underline{K}_t]'[\underline{w}_{t+1} + 2\underline{W}_{t+1}\hat{\underline{\xi}}_t] \tag{B.19}$$

$$c_t = c_{t+1} + \underline{w}_{t+1}'\hat{\underline{\xi}}_t + \hat{\underline{\xi}}_t'\underline{W}_{t+1}\hat{\underline{\xi}}_t \tag{B.20}$$

This completes the inductive proof that (B.9) is the correct cost-to-to from time t. It is possible to simplify (B.18), however, as follows:

$$\underline{W}_t = \underline{Q}_t + \underline{K}_t'\underline{R}_t\underline{K}_t + (\underline{A}_t - \underline{B}_t\underline{K}_t)'\underline{W}_{t+1}(\underline{A}_t - \underline{B}_t\underline{K}_t) \tag{B.21}$$

$$= \underline{Q}_t + \underline{A}_t'\underline{W}_{t+1}\underline{A}_t + \underline{K}_t'[\underline{R}_t + \underline{B}_t'\underline{W}_{t+1}\underline{B}_t]\underline{K}_t - 2\underline{K}_t'\underline{B}_t'\underline{W}_{t+1}\underline{A}_t \tag{B.22}$$

Substituting (B.3) into (B.22) gives

$$\underline{W}_t = \underline{Q}_t + \underline{A}_t'\underline{W}_{t+1}\underline{A}_t + \underline{A}_t'\underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}(\underline{R}_t + \underline{B}_t'\underline{W}_{t+1}\underline{B}_t)(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{A}_t - 2\underline{A}_t'\underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{W}_{t+1}\underline{A}_t \tag{B.23}$$

If one compares (B.23) and the propagation equation for $\underline{M}_t$, (B.4), one sees that $\underline{M}_t$ and $\underline{W}_t$ will be identical if $\underline{M}_{t+1} = \underline{W}_{t+1}$. Since both $\underline{M}_T$ and $\underline{W}_T$ equal $\underline{Q}_T$ (see (B.5) and (B.11)), it follows that $\underline{M}_{T-1}$. This equality proceeds forward in time iductively, giving

$$\underline{W}_t = \underline{M}_t \qquad \text{for all } t \tag{B.24}$$

where $\underline{M}_t$ satisfies (B.5).

Using (B.24) and comparing (B.19), (B.7), (B.8), and (B.12) gives

$$\underline{w}_t = \underline{m}_t = [\underline{A}_t - \underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{A}_t]'[2\underline{M}_{t+1}\hat{\underline{\xi}}_t + \underline{m}_{t+1}] \tag{B.25}$$

and (B.20) becomes

$$c_t = \underline{m}_{t+1}'\hat{\underline{\xi}}_t + \hat{\underline{\xi}}_t'\underline{M}_{t+1}\hat{\underline{\xi}}_t + c_{t+1} \tag{B.26}$$

It only remains to evaluate $I_0(\hat{\underline{x}}_0)$, the cost-to-go from time $t = 0$. From the general form (B.9), it follows that

$$I_0 = \frac{1}{2}[\hat{\underline{x}}_0'\underline{M}_0\hat{\underline{x}}_0 + \underline{m}_0'\hat{\underline{x}}_0 + c_0] \tag{B.27}$$

The constant c_0 can be evaluated in terms of $\{\hat{\underline{\xi}}_t\}$, $\{\underline{M}_t\}$ and $\{\underline{m}_t\}$ by using (B.26). The result is

$$c_0 = \sum_{t=0}^{T-1} \underline{m}_t'\hat{\underline{\xi}}_t + \hat{\underline{\xi}}_t'\underline{M}_{t+1}\hat{\underline{\xi}}_t \tag{B.28}$$

Thus

$$I_0 = \frac{1}{2}[\hat{\underline{x}}_0'\underline{M}_0\hat{\underline{x}}_0 + \underline{m}_0'\hat{\underline{x}}_0 + \sum_{t=0}^{T-1} \hat{\underline{\xi}}_t'\underline{M}_t\hat{\underline{\xi}}_t + \underline{m}_t'\hat{\underline{\xi}}_t] \tag{B.29}$$

which is exactly (B.6).

APPENDIX C

Let $\{\underline{\Sigma}_{t|t}\}$ be a sequence of symmetric n x n matrices satisfying the following set of equations for t = 0, 1, . . . , T

$$\underline{\Sigma}_{t+1|t} = \underline{A}_t \underline{\Sigma}_{t|t} \underline{A}_t' + \underline{\Xi}_t \tag{C.1}$$

$$\underline{\Sigma}_{t+1|t+1} = \underline{\Sigma}_{t+1|t} - \underline{S}_{t+1}(\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}')\underline{S}_{t+1}' \tag{C.2}$$

$$\underline{\Sigma}_{0|-1} = \underline{\Sigma}_0 = \text{given} \tag{C.3}$$

$$\underline{S}_{t+1} = \underline{\Sigma}_{t+1|t}\underline{C}_{t+1}'(\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}')^{-1} \tag{C.4}$$

Let $\{\underline{M}_t\}$ be a set of symmetric n x n matrices satisfying the following set of equations:

$$\underline{M}_t = \underline{Q}_t + \underline{A}_t'\underline{M}_{t+1}\underline{A}_t - \underline{A}_t'\underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{A}_t \tag{C.5}$$

$$\underline{M}_T = \underline{Q}_T = \text{given}$$

and let $\underline{K}_t$ be defined

$$\underline{K}_t = (\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{A}_t \tag{C.6}$$

The object of this appendix is to show that the following two expressions, $\hat{E}_1$ and $\hat{E}_2$, are equal:

$$\hat{E}_1 \triangleq \operatorname{tr}\left\{\underline{M}_0\underline{\Sigma}_0 + \sum_{t=0}^{T-1}\underline{M}_{t+1}(\underline{\Xi}_t + \underline{B}_t\underline{K}_t\underline{\Sigma}_{t|t}\underline{A}_t')\right\} + \sum_{t=0}^{T}\ell_t(\underline{v}_t) \tag{C.7}$$

$$\hat{E}_2 \triangleq \sum_{t=0}^{T} \ell_t(\underline{v}_t) + \mathrm{tr}\{\underline{Q}_T\underline{\Sigma}_T + \sum_{t=0}^{T-1} \underline{Q}_t\underline{\Sigma}_{t|t}\}$$

$$+ \mathrm{tr}\{\underline{M}_0\underline{S}_0[\underline{C}_0\underline{\Sigma}_0\underline{C}_0' + \underline{D}_0\underline{\Theta}_0\underline{D}_0']\underline{S}_0$$

$$+ \sum_{t=0}^{T-1} \underline{M}_{t+1}\underline{S}_{t+1}[\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}']\underline{S}_{t+1}'\} \tag{C.8}$$

First note that the $\ell_t(\underline{v}_t)$ terms in $\hat{E}_1$ and $\hat{E}_2$ are equal. Thus if $\underline{E}_1$ and $\underline{E}_2$ are obtained from $\hat{\underline{E}}_1$ and $\hat{\underline{E}}_2$ respectively by deleting the $\ell_t(\underline{v}_t)$ terms, it will follow that $\hat{E}_1 = \hat{E}_2$ if and only if $E_1 = E_2$.

Consider E_2:

$$E_2 = \mathrm{tr}\Big\{\underline{Q}_T\underline{\Sigma}_T + \underline{M}_0\underline{S}_0(\underline{C}_0\underline{\Sigma}_0\underline{C}_0' + \underline{D}_0\underline{\Theta}_0\underline{D}_0')\underline{S}_0$$

$$+ \sum_{t=0}^{T-1} \underline{M}_{t+1}\underline{S}_{t+1}(\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}')\underline{S}_{t+1}$$

$$+ \underline{Q}_t\underline{\Sigma}_{t|t}\Big\} \tag{C.9}$$

Using (C.2) and (C.1), calculate

$$\underline{S}_{t+1}(\underline{C}_{t+1}\underline{\Sigma}_{t+1|t}\underline{C}_{t+1}' + \underline{D}_{t+1}\underline{\Theta}_{t+1}\underline{D}_{t+1}')\underline{S}_{t+1}'$$

$$= \underline{\Sigma}_{t+1|t} - \underline{\Sigma}_{t+1|t+1} \tag{C.10}$$

$$= \underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t - \underline{\Sigma}_{t+1|t+1} \tag{C.11}$$

The analogous expression for $t + 1 = 0$ is

$$\underline{S}_0(\underline{C}_0\underline{\Sigma}_0\underline{C}_0' + \underline{D}_0\underline{\Theta}_0\underline{D}_0')\underline{S}_0 = \underline{\Sigma}_0 - \underline{\Sigma}_{0|0} \tag{C.12}$$

(see equation (4.3.5) of Chapter IV).

Substituting (C.11) and (C.12) into (C.9) gives

$$E_2 = \mathrm{tr}\{\underline{Q}_T\underline{\Sigma}_{T|T} + \underline{M}_0\underline{\Sigma}_0 - \underline{M}_0\underline{\Sigma}_{0|0} + \sum_{t=0}^{T-1} \underline{Q}_t\underline{\Sigma}_{t|t} + \underline{M}_{t+1}[\underline{A}_t\underline{\Sigma}_{t|t}\underline{A}_t' + \underline{\Xi}_t - \underline{\Sigma}_{t+1|t+1}]\} \quad (C.13)$$

Using the identity

$$\mathrm{tr}[\underline{A}\,\underline{B}\,\underline{C}] = \mathrm{tr}[\underline{C}\,\underline{A}\,\underline{B}] \quad (C.14)$$

gives

$$E_2 = \mathrm{tr}\{\underline{Q}_T\underline{\Sigma}_{T|T} + \underline{M}_0\underline{\Sigma}_0 - \underline{M}_0\underline{\Sigma}_{0|0} + \sum_{t=0}^{T-1} \underline{Q}_t\underline{\Sigma}_{t|t} + \underline{M}_{t+1}\underline{\Xi}_t + \underline{A}_t'\underline{M}_{t+1}\underline{A}_t\underline{\Sigma}_{t|t} - \underline{M}_{t+1}\underline{\Sigma}_{t+1|t+1}\} \quad (C.15)$$

Now use equation (C.5) to eliminate $\underline{Q}_t$:

$$\underline{Q}_t = \underline{M}_t - \underline{A}_t'\underline{M}_{t+1}\underline{A}_t + \underline{A}_t'\underline{M}_{t+1}\underline{B}_t(\underline{R}_t + \underline{B}_t'\underline{M}_{t+1}\underline{B}_t)^{-1}\underline{B}_t'\underline{M}_{t+1}\underline{A}_t \quad (C.16)$$

$$= \underline{M}_t - \underline{A}_t'\underline{M}_{t+1}\underline{A}_t + \underline{A}_t'\underline{M}_{t+1}\underline{B}_t\underline{K}_t \quad (C.17)$$

Putting this into (C.15) gives

$$E_2 = \mathrm{tr}\Big\{\underline{Q}_T\underline{\Sigma}_{T|T} + \underline{M}_0\underline{\Sigma}_0 - \underline{M}_0\underline{\Sigma}_{0|0} + \sum_{t=0}^{T-1} \underline{M}_t\underline{\Sigma}_{t|t} - \underline{A}_t'\underline{M}_{t+1}\underline{A}_t\underline{\Sigma}_{t|t} + \underline{A}_t'\underline{M}_{t+1}\underline{B}_t\underline{K}_t\underline{\Sigma}_{t|t} + \underline{M}_{t+1}\underline{\Xi}_t + \underline{A}_t'\underline{M}_{t+1}\underline{A}_t\underline{\Sigma}_{t|t} - \underline{M}_{t+1}\underline{\Sigma}_{t+1|t+1}\Big\} \quad (C.18)$$

$$= \mathrm{tr}\{\underline{Q}_T\underline{\Sigma}_{T|T} + \underline{M}_0\underline{\Sigma}_0 - \underline{M}_0\underline{\Sigma}_{0|0} + \sum_{t=0}^{T-1} \underline{M}_t\underline{\Sigma}_{t|t}$$

$$+ \underline{A}_t'\underline{M}_{t+1}\underline{B}_t\underline{K}_t\underline{\Sigma}_{t|t} + M_{t+1}\Xi_t - \underline{M}_{t+1}\underline{\Sigma}_{t+1|t+1}\} \qquad \text{(C.19)}$$

$$= \mathrm{tr}\{\underline{Q}_T\underline{\Sigma}_{T|T} + \underline{M}_0\underline{\Sigma}_0 - \underline{M}_0\underline{\Sigma}_{0|0} + \sum_{t=0}^{T-1} \underline{M}_{t+1}(\Xi_t + \underline{B}_t\underline{K}_t\underline{\Sigma}_{t|t}\underline{A}_t')$$

$$+ \sum_{t=0}^{T-1} \underline{M}_t\underline{\Sigma}_{t|t} - \sum_{t=0}^{T-1} \underline{M}_{t+1}\underline{\Sigma}_{t+1|t+1}\} \qquad \text{(C.20)}$$

Using the fact that $\underline{Q}_T = \underline{M}_T$, (C.20) reduces to

$$E_2 = \mathrm{tr}\{\underline{M}_0\underline{\Sigma}_0 + \sum_{t=0}^{T-1} \underline{M}_{t+1}(\Xi_t + \underline{B}_t\underline{K}_t\underline{\Sigma}_{t|t}\underline{A}_t')\} \qquad \text{(C.21)}$$

This is exactly the same as E_1, as may be seen by comparing to equation (C.7) and recalling that $E_1 = \hat{E}_1 - \sum_{t=0}^{T} \ell_t(\underline{v}_t)$.

Q.E.D.

APPENDIX D

LISTING OF PROGRAMS USED TO EVALUATE EQUATION (5.2.44)

```
    ∇ Z←GAUSS X
      Z←*-((X-MEAN)*2)÷2×COV
    ∇
```

```
    ∇ U JOE V;H;INOV;STD;SIG10;MEAN;COV;UU1;HOLD;I1
      ANS←(U*2)+*XHAT00+SIG00÷2
      SIG10←SIG00+1
      SIG11←SIG10÷1+SIG10×V*2
      MEAN←HOLD←0
      COV←SIG11×SIG10×V*2
      H←0.1×STD←COV*÷2
      INOV←MEAN-5×STD
      →(V=0)/DETERM
      LOOP:→(INOV≥MEAN+5×STD)/DONE
      XHAT11←XHAT00+U+INOV
      UU1←U1
      I1←(*XHAT11+SIG11÷2)+(UU1*2)+¯2×UU1
      HOLD←HOLD+H×I1×(GAUSS INOV)
      INOV←INOV+H
      →LOOP
      DONE:HOLD←HOLD÷STD×(○2)*÷2
      ANS←ANS+HOLD
      ANS
      →0
      DETERM:XHAT11←XHAT00+U
      UU1←U1
      ANS←ANS+(*XHAT11+SIG11÷2)+(UU1*2)+¯2×UU1
      ANS
    ∇
```

```
    ∇ Z←U1;N;C;HIGH;LOW;MID
      C←¯0.5××XHAT11+(SIG11+1)÷2
      N←HIGH←0
      LOW←¯10
      LOOP1:→(LOW≤C××LOW)/LOOP2
      LOW←LOW-10
      HIGH←HIGH-10
      →(N>20)/CHOKE
      N←N+1
      →LOOP1
      CHOKE:UNDEFINED
      LOOP2:→(0.001>HIGH-LOW)/DONE
      N←C××MID←(LOW+HIGH)÷2
      →(0.001>|N-MID)/DONE
      →(MID>N)/DOWN
      LOW←MID
      →LOOP2
      DOWN:HIGH←MID
      →LOOP2
      DONE:Z←MID
    ∇
```

APPENDIX E

PROOF OF THEOREM 6.3.1

Suppose that the optimal cost-to-go from time step t+1 to time step T for the problem defined by equations (6.3.8) - (6.3.13) is $\mathcal{J}_{t+1}$ given by

$$\mathcal{J}_{t+1} = [\hat{\underline{x}}'_{t+1|t} \,\vdots\, \underline{w}'_{t+1}] \underline{M}_{t+1} \begin{bmatrix} \underline{x}_{t+1|t} \\ \hline \underline{w}_{t+1} \end{bmatrix} + I_{t+1}(\underline{\Sigma}_{t+1|t}) \qquad \text{(E.1)}$$

where $\underline{M}_{t+1}$ is a deterministic (n+m) x (n+m) matrix, and $I_{t+1}(\cdot)$ is a deterministic function. Then the cost-to-go from time step t is given by

$$\mathcal{J}_t = E\{\underline{x}'_t \underline{Q}_t \underline{x}_t + \underline{u}'_t \underline{R}_t \underline{u}_t + \ell_t v_t + \mathcal{J}_{t+1} | \tilde{Y}_{t-1}\} \qquad \text{(E.2)}$$

The two cases $v_t = 0$ and $v_t = 1$ will be considered separately, and $\mathcal{J}_t$ will be evaluated and optimized over $\underline{u}_t$ for each case.

<u>Case I:</u> $v_t = 0$

If $v_t = 0$, the Kalman filter equations (6.3.18) and (6.3.19) take the simple form

$$\underline{x}_{t+1|t} = \underline{A}_t \underline{x}_{t|t-1} + \underline{B}_t \underline{w}_t \qquad \text{(E.3)}$$

$$\underline{\Sigma}_{t+1|t} = \underline{\Xi}_t + \underline{A}_t \underline{\Sigma}_{t|t-1} \underline{A}'_t \qquad \text{(E.4)}$$

and the $\underline{w}_t$ - equation (6.3.9) becomes

$$\underline{w}_{t+1} = \underline{w}_t \qquad \text{(E.5)}$$

Substituting (E.3) - (E.5) into (E.1) and then (E.1) into (E.2), and using the partitioned form of $\underline{M}_{t+1}$, one obtains

$$\mathcal{J}_t\Big|_{v_t=0} = E\{\underline{x}_t'\underline{Q}_t\underline{x}_t + \underline{u}_t'\underline{R}_t\underline{u}_t + (\underline{A}_t\hat{\underline{x}}_{t|t-1} + \underline{B}_t\underline{w}_t)'\underline{M}_{t+1}^{11}(\underline{A}_t\hat{\underline{x}}_t + \underline{B}_t\underline{w}_t)$$
$$+ (\underline{A}_t\hat{\underline{x}}_{t|t-1} + \underline{B}_t\underline{w}_t)'\underline{M}_{t+1}^{12}\underline{w}_t + \underline{w}_t'\underline{M}_{t+1}^{21}(\underline{A}_t\hat{\underline{x}}_{t|t-1} + \underline{B}_t\underline{w}_t)$$
$$+ \underline{w}_t'\underline{M}_{t+1}^{22}\underline{w}_t + I_{t+1}(\underline{A}_t\underline{\Sigma}_{t|t-1}\underline{A}_t' + \underline{\Xi}_t) \mid \breve{Y}_{t-1}\} \qquad \text{(E. 5)}$$

Nothing inside the expectation is random except the term $\underline{x}_t'\underline{Q}_t\underline{x}_t$, the conditional expected value of which is $\hat{\underline{x}}_{t|t-1}'\underline{Q}_t\hat{\underline{x}}_{t|t-1} + tr\left[\underline{Q}_t\underline{\Sigma}_{t|t-1}\right]$. Thus $\mathcal{J}_t\Big|_{v_t=0}$ may be expressed as follows:

$$\mathcal{J}_t\Big|_{v_t=0} = \hat{\underline{x}}_{t|t-1}'[\underline{Q}_t + \underline{A}_t'\underline{M}_{t+1}^{11}\underline{A}_t]\hat{\underline{x}}_{t|t-1} + \hat{\underline{x}}_{t|t-1}'\underline{A}_t'(\underline{M}_{t+1}^{11}\underline{B}_t + \underline{M}_{t+1}^{12})\underline{w}_t$$
$$+ \underline{w}_t'(\underline{B}_t'\underline{M}_{t+1}^{11} + \underline{M}_{t+1}^{21})\underline{A}_t\underline{x}_{t|t-1}$$
$$+ \underline{w}_t'(\underline{M}_{t+1}^{22} + \underline{M}_{t+1}^{21}\underline{B}_t + \underline{B}_t'\underline{M}_{t+1}^{12} + \underline{B}_t'\underline{M}_{t+1}^{11}\underline{B}_t)\underline{w}_t$$
$$+ I_{t+1}(\underline{A}_t\underline{\Sigma}_{t|t-1}\underline{A}_t' + \underline{\Xi}_t) + tr\left[\underline{Q}_t\underline{\Sigma}_{t|t-1}\right]$$
$$+ \underline{u}_t'\underline{R}_t\underline{u}_t \qquad \text{(E. 6)}$$

The only term in (E.6) which depends on $\underline{u}_t$ is the last one, and since $\underline{R}_t$ is a positive definite matrix, $\mathcal{J}_t$ is minimized over $\underline{u}_t$ by the choice

$$\underline{u}_t^* = \underline{0} \quad \text{if } v_t = 0 \qquad \text{(E. 7)}$$

and the resulting value of $\mathcal{J}_t$, denoted by $\mathcal{V}_t\Big|_{v_t=0}$, is given by

$$\mathcal{V}_t\Big|_{v_t=0} = [\hat{\underline{x}}_{t|t-1}' \,\vdots\, \underline{w}_t']\underline{M}_t'\begin{bmatrix}\hat{\underline{x}}_{t|t-1} \\ \hline \underline{w}_t\end{bmatrix} + I_t(\underline{\Sigma}_{t|t-1}) \qquad \text{(E. 8)}$$

where

$$\underline{M}_t^{11} = \underline{Q}_t + \underline{A}_t'\underline{M}_{t+1}^{11}\underline{A}_t \tag{E.9}$$

$$\underline{M}_t^{12} = \underline{A}_t'(\underline{M}_{t+1}^{11}\underline{B}_t + \underline{M}_{t+1}^{12}) \tag{E.10}$$

$$\underline{M}_t^{21} = (\underline{M}_t^{12})' \tag{E.11}$$

$$\underline{M}_t^{22} = \underline{M}_{t+1}^{22} + \underline{B}_t'\underline{M}_{t+1}^{12} + \underline{M}_{t+1}^{21}\underline{B}_t + \underline{B}_t'\underline{M}_{t+1}^{11}\underline{B}_t \tag{E.12}$$

$$I_t(\underline{\Sigma}_{t|t-1}) = I_{t+1}(\underline{A}_t\underline{\Sigma}_{t|t-1}\underline{A}_t' + \underline{\Xi}_t) + \mathrm{tr}\left[\underline{Q}_t\underline{\Sigma}_{t|t-1}\right] \tag{E.13}$$

Case II: $v_t = 1$

If $v_t = 1$, the Kalman filter equations (6.3.18) - (6.3.19) take the form

$$\hat{\underline{x}}_{t+1|t} = \underline{A}_t\hat{\underline{x}}_{t|t-1} + \underline{B}_t\underline{w}_t + \underline{S}_t(\underline{y}_t - \underline{C}_t\hat{\underline{x}}_{t|t-1}) \tag{E.14}$$

$$\underline{S}_t = \underline{A}_t\underline{\Sigma}_{t|t-1}\underline{C}_t'(\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{C}_t' + \underline{\Theta}_t)^{-1} \tag{E.15}$$

$$\underline{\Sigma}_{t+1|t} = \underline{\Xi}_t + \underline{A}_t\underline{\Sigma}_{t|t-1}\underline{A}_t' - \underline{A}_t\underline{\Sigma}_{t|t-1}\underline{C}_t'(\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{C}_t' + \underline{\Theta}_t)^{-1}\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{A}_t' \tag{E.16}$$

and the $\underline{w}_t$-equation (6.3.9) is

$$\underline{w}_{t+1} = \underline{u}_t \tag{E.17}$$

Using the partitioned form of $\underline{M}_{t+1}$, the cost-to-go $\mathcal{J}_t\big|_{v_t=1}$ given in equation (E.2) may be expressed as follows:

$$\gamma_t \Big|_{v_t=1} = E\{\underline{x}_t \underline{Q}_t \underline{x}_t + \underline{u}_t' \underline{R}_t \underline{u}_t + \ell_t + \hat{\underline{x}}_{t+1|t} \underline{M}_{t+1}^{11} \hat{\underline{x}}_{t+1|t}$$
$$+ \hat{\underline{x}}_{t+1|t}' \underline{M}_{t+1}^{12} \underline{w}_{t+1} + \underline{w}_{t+1}' \underline{M}_{t+1}^{21} \hat{\underline{x}}_{t+1|t}$$
$$+ \underline{w}_{t+1}' \underline{M}_{t+1}^{22} \underline{w}_{t+1} + I_{t+1}(\underline{\Sigma}_{t+1|t}) | \breve{Y}_{t-1}\} \qquad (E.18)$$

The individual terms in (E.18) will now be evaluated using equations (E.14) - (E.17). The first three terms are

$$E\{\underline{x}_t \underline{Q}_t \underline{x}_t + \underline{u}_t \underline{R}_t \underline{u}_t + \ell_t | \breve{Y}_{t-1}\} = \hat{\underline{x}}_{t|t-1}' \underline{Q}_t \hat{\underline{x}}_{t|t-1} + \mathrm{tr}\left[\underline{Q}_t \underline{\Sigma}_{t|t-1}\right]$$
$$+ \underline{u}_t' \underline{R}_t \underline{u}_t + \ell_t \qquad (E.19)$$

This is correct because $\underline{u}_t$ is deterministic given Y_{t-1}. The fourth term in (E.18) is

$$E\{\hat{\underline{x}}_{t+1|t}' \underline{M}_{t+1}^{11} \hat{\underline{x}}_{t+1|t} | \breve{Y}_{t-1}\}$$
$$= E\{(\underline{A}_t \hat{\underline{x}}_{t|t-1} + \underline{B}_t \underline{w}_t + \underline{S}_t(\underline{y}_t - \underline{C}_t \hat{\underline{x}}_{t|t-1}))\underline{M}_{t+1}^{11}$$
$$(\underline{A}_t \hat{\underline{x}}_{t|t-1} + \underline{B}_t \underline{w}_t + \underline{S}_t(\underline{y}_t - \underline{C}_t \hat{\underline{x}}_{t|t-1})) | \breve{Y}_{t-1}\} \qquad (E.20)$$
$$= \hat{\underline{x}}_{t|t-1}' \underline{A}_t' \underline{M}_{t+1}^{11} \underline{A}_t \hat{\underline{x}}_{t|t-1} + \underline{w}_t' \underline{B}_t' \underline{M}_{t+1}^{11} \underline{B}_t \underline{w}_t$$
$$+ 2\hat{\underline{x}}_{t|t-1}' \underline{A}_t' \underline{M}_{t+1}^{11} \underline{B}_t \underline{w}_t$$
$$+ E\{(\underline{y}_t - \underline{C}_t \hat{\underline{x}}_{t|t-1})' \underline{S}_t' \underline{M}_{t+1}^{11} \underline{S}_t (\underline{y}_t - \underline{C}_t \hat{\underline{x}}_{t|t-1}) | \breve{Y}_{t-1}\} \qquad (E.21)$$

The cross terms in (E.20) vanish in going to (E.21) because the random variable $\underline{y}_t - \underline{C}_t \hat{\underline{x}}_{t|t-1}$ has zero mean given $\breve{Y}_{t-1}$. The last term in (E.21) may be written

$$E\{(\underline{y}_t - \underline{C}_t\underline{x}_{t|t-1})\underline{S}_t'\underline{M}_{t+1}^{11}\underline{S}_t(\underline{y}_t - \underline{C}_t\underline{x}_{t|t-1})|\tilde{Y}_{t-1}\}$$

$$= E\{(\underline{\theta}_t + \underline{C}_t(\underline{x}_t - \underline{x}_{t|t-1}))\underline{S}_t'\underline{M}_{t+1}^{11}\underline{S}_t(\underline{\theta}_t + \underline{C}_t(\underline{x}_t - \underline{x}_{t|t-1}))|\tilde{Y}_{t-1}\} \quad (E.22)$$

$$= \mathrm{tr}\left[\underline{S}_t'\underline{M}_{t+1}^{11}\underline{S}_t(\underline{\Theta}_t + \underline{C}_t\underline{\Sigma}_{t|t-1}\underline{C}_t')\right] \quad (E.23)$$

where the cross terms in (E.22) vanish because $\underline{\theta}_t$ and $(\underline{x}_t - \hat{\underline{x}}_{t|t-1})$ are independent zero mean random vectors.

Substituting (E.15) into (E.23) and (E.23) into (E.21) yields

$$E\{\underline{x}_{t+1|t}'\underline{M}_{t+1}^{11}\underline{x}_{t+1|t}|\tilde{Y}_{t-1}\} = \underline{x}_{t|t-1}'\underline{A}_t'\underline{M}_{t+1}^{11}\underline{A}_t\underline{x}_{t|t-1}$$
$$+ \underline{w}_t'\underline{B}_t'\underline{M}_{t+1}^{11}\underline{B}_t\underline{w}_t + 2\underline{x}_{t|t-1}'\underline{A}_t'\underline{M}_{t+1}^{11}\underline{B}_t\underline{w}_t$$
$$+ \mathrm{tr}[\underline{M}_{t+1}^{11}\underline{A}_t\underline{\Sigma}_{t|t-1}\underline{C}_t'(\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{C}_t' + \underline{\Theta}_t)^{-1}\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{A}_t'] \quad (E.24)$$

The fifth term in (E.18) is

$$E\{\hat{\underline{x}}_{t+1|t}'\underline{M}_{t+1}^{12}\underline{w}_{t+1}|\tilde{Y}_{t-1}\}$$

$$= E\{(\underline{A}_t\hat{\underline{x}}_{t|t-1} + \underline{B}_t\underline{w}_t + \underline{S}_t(\underline{y}_t - \underline{C}_t\hat{\underline{x}}_{t|t-1}))\underline{M}_{t+1}^{12}\underline{u}_t|\tilde{Y}_{t-1}\} \quad (E.25)$$

$$= \hat{\underline{x}}_{t|t-1}'\underline{A}_t'\underline{M}_{t+1}^{12}\underline{u}_t + \underline{w}_t'\underline{B}_t'\underline{M}_{t+1}^{12}\underline{u}_t \quad (E.26)$$

The sixth term in (E.8) is the transpose of the fifth. Since these terms are scalars, this means that the sixth term equals the fifth. The last two terms in (E.18) are not random and hence they may be removed from the expectation operation and expressed in terms of $\underline{u}_t$ and $\underline{\Sigma}_{t|t-1}$ using equations (E.16) and (E.17). Carrying this out and collecting all the terms in (E.18), one finally obtains

$$
\begin{aligned}
\mathcal{J}_t\Big|_{v_t=1} = {} & \hat{x}'_{t|t-1} Q_t \hat{x}_{t|t-1} + \mathrm{tr}\left[Q_t \Sigma_{t|t-1}\right] + u'_t R_t u_t + \ell_t \\
& + \hat{x}'_{t|t-1} A'_t M^{11}_{t+1} A_t \hat{x}_{t|t-1} + w'_t B'_t M^{11}_{t+1} B_t w_t \\
& + 2\hat{x}'_{t|t-1} A'_t M^{11}_{t+1} B_t w_t \\
& + \mathrm{tr}\left[M^{11}_{t+1} A_t \Sigma_{t|t-1} C'_t (C_t \Sigma_{t|t-1} C'_t + \Theta_t)^{-1} C_t \Sigma_{t|t-1} A'_t\right] \\
& + 2(\hat{x}_{t|t-1} A'_t + w'_t B'_t) M^{12}_{t+1} u_t + u'_t M^{22}_{t+1} u_t \\
& + I_{t+1}(\Xi_t + A_t \Sigma_{t|t-1} A'_t - S_t(C'_t \Sigma_{t|t-1} C_t + \Theta_t) S'_t)
\end{aligned}
\qquad \text{(E.27)}
$$

Collecting the terms in this expression which depend on u_t gives

$$
\begin{aligned}
\mathcal{J}_{t,u_t \text{ terms}}\Big|_{v_t=1} = {} & u'_t(R_t + M^{22}_{t+1}) u_t \\
& + 2(\hat{x}'_{t|t-1} A'_t + w_t B'_t) M^{12}_{t+1} u_t
\end{aligned}
\qquad \text{(E.28)}
$$

The value of u_t which minimizes this expression is given by

$$
u^*_t = -(R_t + M^{22}_{t+1})^{-1} M^{21}_{t+1} (A_t \hat{x}_{t|t-1} + B_t w_t) \qquad \text{(E.29)}
$$

where the fact that $(M^{12}_{t+1})' = M^{21}_{t+1}$ has been used. Substituting this value back into (E.27) to obtain the minimal value of $\mathcal{J}_t$ when $v_t = 1$, denoted $\mathcal{V}_t\Big|_{v_t=1}$, results in

$$
\begin{aligned}
\mathcal{V}_t\Big|_{v_t=1} = {} & \hat{x}'_{t|t-1}[Q_t + A'_t M^{11}_{t+1} A_t] \hat{x}_{t|t-1} \\
& + w'_t B'_t M^{11}_{t+1} B_t w_t + \hat{x}_{t|t-1} A'_t M^{11}_{t+1} B_t w_t \\
& + w'_t B'_t M^{11}_{t+1} A_t \hat{x}_{t|t-1} + \ell_t + \mathrm{tr}\left[Q_t \Sigma_{t|t-1}\right] \\
& + \mathrm{tr}\left[M^{11}_{t+1} A_t \Sigma_{t|t-1} C'_t (C_t \Sigma_{t|t-1} C'_t + \Theta_t)^{-1} C_t \Sigma_{t|t-1} A'_t\right] \\
& + I_{t+1}(\Xi_t + A_t \Sigma_{t|t-1} A'_t - S_t(C_t \Sigma_{t|t-1} C'_t + \Theta_t) S'_t)
\end{aligned}
$$

$$- (\underline{A}_t\hat{\underline{x}}_{t|t-1} + \underline{B}_t\underline{w}_t)'\underline{M}^{12}_{t+1}(\underline{R}_t + \underline{M}^{22}_{t+1})^{-1}\underline{M}^{21}(A_t\hat{\underline{x}}_{t|t-1} + \underline{B}_t\underline{w}_t)$$

(E.30)

This may be expressed in the form

$$\mathcal{V}_t\Big|_{v_t=1} = \left[\hat{\underline{x}}'_{t|t-1} \vdots \underline{w}'_t\right]\underline{M}_t\begin{bmatrix}\hat{\underline{x}}_{t|t-1}\\ \hline \underline{w}_t\end{bmatrix} + I_t(\underline{\Sigma}_{t|t-1}) \qquad (E.31)$$

where the following equations hold for the partioned sections of $\underline{M}_t$ and for $I_t(\Sigma_{t|t-1})$:

$$\underline{M}^{11}_t = \underline{Q}_t + \underline{A}'_t(\underline{M}^{11}_{t+1} - \underline{M}^{12}_{t+1}(\underline{R}_t + \underline{M}^{22}_{t+1} + \underline{M}^{22})^{-1}\underline{M}^{12})\underline{A}_t \qquad (E.32)$$

$$\underline{M}^{12}_t = \underline{A}'_t(\underline{M}^{11}_{t+1} - \underline{M}^{12}_{t+1}(\underline{R}_t + \underline{M}^{22}_{t+1})^{-1}\underline{M}^{21}_{t+1})\underline{B}_t \qquad (E.33)$$

$$\underline{M}^{21}_t = (\underline{M}^{12}_t)' \qquad (E.34)$$

$$\underline{M}^{22}_t = \underline{B}'_t(\underline{M}^{11}_{t+1} - \underline{M}^{12}_{t+1}(\underline{R}_t + \underline{M}^{22}_{t+1})^{-1}\underline{M}^{21}_{t+1})\underline{B}_t \qquad (E.35)$$

$$\begin{aligned} I_t(\underline{\Sigma}_{t|t}) = {} & \ell_t + tr\left[\underline{Q}_t\underline{\Sigma}_{t|t-1}\right] \\ & + tr\left[\underline{M}^{11}_{t+1}\underline{A}_t\underline{\Sigma}_{t|t-1}\underline{C}'_t(\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{C}'_t + \underline{\Theta}_t)^{-1}\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{A}'_t\right] \\ & + I_{t+1}(\underline{\Xi}_t + \underline{A}_t(\underline{\Sigma}_{t|t-1} - \underline{\Sigma}_{t|t-1}\underline{C}'_t(\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{C}'_t + \underline{\Theta}_t)^{-1}\underline{C}_t\underline{\Sigma}_{t|t-1})\underline{A}'_t) \end{aligned}$$

(E.36)

This completes the analysis of the case $v_t = 1$.

The technique for determining the values of v_t and $\underline{u}_t$ to minimize $\mathcal{J}_t$, given by (E.2), subject to (E.1) may thus be summarized as follows:

(1) Calculate $\mathcal{V}_t\Big|_{v_t=0}$ given by equations (E.8) through (E.13).

(2) Calculate $\vartheta_t\big|_{v_t=1}$ given by equations

(E.31) through (E.36).

(3) If $\vartheta_t\big|_{v_t=0} < \vartheta_t\big|_{v_t=1}$, set $v_t^* = 0$. If $\vartheta_t\big|_{v_t=1} < \vartheta_t\big|_{v_t=0}$,

set $v_t^* = 1$.

(4) Calculate u_t^* as follows:

$$\underline{u}_t^* = -v_t^*(\underline{R}_t + \underline{M}_{t+1}^{22})^{-1}\underline{M}_{t+1}^{21}(\underline{A}_t\hat{\underline{x}}_{t|t-1} + \underline{B}_t\underline{w}_t) \tag{E.37}$$

(5) Express the cost-to-go from step t as follows:

$$\vartheta_t = \left[\hat{\underline{x}}'_{t|t-1} \vdots \underline{w}'_t\right]\underline{M}_t\begin{bmatrix}\underline{x}_{t|t-1}\\ \hline \underline{w}_t\end{bmatrix} + I_t(\underline{\Sigma}_{t|t-1}) \tag{E.38}$$

Where $\underline{M}_t$ and $\underline{I}_t$ are specified by:

$$\underline{M}_t^{11} = \underline{Q}_t + \underline{A}_t'\underline{M}_{t+1}^{11}\underline{A}_t - v_t^*\underline{A}_t'\underline{P}_t\underline{A}_t \tag{E.39}$$

$$\underline{M}_t^{12} = \underline{A}_t'\underline{M}_{t+1}^{11}\underline{B}_t - v_t^*\underline{A}_t'\underline{P}_t\underline{B}_t + (1-v_t^*)\underline{A}_t'\underline{M}_{t+1}^{12} \tag{E.40}$$

$$\underline{M}_t^{21} = (\underline{M}_t^{12})' \tag{E.41}$$

$$\underline{M}_t^{22} = \underline{B}_t'\underline{M}_{t+1}^{11}\underline{B}_t - v_t^*\underline{B}_t'\underline{P}_t\underline{B}_t + (1-v_t^*)\left(\underline{M}_{t+1}^{22} + \underline{B}_t'\underline{M}_{t+1}^{12} + \underline{M}_{t+1}^{21}\underline{B}_t\right) \tag{E.42}$$

$$\underline{P}_t \triangleq \underline{M}_{t+1}^{12}(\underline{R}_t + \underline{M}_{t+1}^{22})^{-1}\underline{M}_{t+1}^{21} \tag{E.43}$$

$$\begin{aligned} I_t = {} & \ell_t v_t^* + \mathrm{tr}\left[\underline{Q}_t\underline{\Sigma}_{t|t-1}\right] \\ & + v_t^* \cdot \mathrm{tr}\left[\underline{M}_{t+1}^{11}\underline{A}_t\underline{\Sigma}_{t|t-1}\underline{C}_t'(\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{C}_t' + \underline{\Theta}_t)^{-1}\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{A}_t'\right] \\ & + I_{t+1}(\underline{\Xi}_t + \underline{A}_t\underline{\Sigma}_{t|t-1}\underline{A}_t' - v_t^*\underline{A}_t\underline{\Sigma}_{t|t-1}\underline{C}_t'(\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{C}_t' + \underline{\Theta}_t)^{-1}\underline{C}_t\underline{\Sigma}_{t|t-1}\underline{A}_t') \end{aligned} \tag{E.44}$$

Equations (E.38) through (E.44) were obtained by collecting the sets of equations (E.8) - (E.13) and (E.31) - (E.36) and including certain factors of v_t^* (which are either zero or one) which serve to specialize the all-inclusive set of equations (E.38) - (E.44) as appropriate.

REFERENCES

[A.1] M. Aoki, Optimization of Stochastic Systems, Academic Press, New York, 1967.

[A.2] K.J. Astrom, Introduction to Stochastic Control Theory, Academic Press, New York, 1970.

[A.3] M. Athans, "The Role of the Linear, Quadratic, Gaussian Problem in Control System Design" IEEE Transactions on Automatic Control, December, 1971 (to appear).

[A.4] M. Athans and F.C. Schweppe, "Optimal Waveform Design via Control Theoretic Concepts", Info. and Control, Vol. 10, no. 4, April, 1967, pp. 335-377.

[A.5] M. Aoki and M.T. Li, "Optimal Discrete-Time Control System with Cost for Observation", IEEE Transactions on Automatic Control, Vol. AC-14, April, 1969, pp. 165-175.

[A.6] M. Athans, "Optimal Measurement Strategies for Linear Stochastic Systems", M.I.T. Lincoln Laboratory Technical Note 1971-12, February 26, 1971.

[A.7] M. Athans and P.L. Falb, Optimal Control, McGraw-Hill, New York, 1966.

[A.8] M. Athans, "The Matrix Minimum Principle", Info. and Control, Vol. 11, No. 5-6, pp. 592-606, November-December, 1967.

[A.9] M. Athans, R.P. Wishner, and A. Bertolini, "Suboptimal State Estimation for Continuous Time Nonlinear Systems from Discrete Noisy Measurements". IEEE Trans. on Auto. Control, Vol. AC-13, pp. 504-514, 1968.

[B.1] D.P. Bertsekas, "Control of Uncertain Systems with a Set-Membership Description of Uncertainty" Ph.D. Thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering, May, 1971.

[B.2] D.P. Bertsekas and I.B. Rhodes, "Recursive State Estimation for a Set-Membership Description of the Uncertainty", IEEE Transactions on Automatic Control, Vol. AC-16, No. April, 1971, pp. 117-127.

[B.3] R. Bellman, Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

[B.4] A.E. Bryson and Y.C. Ho, Applied Optimal Control, Blaisdell Pub. Co., Waltham, Mass, 1969.

[B.5] R.W. Brockett, Finite Dimensional Linear Systems, J. Wiley and Sons, New York, 1970.

[B.6] R.S. Bucy and K.D. Senne, "Digital Synthesis of Nonlinear Filters", Automatica, Vol. 7, No. 3, May, 1971, pp.287-298.

[C.1] C.A. Cooper and N.E. Nahi, "An Optimal Stochastic Control Problem with Observation Cost". 1970 Joint Automatic Control Conference, Atlanta, Georgia, 22-26 June 1970, paper 27-C. pp. 723-728.

[C.2] R. Curry, "Estimation and Control with Quantized Measurements". M.I.T. Press, Cambridge, Mass. 1970.

[C.3] T.J. Cruise, "Communication Utilizing Feedback Channels" Ph.D. Thesis, M.I.T. Dept. of Electrical Engineering, January, 1969.

[D.1] J.L. Doob, Stochastic Processes, Wiley, New York, 1953.

[D.2] S.E. Dreyfus, Dynamic Programming and the Calculus of Variations, Academic Press, 1965.

[F.1] P.A. Frost and T. Kailath, "An Innovations Approach to Least-Squares Estimation - Part III: Nonlinear Estimation in White Gaussian Noise", IEEE Trans. on Auto. Control, Vol. AC-16, No. 3, June, 1971, pp. 217-226.

[J.1] A.H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.

[K.1] H.J. Kushner, Stochastic Stability and Control, Academic Press, New York, 1967.

[K.2] H.J. Kushner, Introduction to Stochastic Control, Holt, Reinhart and Winston, 1971.

[K.3] H. J. Kushner, On the Optimal Timing of Observations for Linear Systems with Unknown Initial States", IEEE Trans. on Auto. Control, Vol. AC-9, 1964, pp. 144-155.

[K.4] R.E. Kalman, "A New Approach to Linear Filtering and Prediction Problems", A.S.M.E.J. of Basic Eng., ser. D, Vol. 82, March, 1960, pp. 35-45.

[K.5] D.L. Kleinman and M. Athans, "The Discrete Minimum Principle with Application to the Linear Regulator Problem", M.I.T. Electronic Systems Laboratory, report ESL-R-260, Feburary, 1966.

[K.6] T. Kailath, "An Innovations Approach to Least-Squares Estimation - Part I: Linear Filtering in Additive White Noise", IEEE Trans. on Auto. Control, Vol. AC-13, December, 1968 pp. 646-654.

[K.7] T. Kailath and P. Frost, "An Innovations Approach to Least Squares Estimation - Part II: Linear Smoothing in Additive White Noise", IEEE Trans. on Auto. Control, Vol. AC-13, December, 1968, pp.655-660.

[K.8] T. Kailath and P. Frost, "Mathematical Modeling of Stochastic Processes", in Stochastic Problems in Control, Symposium of A.A.C.C., at the 1968 J.A.C.C., Ann Arbor, Michigan, 1968.

[L.1] M. Loeve, Probability Theory, Van Nostrand, Princeton, 1961.

[M.1] O. Mayr, "The Origins of Feedback Control", Scientific American, Vol. 223, No. 4, October, 1970, pp. 110-118.

[M.2] L. Meier, J. Peschon, and P.M. Dressler, "Optimal Control of Measurement Subsystems", IEEE Trans. on Auto. Control, Vol. AC-12, 1967, pp. 528-536.

[M.3] Meier, R.E. Larson and T.J. Tether, "Dynamic Programming for Stochastic Control of Discrete Systems" to appear in IEEE Trans. on Auto. Control, Special Issue on Linear-Gaussian-Quadratic Problem, December 1971.

[O.1] J.K. Omura, "Optimal Linear Transmission of Analog Data for Channels with Feedback", IEEE Trans. IT. IT-14, No. 1 January, 1968, pp. 38-43.

[P.1] A. Papoulis, Probability, Random Variables, and Stochastic Processes, MeGraw-Hill, New York, 1965.

[P.2] L.S. Pontryagin, V. Boltyanskii, R. Gankrelidze, and E. Mishchenko, The Mathematical Theory of Optimal Processes, Interscience Publishers, Inc., New York, 1962.

[R.1] I.B. Rhodes, Notes for course 6.214T, "Dynamic Systems, Control, and Optimization, M.I.T. Dept of Electrical Engineering, Spring, 1969. (to appear as a text).

[S.1] F.C. Schweppe, "Recurssive State Estimation; Unknown but Bounded Errors and System Inputs", IEEE Transactions on Automatic Control, Vol. AC-13, No. 1, January, 1968. pp. 22-28.

[S.2] F. C. Schweppe, Uncertain Dynamic Systems, Academic Press, New York, N.Y., (to appear).

[S.3] M.K. Sain, "Control of Linear Systems Accoring to the Minimal Variance Criterion - A New Approach to the Disturbance Problem" IEEE Transactions on Automatic Control, Vol. AC-11, No. 1, January, 1966, pp. 118-121.

[S.4] M.K. Sain and S.R. Liberty, "Some New Computational Results for Minimum Variance Controllers", Proc. of the Seventh Allerton Conf. on Circuits and Systems Theory, October, 1969.

[S.5] A. Sano and M. Terao, "Measurement Optimization in Optimal Process Control", Automatica, Vol. 6, No. 5, September, 1970, pp. 705-714.

[S.6] C. Striebel, "Sufficient Statistics in the Optimum Control of Stochastic Systems", J. of Math. Analy. and Appl., Vol. 12, 1965, pp. 576-592.

[S.7] H.A. Spang, Optimal Control of an Unknown Linear Plant Using Baysian Estimation of Error". IEEE Trans. on Auto. Control, Vol. AC-10, January, 1965, pp. 80-83.

[S.8] J.P.M. Schalkwijk and T. Kailath, "A Coding Scheme for Additive Noise Channels with Feedback" Part I: No Band with Constraint, IEEE Tr. I.T. Vol. IT-12, April, 1966, pp. 172-182.

[S.9] J.P.M. Schakwijk, "A Coding Scheme for Additive Noise Channels with Feedback - Part II: Band-limited Signals", IEEE Trans. IT Vol. IT-12, April, 1966, pp. 183-189.

[T.1] E.T.S. Tse, "On the Optimal Control of Linear Systems with Incomplete Information", Ph.D. Thesis, M.I.T., Dept. of Electrical Engineering, November, 1969.

[T.2] E. Tse and M. Athans, "Adaptive Stochastic Control for Linear Systems, Part I - Solution Methods, Part II - Assymptotic Properties and Simmulation Results", Proc. Decision and Control Conference, University of Texas, Austin, Texas, December, 1970.

[W.1] H.S. Witsenhausen, "Separation of Estimation and Control for Discrete Time Systems, to appear in Proceeding of IEEE, December, 1971.

[W.2] W.M. Wonham, "On the Separation Theorem of Stochastic Control", SIAM J. Control, Vol. 6, No. 2, 1968, pp. 312-326.

[W.3] E. Wong, Stochastic Processes in Information and Dynamical Systems, McGraw-Hill, New York, 1971.

[W.4] R.P. Wishner, J.A. Tabaczynski, and M. Athans, "A Comparison of three Non-linear Filters", Automatica, Vol. 5, pp. 487-496, 1969.

[Z.1] L.A. Zadeh, "Fuzzy Sets", Info. and Control, Vol. 8, 1965, pp. 338-353.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SCIENCE AND ARTS

1861

MENS ET MANUS