June 1976

Report ESL-R-665

NASA Grant NGL-22-009-124

(NASA-CR-149670) CONSIDERATION OF COMPUTER LIMITATIONS IN IMPLEMENTING ON-LINE CONTROLS M.S. Thesis (Massachusetts Inst. of Tech.) 70 p HC A04/MF A01 CSCL 09B

N77-18761

Unclas

G3/60 20452

CONSIDERATION OF COMPUTER LIMITATIONS IN IMPLEMENTING ON-LINE CONTROLS

Gary K Roberts

Electronic Systems Laboratory *Decision and Control Sciences Group*

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139

Department of Electrical Engineering and Computer Science

June, 1976 Report ESL-R-665

CONSIDERATION OF COMPUTER LIMITATIONS IN IMPLEMENTING ON-LINE CONTROLS

by

Gary K. Roberts

This report is based on the unaltered thesis of Gary Kenneth Roberts, submitted in partial fulfillment of the requirements for the degrees of Master of Science and Electrical Engineer at the Massachusetts Institute of Technology in June, 1976. This research was conducted at the Decision and Control Sciences Group of the M.I.T. Electronic Systems Laboratory with support provided by the NASA Ames Research Center under grant NGL-22-009-124.

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

CONSIDERATION OF COMPUTER LIMITATIONS IN

IMPLEMENTING ON-LINE CONTROLS

by

Gary Kenneth Roberts

SB, Massachusetts Institute of Technology

(1973)

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREES OF

MASTER OF SCIENCE

AND

ELECTRICAL ENGINEER

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June, 1976

Signature of Author Gary Kenneth Roberts ....
Department of Electrical Engineering
and Computer-Science, May 7, 1976

Certified by .... Alan S. Willsky ....
Thesis Supervisor

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

CONSIDERATION OF COMPUTER LIMITATIONS IN

IMPLEMENTING ON-LINE CONTROLS

by

Gary Kenneth Roberts

Submitted to the Department of Electrical Engineering and Computer Science on May 7, 1976, in partial fulfillment of the requirements for the Degrees of Master of Science and Electrical Engineer.

ABSTRACT

The digital computer has become an important tool in the on-line optimal control of large systems. Because computers are discrete-time in nature, it is necessary to select an interval of discretization when controlling a continuous-time system; this introduces another parameter over which to optimize.

In this thesis, we discuss how the characteristics of a particular computer, such as multiplication time, will determine the range of choice for this parameter, and how such a choice will affect system performance. We also discuss how a suboptimal control, involving less computation, will allow us to control the system at more frequent intervals and thus, in some cases, lead to better results than if we used the usual optimal control.

We begin by making a formal statement of the optimal control problem which includes the interval of discretization as an optimization parameter, and extend this to include selection of a control algorithm as part of the optimization procedure. We show how the performance of the scalar linear system depends on the discretization interval. We go on to develop discrete-time versions of the output feedback regulator and an optimal compensator, and use these results in presenting an example of a system for which fast partial-state feedback control better minimizes a quadratic cost than either a full-state feedback control or a compensator.

Thesis Supervisor: Alan S. Willsky

Title: Assistant Professor of Electrical Engineering

ACKNOWLEDGMENTS

I would herein like to thank Professor Alan S. Willsky, my Advisor and Thesis Supervisor, for his help and patience until the bitter end.

Thanks also are due to Professors Tim Johnson and Nils Sandell, and my fellow students, particularly Steve Young and the ubiquitous Wolf Kohn, for their insights and suggestions. I am especially indebted to Art Giordani of ESL Drafting for the figures, ESL Accounting and Purchasing for their tea, and Barbara Peacock of the ESL for her sympathy. I won't even mention the inspired typing of Susanna Natti.

I would also like to acknowledge the NASA/Ames Research Center under grant NGL-22-009-124 for providing me with the computer support for this research effort.

Table of Contents

Chapter I

Introduction

1.1 Motivation

The digital computer has become an essential tool in the solution of practical optimal control problems. Even "simple" lower-order systems require an inordinate amount of computation (such as in the solutions of the Riccati and Kalman Filter equations in [1] and [4]) in order to achieve the optimal trajectory. For a large-scale system, the problems are even greater. In particular, it may be desirable to use the computer "on-line" to control the system, that is, to include the computer in the feedback loop itself.

Naturally, an arrangement of this kind gives rise to new problems and limitations, stemming directly from the special nature of the computer itself. For example, let us consider the deterministic linear regulator problem [3], in which the system is described by the equations

$$\dot{\underline{x}}(t) = \underline{F}(t)\underline{x}(t) + \underline{G}(t)\underline{u}(t) \tag{1.1}$$

$$\underline{y}(t) = \underline{H}(t)\underline{x}(t), \tag{1.2}$$

in which we wish to determine

$$\underline{u}(t) = \underline{g}(\underline{y}(t)) \tag{1.3}$$

in order to minimize a given functional $J(\underline{x}(\cdot), t_0, t_1)$, where $[t_0, t_1]$ represents some interval of time. Often, the functional form of $g(\cdot)$

in (1.3) will also be restricted. In any case, the state $\underline{x}(t)$ of the system will have a constantly changing value, and consequently, so will the optimal control $\underline{u}^*(t)$. This cannot be reconciled with the discrete-time nature of the computer; the machine can in no way continuously monitor the output $\underline{y}(t)$, nor can it continuously change the input.

The usual solution to this problem is to <u>sample</u> the plant output at a certain select interval Δ of time, and similarly, calculate a control at intervals, using an external device to hold the actual input constant over each period of time. This is illustrated in Figure 1.

In a sense, what we need to do is construct a discrete-time system

$$\underline{x}_{k+1} = \underline{\Phi}(k)\underline{x}_k + \underline{M}(k)\underline{u}_k \tag{1.4}$$

$$\underline{y}_k = \underline{C}(k)\underline{x}_k \tag{1.5}$$

for which the behavior of $\underline{x}_k$ is in some sense "close" to the behavior of $\underline{x}(t)$ when the value of $\underline{x}_k$ is compared directly with that of $\underline{x}(k\Delta)$. A system in which the states of (1.1) and (1.5) are <u>equal</u> at these points is in fact determined by Levis [10].

It should not be surprising, though, that the selection of the sampling interval Δ will affect the performance of our on-line control. This thesis will discuss ways in which the best choice for Δ depends on the characteristics of the computer facility upon which the control is implemented, and how this dependence in turn affects the choice of a control algorithm.

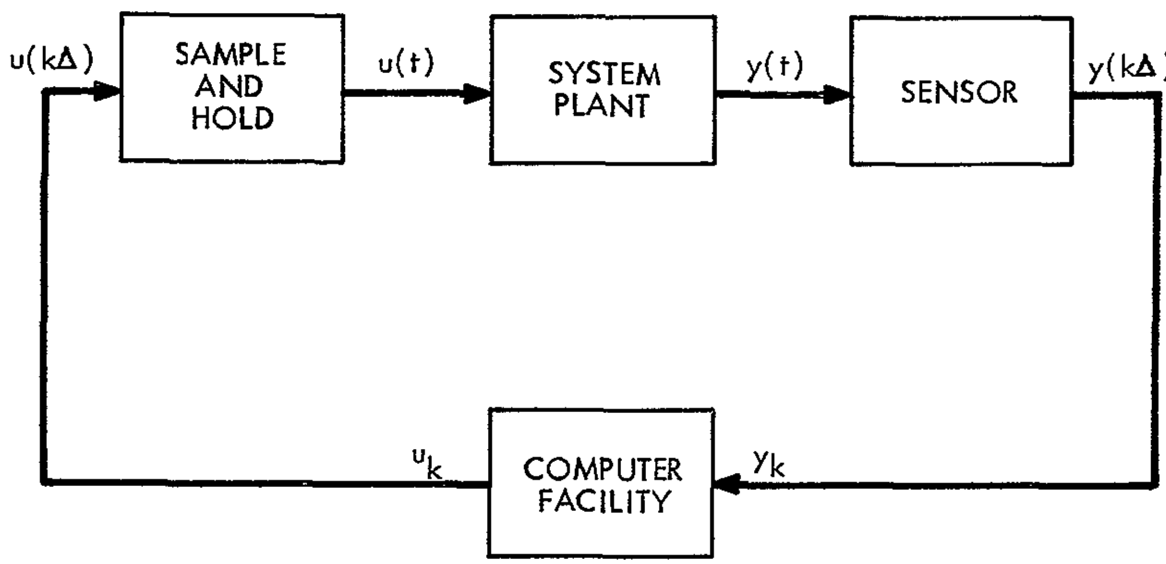

Fig.1 Inclusion of the Computer in the Feedback Loop

1.2 Optimization on the Digital Computer

When an optimal control for a discrete-time system is determined on-line by a computer, there is a certain computation time associated with the calculation of each step. Thus a certain minimal interval of discretization can be associated with a given control algorithm. For example, suppose a computer takes τ seconds to perform any simple computation (like multiplication), and for a particular algorithm, n simple computations are needed to compute the optimal control. If we use Δ to symbolize the interval of discretization for a discrete-time model, then for the aforementioned algorithm, the minimum Δ that is physically possible is $\Delta_{min} = n\tau$ seconds. If we denote the control algorithm as i, the possible range for Δ is $[\Delta_{min(i)}, \infty)$.

For the remainder of this work, we will restrict our attention to time-invariant systems, and consider only steady-state controls. This gives us the advantage of being able to associate with each algorithm a very simple parameter set. For example, in such a case, the feedback matrix given by the familiar Riccati equation becomes an algebraic equation with a constant matrix solution, rather than one that is a function of time.

Let us state this in a precise manner. Henceforth, for convenience, we will no longer underline vectors and matrices. The reader should assume all upper case Greek and Roman letters to be matrices and all lower case letters to be vectors, except for units of time (such as Δ or t), indices (i, j, k), or where otherwise specified.

We will consider deterministic systems of the form

$$\dot{x}(t) = f(x(t), u(t), t) \tag{1.6}$$

$$y(t) = h(x(t), t); \tag{1.7}$$

a feedback control algorithm i will be of the form

$$u(t) = g_i(z(t_k); K), \tag{1.8}$$

for $t_k \leq t < t_{k+1}$, where $\Delta = t_{k+1} - t_k$, K is the set of constant parameters that we are to optimize over in the control law, and $z(t_k)$ is the information that may be obtained from y(t), for $t \leq t_k$.

Now let $\mathcal{A}$ be an index set of control laws. We can characterize the usual optimization problem [10] as follows. If we are given $i \in \mathcal{A}$, and a particular $\Delta \in [\Delta_{min(i)}, \infty)$, find the K that minimizes a given cost functional $J(i, \Delta, K)$, where K is a parameter set as mentioned above.

Using this format to state the problem suggests a generalization that will be the crux of this thesis. In the case of the linear system with quadratic cost, where the control is based on a full state observation, and is a linear feedback law, Levis [10] has raised the question of how to determine the optimal interval of discretization for a given control algorithm. He has pointed out that for some systems, costs <u>do not</u> decrease monotonically as Δ decreases (see Chapter II for a further discussion); in fact, he was unable to determine the optimal Δ for the problem he considered. Let us state this question as follows: suppose we are able to minimize $J(i, \Delta, K)$ by finding a suitable $K = K^*$, as in the problem just mentioned; since this is for a particular i and Δ, let us then say that $K^* = K^*(i, \Delta)$. In other words, for a given i and Δ, there is a "best" K, $K^*(i, \Delta)$. Then for a given $i \in \mathcal{I}$, we wish to find the $\Delta \in [\Delta_{\min(i)}, \infty)$ that minimizes

$$J(i, \Delta, K^*(i, \Delta)). \tag{1.9}$$

Now since $\Delta_{\min(i)}$ is certainly a function of i alone (for a given computer), the Δ that minimizes equation (1.9) is also a function of i, that is, $\Delta = \Delta_{f(i)}$ (as we shall see in Chapter II, it is <u>not</u> always true that $\Delta_{f(i)} = \Delta_{\min(i)}$). Then the crucial question becomes: what is the control algorithm i that minimizes the cost? In other words, find the $i \in \mathcal{I}$ that minimizes

$$J(i, \Delta_{f(i)}, K^*(i, \Delta_{f(i)})). \tag{1.10}$$

At this time, there does not appear to be a neat general approach to the solution of this problem. The purpose of this work is to provide an impetus and suggest a start for investigation along this line.

1.3 The Scope of this Work

Because this area is extremely broad, too broad, in fact, to develop a unified theory, we will be content in this document mainly to illustrate some of the points mentioned in the previous subsections. In Chapter II, we will discuss the selection of the optimal sampling rate for a full-state feedback law, and present, as an example, a simple class of systems for which the optimal Δ can be solved for explicitly. In Chapters III and IV, we will derive the sampled-data versions of two alternative control laws, expressing each as a function of the sampling rate Δ. In particular, Chapter III will deal with partial-state linear feedback controls, while Chapter IV will give results on optimal reduced-order compensators. In Chapter V, we will give a specific example of a system for which, due to our greater freedom of choice in selecting Δ that results from computational considerations, we actually are able to decrease a system cost by choosing a "suboptimal" control algorithm over an "optimal" one.

Chapter II

On Choosing the Optimal Δ for a Given Algorithm: The Scalar Case

For a given control algorithm, it is clear that the continuous controller (i.e., $\Delta \to 0$) will minimize the system cost better than any particular discrete-time controller. Thus it might appear that the problem of selecting an optimizing Δ for a computer control is trivial, that the best choice is $\Delta = \Delta_{min}$. As Levis [10] has shown, this may not always be the case. For the second order system

$$\frac{d}{dt}\underline{x}(t) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\underline{x}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u(t), \qquad (2.1)$$

which is completely controllable, sampling rates with periods that are multiples of π lead to uncontrollable discrete time systems, while all other sampling rates lead to controllable systems. Thus, plotting J* as a function of Δ would give something like Figure 2.

Even for such a system, decreasing the sampling period will sometimes give improved results. It is true that

$$J^*(\text{continuous time}) = \liminf_{\Delta} J^*(\Delta), \qquad (2.2)$$

that is, for any $\Delta > 0$, there exists a $\Delta_1 < \Delta$ such that $J^*(\Delta_1) \leq J^*(\Delta)$. In particular, if n is a positive integer, then $J^*(\frac{\Delta}{n}) \leq J^*(\Delta)$, since the control which optimizes the system for Δ is also available when $\Delta_1 = \frac{\Delta}{n}$.

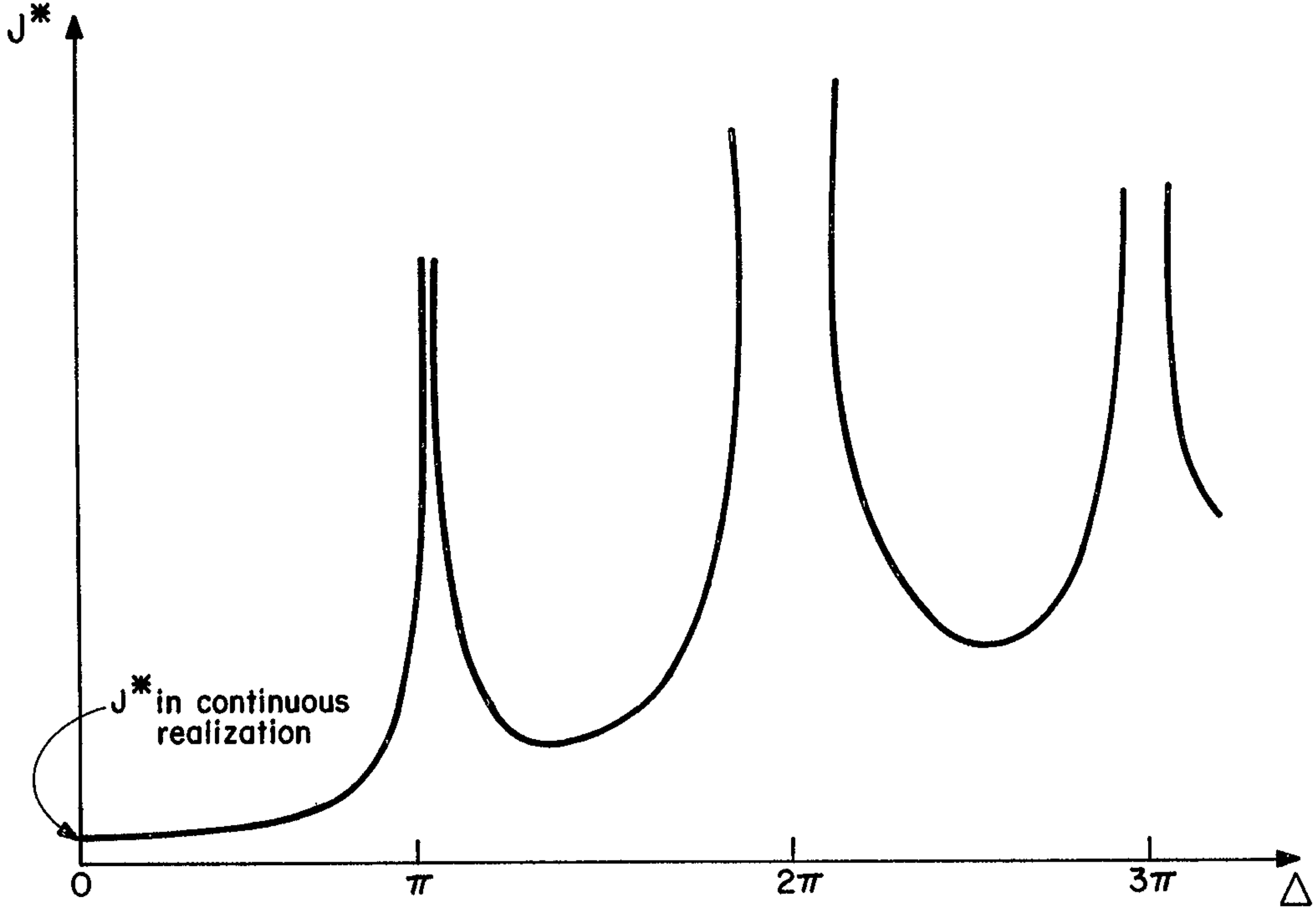

Fig.2 Possible Optimal Cost of the Harmonic Oscillator as a Function of the Sampling Period (see ref [10])

Not all systems are as ill-behaved as that of Figure 2. For the scalar case, that is, the system in which $u \in \mathbb{R}$ and $x \in \mathbb{R}$, we can actually determine the interval of discretization Δ which minimizes J. If $J^*(\Delta_1)$ is the minimal value of J for a given $\Delta = \Delta_1$, then we can show that·

$$\frac{\partial J^*(\Delta)}{\partial \Delta} > 0. \tag{2.3}$$

Thus, the optimal value of Δ is Δ_{min}. Since in this case, the feedback matrix is a scalar, our online computation involves only one multiplication. Thus, $\Delta_{min} = \tau_M$, where τ_M is the time needed for a single multiplication.

We discretize the system

$$\frac{d}{dt} x(t) = ax(t) + bu(t), \; x(0) = x_0, \tag{2.4}$$

with cost functional

$$J = \lim_{T \to \infty} \frac{1}{2} \int_0^T (qx^2 + ru^2)dt, \tag{2.5}$$

where $q, r > 0$, according to the solution of Levis [10]. For a given Δ, holding u constant over each interval, J in (2.5) becomes

$$J = \lim_{N \to \infty} \frac{1}{2} \sum_{k=0}^{N-1} [(x(k\Delta))^2 Q + 2x(k\Delta)Mu(k\Delta) + (u(k\Delta))^2 R], \tag{2.6}$$

where Q, M and R are scalars as follows:

$$Q = \int_0^\Delta e^{2at} qdt = \frac{q}{2a} (e^{a\Delta} + 1)(e^{a\Delta} - 1) \tag{2.7}$$

$$M = \int_0^\Delta bq\, e^{at} \left[\int_0^t e^{as} ds \right] dt = \frac{bq}{2a^2} (e^{a\Delta} - 1)^2 \tag{2.8}$$

$$R = \int_0^\Delta [r + q(b \int_0^t e^{as} ds)^2] dt$$

$$= a\Delta \left(\frac{ra^2 + qb^2}{a^3}\right) + \frac{qb^2}{2a^3} (e^{a\Delta} - 3)(e^{a\Delta} - 1) \tag{2.9}$$

The state equation becomes

$$x_{k+1} = \phi x_k + Du_k, \tag{2.10}$$

where

$$\phi = e^{a\Delta}, \tag{2.11}$$

and

$$D = \int_0^\Delta \phi(t,0) b dt = \frac{b}{a} (e^{a\Delta} - 1). \tag{2.12}$$

Now in steady state, the Riccati equation becomes

$$0 = -\hat{K}^2 [\phi - \frac{DM}{R}]^2 D^2 + (\phi - \frac{DM}{R})^2 (R + D^2\hat{K})\hat{K}$$

$$- \hat{K}(R + D^2\hat{K}) + (Q - \frac{M^2}{R})(R + D^2\hat{K}), \tag{2.13}$$

which simplifies to

$$0 = -D^2\hat{K}^2 + ((\phi^2 - 1)R - 2DM\phi + QD^2)\hat{K} + (QR - M^2). \tag{2.14}$$

If

$$\hat{B} = (\phi^2 - 1)R - 2DM\phi + QD^2, \tag{2.15}$$

then

$$\hat{K} = \frac{(\phi^2-1)R - 2DM\phi + QD^2}{2D^2} + \sqrt{\frac{\hat{B}^2 + 4D^2QR - 4M^2D^2}{-4D^2}} \tag{2.16}$$

Now

$$\frac{\hat{B}}{2D^2} = \frac{(\phi^2-1)R}{2\frac{b^2}{a^2}(\phi-1)^2} - \frac{M\phi}{D} + \frac{Q}{2}$$

$$= \frac{a^2(\phi+1)R}{2b^2(\phi-1)} - \frac{bq(\phi-1)^2\phi}{2a^2\frac{b}{a}(\phi-1)} + \frac{q(\phi^2-1)}{4a}$$

$$= \frac{(\phi+1)}{(\phi-1)}\left(\frac{ra^2+qb^2}{2ab^2}\right)(a\Delta) + \frac{q(\phi+1)(\phi-3)}{4a} - \frac{q(\phi^2-2\phi+1)}{4a}$$

$$= \frac{(\phi+1)}{(\phi-1)}\left(\frac{ra^2+qb^2}{2ab^2}\right)(a\Delta) - \frac{q}{a} \tag{2.17}$$

If we denote the second term in (2.16) by $\sqrt{T}$, then

$$T = \frac{(\phi+1)^2}{(\phi-1)^2}\left(\frac{ra^2+qb^2}{2ab^2}\right)^2(a\Delta)^2 + \frac{q^2}{a^2} - \frac{q(\phi+1)(a\Delta)(ra^2+qb^2)}{a^2b^2(\phi-1)}$$

$$+ \frac{\frac{q}{2a}(\phi^2-1)\,\frac{(a\Delta)(ra^2+qb^2)}{a^3} + \frac{q^2b^2}{4a^4}(\phi^2-1)(\phi-3)(\phi-1)}{\frac{b^2}{a^2}(\phi-1)^2}$$

$$- \frac{\frac{b^2q^2}{4a^4}(\phi-1)^4}{\frac{b^2}{a^2}(\phi-1)^2}$$

$$= \left(\frac{ra^2+qb^2}{2ab^2}\right)^2\frac{(\phi+1)^2}{(\phi-1)^2}(a\Delta)^2 + \frac{q^2}{a^2} - \frac{q(ra^2+qb^2)(\phi+1)(a\Delta)}{a^2b^2(\phi-1)}$$

$$+\frac{q(ra^2+qb^2\ (\phi+1)(a\Delta)}{2b^2a^2(\phi-1)}+\frac{q^2(\phi+1)(\phi-3)}{4a^2}+\frac{q^2(\phi-1)^2}{4a^2}$$

$$=\left(\frac{ra^2+qb^2}{2ab^2}\right)^2\frac{(\phi+1)^2}{(\phi-1)^2}(a\Delta)^2-\frac{q(ra^2+qb^2)(\phi+1)(a\Delta)}{2a^2b^2(\phi-1)}\quad . \tag{2.18}$$

Note that

$$\frac{\phi-1}{a\Delta}=\frac{e^{a\Delta}-1}{a\Delta}=\frac{(1+a\Delta+\frac{(a\Delta)^2}{2!}+\ldots)-1}{a\Delta}=1+\frac{a\Delta}{2!}+\frac{(a\Delta)^2}{3!}+\ldots \tag{2.19}$$

which is 1 when $a\Delta = 0$. Thus,

$$\lim_{\Delta\to 0}\hat{K}=\frac{ra^2+qb^2}{ab^2}-\frac{q}{a}+\sqrt{\left(\frac{ra^2+qb^2}{ab^2}\right)^2-\frac{q(ra^2+qb^2)}{a^2b^2}}$$

$$=\frac{ra}{b^2}+\frac{r}{b}\sqrt{\frac{a^2}{b^2}+\frac{q}{r}} \tag{2.20}$$

which is the continuous time solution derived in [1].

Now it is well known that

$$J^*(\Delta)=\frac{1}{2}x^2(0)\hat{K}. \tag{2.21}$$

It is straightforward to show (see Appendix A) that

$$\frac{\partial\hat{K}}{\partial\Delta}\geq 0, \tag{2.22}$$

and so, by (2.21),

$$\frac{\partial J^*}{\partial\Delta}(\Delta)\geq 0 \tag{2.23}$$

for all $\Delta \geq 0$.

We have shown, then, that for the scalar system described in equations (2.4) and (2.5), the optimizing Δ is Δ_{min}. Since the feedback control law is of the form

$$u = -kx, \tag{2.24}$$

Δ_{min} is just the time a controlling computer needs for a simple multiplication, τ_M. In other words, $\Delta = \tau_M$.

Chapter III

The Discrete Time Output Regulator

In order to optimize a cost functional over a class of control algorithms, as discussed in Chapter I, it is, of course, necessary to develop the appropriate set of algorithms. The most basic such control that comes to mind is the linear, full-state observation, sampled-data control developed by Levis [10]. His solution is to control the system

$$\dot{x}(t) = Ax(t) + Bu(t);\; x(t_0) = x_0 \tag{3.1}$$

in order to minimize the functional

$$J = \int_0^\infty (\underline{x}'\hat{Q}\underline{x} + \underline{u}'\hat{R}\underline{u})\,dt \tag{3.2}$$

by approximating (3.1) by a discrete-time system of the form

$$x_{k+1} = \Phi x_k + Du_k \tag{3.3}$$

$$u_k = -Gx_k \tag{3.4}$$

with cost functional

$$J = \sum_{i=0}^{\infty} (x_i'Qx_i + u_i'Ru_i + 2x_i'Mu_i) \tag{3.5}$$

where $u_i = u(i\Delta)$ and $x_i = x(i\Delta)$ for a given Δ, and where the parameters Φ, D, Q, R and M are given as follows

$$\Phi = e^{A\Delta} \tag{3.6}$$

$$D = \int_0^\Delta e^{At}B\,dt \tag{3.7}$$

$$Q = \int_0^{\Delta} e^{A't} Q e^{At} dt \tag{3.8}$$

$$R = \hat{R}\Delta + B' \int_0^{\Delta} [\int_0^t e^{A's} ds] A [\int_0^t e^{As} ds] dt \cdot B \tag{3.9}$$

$$M = \int_0^{\Delta} e^{A't} \hat{Q} [\int_0^t e^{As} ds] B dt. \tag{3.10}$$

Levis found that the optimal value for G is given by

$$G = R^{-1}M' + (R + D'\hat{K}D)^{-1} D'\hat{K}(\Phi - DR^{-1}M') \tag{3.11}$$

where

$$\hat{K} = \lim_{k \to \infty} K_k \tag{3.12}$$

is the steady-state solution of

$$K_k = [\phi - DR^{-1}M']' [K_{k+1} - K_{k+1} D(R + D'K_{k+1}D)^{-1} D'K_{k+1}] \\ [\Phi - DR^{-1}M'] + [Q - MR^{-1}M'] \tag{3.13}$$

with boundary condition

$$K_N = 0. \tag{3.14}$$

The cost in this case is $\frac{1}{2} x_0'Kx_0$.

Now if $u \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$, and τ_M and τ_A are the times needed for a computer to perform an elementary multiplication and addition, respectively, then the minimum online time needed to perform the vector multiplication $G \cdot x_i$ is $mn\tau_M + m(n-1)\tau_A$, which is approximately $mn\tau_M$ if $\tau_M >> \tau_A$ (we are assuming, of course, that all of the elements of G are non-zero - as they would usually be). Thus the minimum interval of discretization is $\Delta_{min} = mn\tau_M$.

Let us consider the controllable 2-dimensional system

$$\dot{x}(t) = \begin{bmatrix} -1 & \varepsilon \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t). \tag{3.15}$$

We might suspect that if ε is sufficiently small, the behavior of this system would resemble that of

$$\dot{x}(t) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \quad , \tag{3.16}$$

which is uncontrollable, but which can be stabilized. If we define

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad , \tag{3.17}$$

then in system (3.16), x_1 will decay to zero regardless of input, while x_2 will behave like the scalar system

$$\dot{x}_2(t) = x_2(t) + u(t), \tag{3.18}$$

and, as was shown in Chapter II, a decreasing Δ will result in a monotonically decreasing cost associated with $x_2(t)$. Thus, the optimal linear control for (3.16) will be of the form

$$u = k\, x_2(t), \tag{3.19}$$

and the optimal Δ is $\Delta = \Delta_{min}$. In the case of system (3.15), for a given Δ, the optimal control will involve two multiplications; as we shall prove in Chapter V, however, a better result can be obtained by implementing the one-multiplication law of (3.19), if ε is sufficiently small.

Clearly, then, in some systems many of the optimal gains are small, and in higher-dimensional systems a great deal of computational simplicity may be obtained by setting these gains to zero. Thus, excellent performance can sometimes be obtained by feeding back only a <u>subset</u> of the states, and, given the simpler form of the controller, using a much smaller Δ in order to achieve performance superior to the "optimal". With this motivation, we will develop an optimal constant discrete-time partial-state feedback control.

The following results are parallel to the continuous-time work of Levine [9]. We include a complete derivation. The equations defining the solution are (3.42)-(3.44).

Given the system

$$x_{k+1} = \phi x_k + Du_k \tag{3.20}$$

$$y_k = Cx_k, \tag{3.21}$$

we wish to select F in the feedback law

$$u_k = -Fy_k \tag{3.22}$$

in order to minimize (3.5), which leads to a cost

$$J(F) = \frac{1}{2}\,\mathrm{tr}\left\{\sum_{i=0}^{\infty} (\phi_0^{\,i})'[Q-2MFC + C'F'RFC]\phi_0^{\,i}\right\}, \tag{3.23}$$

where

$$\phi_0 = \phi - DFC. \tag{3.24}$$

We use the "trace" form in (3.23) to avoid developing a feedback law that depends on the initial state of the system; our solution will be optimal in an "average" sense ([6], [7], [9]).

We will find the following identity, due to Kleinman [5] useful. If we have a function $J(F)$ which satisfies the condition

$$J(F + \varepsilon\Delta F) - J(F) = \varepsilon \mathrm{tr}[M(F)\Delta F] + o(\varepsilon), \tag{3.25}$$

then,

$$\frac{\partial J(F)}{\partial F} = M'(F). \tag{3.26}$$

We proceed as follows. for small ε we have $\varepsilon^n \approx 0$ for $n \geq 2$, so we may approximate

$$J(F + \varepsilon\Delta F) \approx \frac{1}{2} \sum_{i=0}^{\infty} \mathrm{tr}\Big\{[\phi_0 - \varepsilon D\Delta FC]'^{i} \cdot$$

$$[Q-2M(F+\varepsilon\Delta F)C+C'F'RFC + \varepsilon C'F'R\Delta FC + \varepsilon C'\Delta F'RFC]\,[\phi_0-\varepsilon D\Delta FC]^{i}\Big\}. \tag{3.27}$$

To first order, for $i \geq 1$,

$$[\phi_0 - \varepsilon D\Delta FC]^{i} \approx \phi_0^{\,i} - \varepsilon\Big\{\phi_0^{\,i-1}D\Delta FC + \phi_0^{\,i-2}D\Delta FC\phi_0 + \ldots + D\Delta FC\phi_0^{\,i-1}\Big\} \tag{3.28}$$

(The expression in (3.28) is the identity when $i = 0$.)

We will call the expression on the right hand side $\phi_0^{\,i} - \varepsilon A_i$.

Now we rewrite (3.27) as

$$J(F+\varepsilon\Delta F) \approx \frac{1}{2} \Big[\sum_{i=1}^{\infty} \mathrm{tr}\Big\{[\phi_0^{\,i} - \varepsilon A_i]'[Q-2MFC+C'FRFC]\phi_0^{\,i}$$

$$+ \varepsilon\phi_0^{\,i}[-2M\Delta FC + C'F'R\Delta FC]A_i\Big\} + \mathrm{tr}\Big\{Q-2M(F+\Delta F\varepsilon)C$$

$$+ C'FRFC + \varepsilon C'F'R\Delta FC + \varepsilon C'\Delta F'RFC\Big\}, \tag{3.29}$$

so,

$$J(F + \varepsilon\Delta F) - J(F) = \frac{1}{2} \sum_{i=1}^{\infty} \mathrm{tr}\{-\varepsilon A_i [Q - 2MFC + C'F'RFC]\phi_0^{\,i}$$

$$+ \varepsilon\phi_0^{\,i'} [-2M\Delta FC + C'F'R\Delta FC + C'\Delta F'RFC]\phi_0^{\,i}$$

$$- \varepsilon\phi_0^{\,i'} [Q - 2MFC + C'F'RFC]A_i\}$$

$$+ \mathrm{tr}\{\varepsilon(-2M\Delta FC + C'F'R\Delta FC + C'\Delta F'RFC)\}]. \tag{3.30}$$

Now, if we let

$$Q_0 = Q - MFC - C'F'M' + C'F'RFC, \tag{3.31}$$

and use the trace property $\mathrm{tr}\{X\Delta FY\} = \mathrm{tr}\{YX\Delta F\}$ to get each term to end in ΔF, we get

$$J(F + \varepsilon\Delta F) - J(F) = \frac{\varepsilon}{2}\,\mathrm{tr}\{-2CM\Delta F + 2CC'F'R\Delta F\}$$

$$+ \frac{\varepsilon}{2}\,\mathrm{tr}\,\{\sum_{i=1}^{\infty} -2C\phi_0^{\,i}\phi_0^{\,i'}M\Delta F + 2C\phi_0^{\,i}\phi_0^{\,i'}C'F'R\Delta F]$$

$$- 2\varepsilon\sum_{i=1}^{\infty} C\phi_0^{\,i'}Q_0\phi_0^{\,i-1}D\Delta F + C\phi_0\phi_0^{\,i'}Q_0\phi_i^{\,i-2}D\Delta F$$

$$+ C\phi_0^{\,2}\phi_0^{\,i'}Q_0\phi_i^{\,i-3}D\Delta F + \ldots + C\phi_0^{\,i-1}\phi_0^{\,i'}Q_0 D\Delta F)$$

$$= \varepsilon\ \mathrm{tr}\,\{\sum_{i=0}^{\infty} [C\phi_0^{\,i}\phi_0^{\,i'}C'F'R\Delta F - C\phi_0^{\,i}\phi_0^{\,i'}M\Delta F]$$

$$- \sum_{i=1}^{\infty}\ \sum_{j=0}^{i-1}\ C\phi_0^{\,j}\phi_0^{\,i'}Q_0\phi_0^{\,i-1-j}D\Delta F\}, \tag{3.32}$$

so

$$\frac{\partial J}{\partial F} = \sum_{i=0}^{\infty} (RFC\phi_0^{\,i}\phi_0^{\,i'}C' - M'\phi_0^{\,i}\phi_0^{\,i'}C'$$

$$- \sum_{i=1}^{\infty}\ \sum_{j=0}^{i-1}\ D'\phi_0^{\,i-1-j'}Q_0\phi_0^{\,i}\phi_0^{\,j'}C'. \tag{3.33}$$

Now, let

$$L = \sum_{i=0}^{\infty} \phi_0^{\,i}\phi_0^{\,i'}, \text{ and} \tag{3.34}$$

$$K = \sum_{i=0}^{\infty} \phi_0^{\,i'} Q_0 \phi_0^{\,i}, \text{ so that} \tag{3.35}$$

$$\phi_0 L \phi_0{}' = L - I, \text{ and} \tag{3.36}$$

$$\phi_0{}' K \phi_0 = K - Q_0. \tag{3.37}$$

Then if we set $\frac{\partial J}{\partial F} = 0$ in order to find the minimum for J, equation (3.33) becomes

$$0 = RFCLC' - M'LC' - \sum_{i=1}^{\infty} \sum_{j=0}^{i-1} D'(\phi_0^{\,i-1-j})'Q_0\phi_0^{\,i}(\phi_0^{\,j})'C' \tag{3.38}$$

But

$$\sum_{i=1}^{\infty} \sum_{j=0}^{i-1} (\phi_0^{\,i-1-j})'Q_0\phi_0^{\,i}(\phi_0^{\,j})' = \sum_{j=0}^{\infty} \sum_{i=j+1}^{\infty} \phi_0^{\,(i-1-j)'}Q_0\phi_0^{\,i}(\phi_0^{\,j})'$$

$$= \sum_{j=0}^{\infty} \sum_{i=j+1}^{\infty} \phi_0^{\,(i-1-j)'}Q_0\phi_0^{\,(i-j)}\phi_0^{\,j}(\phi_0^{\,j})' \tag{3.39}$$

If we let $k = i-j-1$, then (3.39) becomes

$$\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} (\phi_0^{\,k})'Q_0\phi_0^{\,k+1}\phi_0^{\,j}(\phi_0^{\,j'}) = \sum_{k=0}^{\infty} (\phi_0^{\,k})'Q_0\phi_0^{\,k+1} \sum_{j=0}^{\infty} \phi_0^{\,j}(\phi_0^{\,j})'$$

$$= K\phi_0 L. \tag{3.40}$$

Then substituting in equation (3.38) gives

$$0 = RFCLC' - M'LC' - D'K\phi_0 LC'$$

$$= RFCLC' - M'LC' - D'K\phi LC' + D'KDFCLC', \tag{3.41}$$

so the solution is

$$F = (R + D'KD)^{-1}(M' + D'K\phi)LC'(CLC')^{-1}, \tag{3.42}$$

where

$$K = \sum_{i=0}^{\infty} \phi_0^{i'} Q_0 \phi_0^{i}, \tag{3.43}$$

and

$$L = \sum_{i=0}^{\infty} \phi_0^{i} \phi_0^{i'}, \tag{3.44}$$

where

$$\phi_0 = \phi - DFC \tag{3.45}$$

and

$$Q_0 = Q - MFC - C'F'M' + C'F'RFC. \tag{3.46}$$

As we note in Appendix C, equations (3.42)-(3.46) reduce to the usual discrete-time solution [10] when $C = I$.

Now, the optimal cost for the partial-state feedback controller is

$$J^* = \frac{1}{2}\,\mathrm{tr}\,\{x_0'Kx_0\}, \tag{3.47}$$

where K is as given in (3.37). We might choose to compare the full- and partial-state feedback laws by calculating Φ, D, Q, R, and M for a given Δ, solve for the optimal full-state feedback solution, re-calculate the parameter set for appropriately reduced values of Δ (reflecting the reduced computation time), and determine the optimal part-state solution for these values of Δ.

Chapter IV

The Discrete Optimal Minimal-Order Compensator

One technique used to stabilize a plant system for which only part of the state can be observed is the construction of a dynamic compensator ([2], [12], [13]), i.e., a system that estimates the unknown part of the plant state based on the entire past history of plant observations. In some cases [2], the plant is not output stabilizable, and use of a compensator is a necessity. If the system is output stabilizable, though, we may choose either a simple output feedback controller or a more computationally complex compensator. We would like to know, then, when to use the "faster" output feedback or the "more accurate" dynamic compensation.

First, however, we must develop a discrete-time version of the optimal compensator. Our work here parallels the continuous-time derivation of Blanvillain [2]. The results are summarized in equations (4.51)-(4.55).

Let us consider the system

$$x(k+1) = Ax(k) + Bu(k) \tag{4.1}$$

with observation

$$y(k) = Cx(k) \tag{4.2}$$

where the vectors and matrices are defined as

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{4.3}$$

with $x_1 \in \mathbb{R}^m$, $x_2 \in \mathbb{R}^{n-m}$,

$$C = \begin{bmatrix} I_{m\times m} & 0_{m\times(n-m)} \end{bmatrix}, \tag{4.4}$$

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{4.5}$$

with $A_{11} \in \mathbb{R}^m \times \mathbb{R}^m$, $A_{22} \in \mathbb{R}^{n-m} \times \mathbb{R}^{n-m}$, $A_{12} \in \mathbb{R}^{n-m} \times \mathbb{R}^m$, $A_{21} \in \mathbb{R}^m \times \mathbb{R}^{n-m}$,

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \tag{4.6}$$

with $B_1 \in \mathbb{R}^m$, $B_2 \in \mathbb{R}^{n-m}$.

In the above, we have assumed a canonical form for the system, insisting that the observation matrix is of a particular form, as in (4.4). In fact, Blanvillain has shown, for continuous-time, that every linear system is indeed equivalent to a system of this form, if one eliminates redundant observations. The argument is easily extended to the discrete-time case.

We would like to choose a cost functional to minimize, using a linear feedback control. Unfortunately, as in the continuous-time case, using a cost of the form

$$\sum_{i=0}^{\infty} (x(i)'Qx(i) + u(i)'Ru(i) - 2x(i)'Mu(i)) \tag{4.7}$$

would lead to a control law that depends on the <u>initial</u> state $x(0) = x_0$, that is, the optimal linear control would not be a strict feedback law. We use instead a cost that leads to a feedback law that is optimal over the average of a set of possible initial states; it is of the form

$$\hat{J} = E\{\sum_{i=0}^{\infty} (x(i)'Qx(i) + u(i)'Ru(i) - 2x(i)'Mu(i))\}. \tag{4.8}$$

We will additionally treat x_0 as a random variable with statistics

$$E\{x_0\} = 0 \ ; \ E\{x_0 x_0'\} = \Sigma_0. \tag{4.9}$$

Now, it should be noted that the system (4.1)-(4.6) can be written as follows·

$$x_1(k+1) = A_{11}x_1(k) + A_{12}x_2(k) + B_1u(k) \tag{4.10}$$

$$x_2(k+1) = A_{21}x_1(k) + A_{22}x_2(k) + B_2u(k) \tag{4.11}$$

$$y(k) = x_1(k). \tag{4.12}$$

Since $x_1(k)$ can be observed exactly, and $u(k)$ is known, we may consider the input to be

$$v(k) = A_{21}x_1(k) + B_2u(k) \tag{4.13}$$

with observation

$$y_2(k) = x_1(k+1) - A_{11}x_1(k) - B_1u(k) \tag{4.14}$$

so the system (4.1)-(4.2) may be written as

$$x_2(k+1) = A_{22}x_2(k) + v(k) \tag{4.15}$$

$$y_2(k) = A_{12}x_2(k). \tag{4.16}$$

If we assume the compensator is of the form

$$\hat{x}_2(k+1) = F\hat{x}_2(k) + Lv(k) + Hy_2(k) \tag{4.17}$$

then we choose [12] $L = I$, and

$$F = A_{22} - HA_{12}, \tag{4.18}$$

where H is a design matrix; it can be shown ([2],[12]) that the error $e = x_2 - \hat{x}_2$ is described by

$$e(k+1) = (A_{22} - HA_{12})e(k) = Fe(k), \tag{4.19}$$

and thus H may be adjusted [15] for arbitrary error dynamics (assuming observability of the original system).

Now one problem with the formulation (4.17) is that $y_2(k)$ is a function of the future state $x_1(k+1)$ (see 4.14). To eliminate this, we instead use for our compensator system the variable

$$z(k) = \hat{x}_2(k) - Hx_1(k), \text{ so} \tag{4.20}$$

$$z(k+1) = Fz(k) + Px_1(k) + Du(k), \tag{4.21}$$

where F is as in (4.15), and

$$P = FH - HA_{11} + A_{21}, \tag{4.22}$$

$$D = B_2 - HB_1. \tag{4.23}$$

This is in fact the minimal Luenberger observer. To optimize the cost (4.8) over H, we will for the sake of computational simplicity reconsider this in terms of the augmented state

$$s(k) = \begin{bmatrix} x(k) \\ e(k) \end{bmatrix} , \quad s_0 = \begin{bmatrix} x_0 \\ e_0 \end{bmatrix} \tag{4.24}$$

where

$$e(k) = x_2(k) - \hat{x}_2(k) = x_2(k) - Hx_1(k) - z(k). \tag{4.25}$$

We will assume at this point that a "separability" property holds, i.e., that if $u(k)$ is to depend on $x_1(k)$ and $\hat{x}_2(k)$, as in

$$u(k) = G\hat{x}(k), \tag{4.26}$$

then the optimal values for G and H may be designed independently of each other. This means, then, that the optimal feedback G* is exactly the optimal feedback matrix for the system (4.1) when the full-state is observed. Blanvillain [2] justifies this assumption by showing that it leads to an optimal compensator design matrix H* that is independent of Q and R, two matrices on which G* explicitly depends. We feel, however, that this argument is circular, and much prefer the reasoning of Miller [13], who shows that for any given H*, use of any other feedback matrix will lead to an increase in cost. The proof is mechanical, and easily extends to the discrete-time case; we will dispense with it here.

We may now express the input as

$$u(k) = G^*x(k) - G^*Ne(k), \tag{4.27}$$

where

$$N = \begin{bmatrix} 0_{m\times(n-m)} \\ I_{(n-m)\times(n-m)} \end{bmatrix} \tag{4.28}$$

The closed loop system may be then described

$$s(k+1) = \Gamma x(k) = \Gamma^{k+1} s_0, \tag{4.29}$$

where

$$\Gamma = \begin{bmatrix} A + BG^* & -BG^*N \\ 0 & F \end{bmatrix} \tag{4.30}$$

The cost (4.8) is rewritten as

$$\hat{J}(H) = E\{\sum_0^\infty s'(k)\Omega s(k)\} \tag{4.31}$$

in which

$$\Omega = \begin{bmatrix} Q + G^{*\prime}RG^* + MG^* + G^{*\prime}M' & -G^{*\prime}RG^*N - MG^*N \\ -N'G^{*\prime}RG^* - N'G^{*\prime}M' & N'G^{*\prime}RG^*N \end{bmatrix} \tag{4.32}$$

If we let $\Xi_0 = E\{s_0 s_0'\}$, then the cost becomes

$$\hat{J}(H) = tr\{\sum_0^\infty \Gamma'^k \Omega \Gamma^k \Xi_0\}, \tag{4.33}$$

which suggests a Lyapunov equation; so if Λ is the solution of

$$\Gamma'\Lambda\Gamma + \Omega = \Lambda, \tag{4.34}$$

then the cost becomes

$$\hat{J}(H) = tr(\Lambda\Xi_0). \tag{4.35}$$

From this point, the derivation of conditions for an optimal H* is fairly mechanical, and we will present only the skeleton of the argument. For a more detailed exposition, see Appendix B.

Our strategy will be to isolate the portions of (4.35) that depend explicitly on H. From equations (4.9) and (4.25) we can show that

$$\Xi_0 = \begin{bmatrix} \Sigma_0 & \Xi_{12} \\ \Xi_{12}' & E_0 \end{bmatrix} \tag{4.36}$$

where

$$E_0 = E\{e_0 e_0'\} = H\Sigma_{11}H' + \Sigma_{22} - H\Sigma_{12} - \Sigma_{12}'H' \tag{4.37}$$

when Σ_0 is written

$$\Sigma_0 = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{bmatrix}. \tag{4.38}$$

Although both Ξ_{12} and E_0 are functions of H, it will be seen that Ξ_{12} does not contribute to the cost (4.35).

As to the solution of Lyapunov equation (4.34), that is,

$$\Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}' & \Lambda_{22} \end{bmatrix}, \tag{4.39}$$

it is trivial to show that Λ_{11} is the usual Riccati solution K of (3.14), given our earlier assumption of separability. This is rather unsurprising in light of the fact that our assumed feedback G* is dependent only on K and the system parameters (see (3.12)). This result leads to the equation for Λ_{12} reducing to

$$\Lambda_{12} = (A' + G^{*\prime}B')\Lambda_{12}F, \tag{4.40}$$

and if we assume that G^* is chosen for a stable plant, and H^*, upon which F depends ((4.18)), is chosen to stabilize the error ((4.19)), then

$$||\Lambda_{12}|| = 0. \tag{4.41}$$

This enables us to rewrite the cost (4.35) as

$$\hat{J}(H) = tr(K\Sigma_0) + tr(\Lambda_{22}E_0), \tag{4.42}$$

where Λ_{22} turns out to be

$$N'G^{*\prime}(B'KB + R)G^*N + F'\Lambda_{22}F = \Lambda_{22}. \tag{4.43}$$

Since only the second term of (4.42) is dependent on H, we will minimize the <u>added cost</u>

$$J(H) = tr(\Lambda_{22}E_0). \tag{4.44}$$

If we write (4.43) as

$$M_0 + F'\Lambda_{22}F = \Lambda_{22}, \tag{4.45}$$

noting that M_0 is independent of H, then since Λ_{22} is the solution of a Lyapunov equation, (4.44) becomes

$$\begin{aligned} J(H) &= tr\,\{(\sum_{i=0}^{\infty} (F^i)'M_0F^i)E_0\} \\ &= tr\{M_0 \sum_{i=0}^{\infty} F^iE_0(F^i)'\} \end{aligned} \tag{4.46}$$

As we intend to minimize J according to

$$\left.\frac{\partial J}{\partial H}\right|_{H^*} = 0, \tag{4.47}$$

we may again, as in Chapter III, make use of Kleinman's lemma ((3.6)-(3.7)) by getting an expression for $J(H + \varepsilon\Delta H) - J(H)$ in the appropriate form. It then turns out that

$$H^* = (\Sigma_{12}' + A_{22}WA_{12}')(\Sigma_{11} + A_{12}WA_{12}')^{-1}, \tag{4.48}$$

where W is the expression

$$W = \sum_{i=0}^{\infty} F^i E_0 (F^i)', \tag{4.49}$$

which naturally appears in (4.46); by substituting H* into the Lyapunov equation associated with (4.49), we get

$$W = A_{22}WA_{22}' - (\Sigma_{12}' + A_{22}WA_{12}')(\Sigma_{11} + A_{12}WA_{12})^{-1}(A_{12}WA_{22}' + \Sigma_{12}) + \Sigma_{22} \tag{4.50}$$

Thus, the minimal increase in cost due to the compensator is

$$J^* = \mathrm{tr}(\Lambda_{22}^* E_0^*), \tag{4.51}$$

where

$$E_0 = H^*\Sigma_{11}H^{*\prime} + \Sigma_{22} - H^*\Sigma_{12} - \Sigma_{12}'H^{*\prime}, \tag{4.52}$$

$$(A_{22} - H^*A_{12})\Lambda_{22}^*(A_{22}' - A_{12}'H^{*\prime}) + \Lambda_{22}^* = N'K(B'KB + R)KN \tag{4.53}$$

$$H^* = (\Sigma_{12}' + A_{22}WA_{12}')(\Sigma_{11} + A_{12}WA_{12}')^{-1}, \tag{4.54}$$

and

$$W = A_{22}WA_{22}' + \Sigma_{22} - (\Sigma_{12}' + A_{22}WA_{12}')(\Sigma_{11} + A_{12}WA_{12}')^{-1} \cdot (A_{12}WA_{22}' + \Sigma_{12}) \qquad (4.55)$$

These equations, of course, may all be solved offline. However, the online construction of the state estimate $\hat{x}$, needed to implement a feedback control, may be very time consuming. An examination of equations (4.20)-(4.23) show that $(n-m)(n+m+1)$ multiplications will be needed for state reconstruction, in addition to n for feedback; only m will be needed for partial-state feedback.

Chapter V

An Example: Full- vs. Partial-State Feedback

As we noted in Chapter III, we may sample outputs, and calculate optimal controls, at a greater rate when we use a partial-state feedback algorithm than when we use full-state feedback. In other words, even if we are physically able to observe the full state of a system, we might feed back a function of only part of the state, hence simplifying computation. In such a case, we will lose knowledge of part of the state that would have been available at a slower rate; in return, however, we gain more frequent knowledge of the part of the state that we do access.

Kokotovic [8] describes a class of systems, called "ε-coupled", which provide a number of illustrative examples. A large system is said to be ε-coupled if it splits into several independent subsystems when a scalar parameter ε is zero. As Kokotovic maintains, the amount of computation for continuous-time control is greatly reduced if we calculate the optimal regulators for the decoupled subsystems rather than for the full system; moreover, the performance of the suboptimal control is fairly near that of the optimal. Since we are in fact concerned with a sampled-control, however, the suboptimal control will lead to a shorter possible sampling period and perhaps a better result than in the full system control.

In particular, if decoupling of a completely controllable system results in one subsystem that is uncontrollable but stable, and one that is controllable, then it might be advantageous to feed back only the states that belong to the latter subsystem. We will consider a simple 2-dimensional example of this. We look at the system

$$\dot{x}(t) = \begin{bmatrix} -1 & \varepsilon \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t), \tag{5.1}$$

which is controllable for all $\varepsilon \neq 0$. If ε is very small, we might expect that (5.1) will act very much like

$$\dot{x}(t) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t). \tag{5.2}$$

In fact, this new system (5.2) is output stabilizable; if we let

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \tag{5.3}$$

then $x_1(t)$ tends to zero regardless of the input, while the behavior of $x_2(t)$ is described by the scalar equation

$$\dot{x}_2(t) = x(t) + u(t), \tag{5.4}$$

which is obviously controllable.

Since only the cost associated with $x_2(t)$ is determined by the feedback law, the cost criterion, of the form

$$J = \int_0^\infty [x'(t) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x(t) + ru(t)^2]dt, \tag{5.5}$$

will be minimized by choosing as fast a sampling rate as possible; it follows from Chapter II that J will be a monotonic function of the sampling interval Δ. What we will show is that if ε in system (5.1) is sufficiently small, the sampling interval chosen in order to minimize a cost of the form (5.5) should also be as small as possible.

We will assume for the sake of computational simplicity that $\varepsilon^2 \approx 0$. We define the matrices

$$A = \begin{bmatrix} -1 & \varepsilon \\ 0 & 1 \end{bmatrix} \tag{5.6}$$

and

$$B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{5.7}$$

and using the usual series definition

$$e^{AT} = I + TA + \frac{T^2}{2!} A^2 + \frac{T^3}{3!} A^3 + \frac{T^4}{4!} A^4 + \ldots \tag{5.8}$$

we obtain

$$\begin{aligned} e^{AT} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + T \begin{bmatrix} -1 & \varepsilon \\ 0 & 1 \end{bmatrix} + \frac{T^2}{2!} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{T^3}{3!} \begin{bmatrix} -1 & \varepsilon \\ 0 & 1 \end{bmatrix} + \ldots \\ &= \begin{bmatrix} e^{-T} & \varepsilon \sinh T \\ 0 & e^{T} \end{bmatrix}. \end{aligned} \tag{5.9}$$

Now using the definitions of (3.7)-(3.11), we have

$$\Phi = e^{A\Delta} = \begin{bmatrix} e^{-\Delta} & \varepsilon \sinh\Delta \\ 0 & e^{\Delta} \end{bmatrix}, \tag{5.10}$$

$$D = \begin{bmatrix} \varepsilon\cosh\Delta \\ e^{\Delta} - 1 \end{bmatrix}, \tag{5.11}$$

$$\hat{Q} = \begin{bmatrix} \frac{1-e^{-2\Delta}}{2} & \frac{\varepsilon}{2}\left(\frac{2\Delta+e^{-2\Delta}-1}{2}\right) \\ \frac{\varepsilon}{2}\left(\frac{2\Delta+e^{-2\Delta}-1}{2}\right) & \frac{e^{2\Delta}-1}{2} \end{bmatrix} \tag{5.12}$$

$$M = \begin{bmatrix} \frac{\varepsilon}{2}\left(\frac{2\Delta+1-e^{-2\Delta}}{2}\right) \\ \frac{1}{2}\,(e^{\Delta}-1)^2 \end{bmatrix}, \tag{5.13}$$

and

$$\hat{R} = (r+1)\Delta + \frac{(e^{\Delta}-1)(e^{\Delta}-3)}{2}\,. \tag{5.14}$$

We should note that the discrete time system

$$x(k+1) = \Phi x(k) + Du(k) \tag{5.15}$$

is completely controllable, since

$$\Phi D = \begin{bmatrix} \varepsilon(e^{-\Delta}\cosh\Delta + e^{\Delta}\sinh\Delta - \sinh\Delta) \\ e^{\Delta}(e^{\Delta}-1) \end{bmatrix}, \tag{5.16}$$

so

$$\det\begin{bmatrix} \Phi D & \Phi \end{bmatrix} = \varepsilon(1-\cosh(2\Delta)), \tag{5.17}$$

which is nonzero when $\varepsilon \neq 0$ and $\Delta \neq 0$.

Now suppose we would like to determine the form of the minimal cost if we feed back only part of the state, i.e., x_2. In this derivation and the one that follows we will make liberal use of the approximation

$$\frac{z}{a+\varepsilon b} = \frac{z(a-\varepsilon b)}{a^2-\varepsilon^2 b^2} \approx \frac{za-\varepsilon zb}{a^2} . \tag{5.18}$$

Looking at (3.23), if we let

$$\hat{K} = \begin{bmatrix} k_1 & k_2 \\ k_2 & k_3 \end{bmatrix} , \tag{5.19}$$

then the expression D'KD becomes

$$D'KD = 2\varepsilon k_2 \cosh\Delta(e^{\Delta}-1)+k_3(e^{\Delta}-1)^2, \tag{5.20}$$

and $D'K\Phi$ is the matrix

$$D'K\Phi = \begin{bmatrix} \varepsilon k_1 e^{-\Delta}\cosh\Delta + k_2(1-e^{-\Delta}) & \varepsilon k_2(e^{2\Delta}\sinh\Delta) + k_3(e^{2\Delta}- e^{\Delta}) \end{bmatrix}. \tag{5.21}$$

Then,

$$M' + D'K\Phi = \left[\begin{array}{c|c} \varepsilon \frac{k_1}{2} (1+e^{-2\Delta}) + \frac{2\Delta-e^{-2\Delta}+1}{4} & \varepsilon(k_2(e^{2\Delta}- \sinh\Delta)) \\ + k_2(1-e^{-\Delta}) & +k_3(e^{2\Delta}-e^{\Delta})+\frac{(e^{\Delta}-1)^2}{2} \end{array}\right]. \tag{5.22}$$

Now of course we can define

$$C = [0 \quad 1] \tag{5.23}$$

for use in equation (3.23). If we define

$$L = \begin{bmatrix} \ell_1 & \ell_2 \\ \ell_2 & \ell_3 \end{bmatrix} , \tag{5.24}$$

then

$$LC'(CLC')^{-1} = \begin{bmatrix} \frac{\ell_2}{\ell_3} \\ \cdots \\ 1 \end{bmatrix}. \tag{5.25}$$

We can develop a property of L which will simplify things considerably. We know from equation (3.23) (repeated here) that

$$F = (R + D'KD)^{-1}(M' + D'R\Phi)LC'(CLC')^{-1}, \tag{5.26}$$

and in this case, F will be a scalar. Then

$$\Phi_0 \triangleq \Phi - DFC \tag{5.27}$$

becomes

$$\Phi_0 = \begin{bmatrix} e^{-\Delta} & \varepsilon(\sinh\Delta - F\cosh\Delta) \\ 0 & e^{\Delta} - F(e^{\Delta} - 1) \end{bmatrix}, \tag{5.28}$$

and using the definition of L in equation (3.15),

$$L = \sum_{i=0}^{\infty} (\Phi_0^{\,i})(\Phi_0^{\,i})', \tag{5.29}$$

it is easily determined that

$$\ell_1 = 1 + e^{-2\Delta} + e^{-4\Delta} + \ldots = \frac{1}{1-e^{-2\Delta}}, \tag{5.30}$$

and that ℓ_2 may be written as

$$\ell_2 = \varepsilon\hat{\ell}_2, \tag{5.31}$$

where $\hat{\ell}_2$ is a function of Δ; that is, we will be able to make the approximation

$$\varepsilon \ell_2 \approx 0. \tag{5.32}$$

Moreover, we can derive a similar property for K. Recalling that the definition (3.16) is

$$K = \sum_{i=0}^{\infty} (\Phi_0{}^i)' Q_0 \Phi_0{}^i , \tag{5.33}$$

where

$$Q_0 = \hat{Q} - MFC - C'F'M' + C'F'\hat{R}FC$$

$$= \begin{bmatrix} \frac{1-e^{-2\Delta}}{2} & \frac{\varepsilon}{4}[2\Delta+e^{-2\Delta}-1-F(2\Delta+1-e^{-2\Delta})] \\ \frac{\varepsilon}{4}[2\Delta+e^{-2\Delta}-1-F(2\Delta+1-e^{-2\Delta})] & \hat{R}F^2+\frac{(e^{2\Delta}-1)}{2}-F(e^{\Delta}-1)^2 \end{bmatrix} \tag{5.34}$$

we can readily see that

$$k_1 = \left(\frac{1-e^{-2\Delta}}{2}\right)(1 + e^{-2\Delta} + e^{-4\Delta} + \dots) = \frac{1-e^{-2\Delta}}{2(1-e^{-2\Delta})} = \frac{1}{2} \tag{5:35}$$

and

$$k_2 = \varepsilon \tilde{k}_2 \tag{5.36}$$

where $\tilde{k}_2$ is <u>not</u> a function of ε.

Now we maintain that k_3 is the solution of the Riccati-like equation associated with discrete-time linear control of the scalar system

$$\dot{x}(t) = x(t) + u(t) \ , \ y(t) = x(t), \tag{5.37}$$

with cost functional

$$J_2 = \int_0^\infty [x(t)^2 + ru(t)^2]dt. \tag{5.38}$$

By setting $b = q = a = 1$ in equations (2.5)-(2.10), the form of the Riccati-like equation (2.12) becomes

$$\begin{aligned} 0 = &-(e^\Delta-1)^2k^2 + [(e^{2\Delta}-1)((r+1)\Delta + \frac{1}{2}(e^\Delta-1)(e^\Delta-3)) \\ &-(e^\Delta-1)^3e^\Delta + \frac{1}{2}(e^\Delta+1)(e^\Delta-1)^3]k \\ &+ \frac{1}{2}(e^{2\Delta}-1)[(r+1)\Delta + \frac{1}{2}(e^\Delta-1)(e^\Delta-3)] - \frac{1}{4}(e^\Delta-1)^4, \end{aligned} \tag{5.39}$$

or,

$$0 = -(e^\Delta-1)k^2 + [(r+1)\Delta(e^\Delta+1) - 2(e^\Delta-1)]k + \frac{1}{2}(e^\Delta+1)(r+1)\Delta+e^\Delta-1. \tag{5.40}$$

In the two-dimensional case, the equation describing k_3 that results from the Lyapunov equation

$$\phi_0{}'K\phi_0 = K - Q_0 \tag{5.41}$$

is

$$k_3[e^\Delta - F(e^\Delta-1)]^2 \approx k_3 - \hat{R}F^2 + \frac{1}{2}(1-e^{2\Delta}) + F(e^\Delta-1)^2, \tag{5.42}$$

where $\hat{R}$ is as in (5.14).

Using the rules $\varepsilon k_2 \approx 0$, $\varepsilon \ell_2 \approx 0$, and $k_2\ell_2 \approx 0$, we can make the approximations from (5.20) and (5.22),

$$\hat{R} + D'KD \approx (r+1)\Delta + (e^\Delta-1)^2k_3 + \frac{1}{2}(e^\Delta-1)(e^\Delta-3), \tag{5.43}$$

$$M' + D'K\phi \approx \left[\varepsilon\left(\frac{2\Delta+e^{-2\Delta}+3}{4}\right) + k_2(1-e^{-\Delta}) + k_3e^{\Delta}(e^{\Delta}-1) + \frac{(e^{\Delta}-1)^2}{2}\right], \tag{5.44}$$

and so, from (5.26),

$$F \approx \frac{k_3e^{\Delta}(e^{\Delta}-1) + \frac{1}{2}(e^{\Delta}-1)^2}{(r+1)\Delta+(e^{\Delta}-1)\left[\frac{1}{2}(e^{\Delta}-3)+(e^{\Delta}-1)k_3\right]} \tag{5.45}$$

Now direct substitution of (5.45) into (5.41) would give us a cubic expression in k_3; we notice, however, that in the subexpression

$$k_3F^2(e^{\Delta}-1)^2 + \hat{R}F^2 = F^2(k_3(e^{\Delta}-1)^2 + \hat{R}), \tag{5.46}$$

the coefficient of F^2 is just the denominator of F; in fact, (5.46) is just

$$F(k_3e^{\Delta}(e^{\Delta}-1) + \frac{1}{2}(e^{\Delta}-1)^2). \tag{5.47}$$

Equation (5.42) is then just

$$\begin{aligned} k_3e^{2\Delta} - 2k_3F(e^{\Delta}-1)e^{\Delta} + F(k_3e^{\Delta}(e^{\Delta}-1) + \tfrac{1}{2}(e^{\Delta}-1)^2) \\ = k_3 + \tfrac{1}{2}(1-e^{2\Delta}) + F(e^{\Delta}-1)^2 \end{aligned} \tag{5.48}$$

which simplifies to

$$k_3(e^{2\Delta}-1) - Fk_3e^{\Delta}(e^{\Delta}-1) = \frac{1}{2}(1 - e^{2\Delta}) + \frac{1}{2}(e^{\Delta}-1)^2F. \tag{5.49}$$

If we divide by $(e^{\Delta}-1)$, and substitute (5.45) into (5.49), we have

$$\begin{aligned} -k_3^2(e^{\Delta}-1) + k_3[(e^{\Delta}+1)(r+1)\Delta + \tfrac{1}{2}(e^{2\Delta}-1)(e^{\Delta}-3) - \tfrac{1}{2}e^{\Delta}(e^{\Delta}-1)^2] \\ + \tfrac{1}{2}(e^{\Delta}+1)(r+1)\Delta + \tfrac{1}{4}(e^{2\Delta}-1)(e^{\Delta}-3) + \tfrac{1}{2}k_3(e^{\Delta}+1)(e^{\Delta}-1)^2 \\ - \tfrac{1}{4}(e^{\Delta}-1)^3 - \tfrac{1}{2}k_3e^{\Delta}(e^{\Delta}-1)^2 = \end{aligned}$$

$$= k_3^2(e^{\Delta}-1) + k_3[(e^{\Delta}+1)(r+1)\Delta - 2(e^{\Delta}-1)]$$

$$+ \frac{1}{2}(e^{\Delta}+1)(r+1)\Delta - 1 + e^{\Delta} = 0, \tag{5.50}$$

which is exactly the expression in (5.40). Thus, if we let $k(\Delta)$ be the solution of (5.40), the matrix K becomes

$$K = \begin{bmatrix} \frac{1}{2} & \varepsilon\tilde{k}_2 \\ \varepsilon\tilde{k}_2 & k(\Delta) \end{bmatrix}, \tag{5.51}$$

and if the initial state is given as

$$\underline{x}(0) = \begin{bmatrix} x_{01} \\ x_{02} \end{bmatrix}, \tag{5.52}$$

the resultant cost is

$$J_p(\Delta) = \frac{1}{2}\left[\frac{1}{2}x_{01}^2 + 2\varepsilon k_2 x_{01}x_{02} + k(\Delta)x_{02}^2\right]. \tag{5.53}$$

We now go to the case of full-state feedback control. Here, F*, the feedback matrix, will be 2-dimensional, e.g.,

$$F^* = [f_1 \quad f_2]; \tag{5.54}$$

we will also denote the steady-state solution of the discrete Riccati equation (3.14) as

$$K^* = \begin{bmatrix} k_1^* & k_2^* \\ k_2^* & k_3^* \end{bmatrix}. \tag{5.55}$$

Unfortunately, in this case we will not have any terms that are zero in the closed loop transition matrix, unlike the previous case of partial-state feedback. Thus, a direct attempt to solve (3.14) will result in three coupled non-linear equations. We can, however, make the simplifying approximations

$$K_1^* \approx \overline{k}_1 + \varepsilon\hat{k}_1 \tag{5.56}$$

$$K_2^* \approx \overline{k}_2 + \varepsilon\hat{k}_2 \tag{5.57}$$

$$K_3^* \approx \overline{k}_3 + \varepsilon\hat{k}_3, \tag{5.58}$$

where $\overline{k}_1$, $\hat{k}_1$, $\overline{k}_2$, $\hat{k}_2$, $\overline{k}_3$, $\hat{k}_3$, are independent of ε, and need to be determined. These follow from the Taylor series expansions

$$k_i(\varepsilon) = k_i(0) + \varepsilon\frac{\partial k_i}{\partial\varepsilon}(0) + \frac{\varepsilon^2}{2!}\frac{\partial^2 k_i}{\partial\varepsilon^2}(0) + \ldots \tag{5.59}$$

and our assumption that $\varepsilon^j \approx 0$ for $j \geq 2$. Note that we are implicitly assuming that $k_i(\varepsilon)$ is analytic in some region around $\varepsilon = 0$ (see Comment 1 at the end of this chapter).

In fact, since $\overline{k}_i = k_i(0)$, we can easily see that $\overline{k}_1 = \frac{1}{2}$, $\overline{k}_2 = 0$, and $\overline{k}_3 = k(\Delta)$. Then the resultant cost becomes

$$J_f(\Delta) = \frac{1}{2}\left[\left(\frac{1}{2} + \varepsilon\hat{k}_1\right)x_{01}^2 + 2\varepsilon\hat{k}_2 x_{01}x_{02} + (k(\Delta) + \varepsilon\hat{k}_3)x_{02}^2\right]. \tag{5.60}$$

We can now compare costs (5.53) and (5.60). First we make the reasonable assumption that we have preselected a <u>maximum</u> sampling period, that is, Δ is bounded above as well as below for both the full and partial-state feedback algorithms. Then, since the discrete time system (5.15) is controllable for all $\Delta \neq 0$ and $\varepsilon \neq 0$, we know $|\hat{k}_1|$, $|\hat{k}_2|$, $|\hat{k}_3|$ and $|\hat{k}_2|$ each have an upper bound over the acceptable range of Δ. Suppose, then, that the sampling rate for full-state feedback is

Δ^* and that for part-state feedback is $\frac{1}{2}\Delta^*$. Then the difference in the two costs is

$$J_f(\Delta^*) - J_p(\tfrac{1}{2}\Delta^*) = \frac{1}{2}\Big[\varepsilon\hat{k}_1 x_{01}^2 + 2\varepsilon(\hat{k}_2-\tilde{k}_2)x_{01}x_{02} + \varepsilon\hat{k}_3 x_{02}^2 + (k(\Delta^*) - k(\tfrac{1}{2}\Delta^*))x_{02}^2\Big] \tag{5.61}$$

Now, as we showed in Chapter II,

$$k(\Delta^*) - k(\tfrac{1}{2}\Delta^*) > 0; \tag{5.62}$$

we will define

$$\delta \triangleq k(\Delta^*) - k(\tfrac{1}{2}\Delta^*). \tag{5.63}$$

Thus, for any sampling period Δ^* and any bounded set of initial states such that for some system of the form (5.1) we have

$$\varepsilon(\hat{k}_1 \frac{x_{01}^2}{x_{02}^2} + 2(\hat{k}_2-\tilde{k}_2)\frac{x_{01}}{x_{02}} + \hat{k}_3) < \delta, \tag{5.64}$$

we may conclude that

$$J_f(\Delta^*) - J_p(\tfrac{1}{2}\Delta^*) > 0. \tag{5.64}$$

In other words, there will be some system (characterized by ε) for which (5.65) holds. Of course, ε must be sufficiently small for our approximation $\varepsilon^2 \approx 0$ to hold.

<u>Comments</u>

1. Our assumption that $K(\varepsilon)$ is analytic is indeed justified. Results of this nature for the differential Riccati equation follow directly from a theorem of Pontryagin [14] (pp. 170-181; also see

Kokotovic [8]). We may extend these results as follows: consider the discrete-time equation (3.14) to be of the form $K_i = f(K_{i+1})$. Then $\hat{K} = f(\hat{K})$ is the steady state solution. Now consider the equation $\frac{dK(t)}{dt} = f(K(t)) - K(t)$. From Pontryagin's conditions, each solution $K(t)$ of this equation is an analytic function of ε. But surely one solution is $K(t) = \hat{K}$, that is, $K(t)$ is constant. Thus, $\hat{K}$ is an analytic function of ε.

2. The conclusion of this chapter can be extended to the more general system

$$\dot{x}(t) = \begin{bmatrix} b & \varepsilon \\ 0 & a \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t),$$

where $b \leq 0$. Then, the new Φ matrix takes on the form

$$\Phi(\Delta) = \begin{bmatrix} e^{b\Delta} & \varepsilon f(a,b,\Delta) \\ 0 & e^{a\Delta} \end{bmatrix},$$

where $f(a,b,\Delta) = \frac{1}{a}\sinh(a\Delta)$, if $b = -a$, and $f(a,b,\Delta) = \frac{e^{b\Delta}-e^{a\Delta}}{b-a}$, if $|b| \neq |a|$. In either case, the reasoning of this chapter holds up in every detail.

3. In fact, there are apparently n-dimensional analogues of this example. We will return to this in Chapter VI. The example of this chapter is also relevant to the dynamic compensator of Chapter IV, if we suppose that, for (5.1), we are only able to observe $x_2(t)$. In this case, partial-state feedback is still superior to compensation, as compensation could not give us even as good a result as the optimal

full-state feedback law.

Chapter VI

Conclusion

The main contribution of this thesis, in our opinion, has been to formally state how the fundamental characteristics of digital computation cause the sampling interval Δ, as well as the usual feedback matrix, to be an optimization parameter. In this light, we have developed several on-line control algorithms, particularly a partial-state feedback law (Chapter III), and an output compensator (Chapter IV). We were able to explicitly state how, due to the question of relative computational complexity, one algorithm might be preferable to another; we derived an example (Chapter V) of a system where this was clearly illustrated, in order to justify our general point of view.

Since this work is intended mainly to provide a framework for further research, perhaps it is more appropriate to discuss the key questions that remain to be considered. In particular, we have not been able to devise a method for selecting the optimal sampling interval $\Delta = \Delta^*$ for a general class of systems, even for a particular control algorithm. One reason for this, as we have mentioned is the non-monotonicity of cost with respect to Δ. While this fact is disappointing, however,, it is not a fundamental obstacle to solution of the above problem. In many situations that occur in the theory of optimal control, it is possible to optimize a control over two parameters; we do not usually depend on the fact that the optimal choice for one parameter is an extremum of its permissible range. A far more serious question is that of "separability". The technique used to develop the optimal compensator (Chapter IV), for example, relied on the fact that the observer and feedback parts could be designed independently. In the case of Δ and F, however, it is foolish to talk about separability; the optimal F fundamentally depends on our choice of Δ.

The example of Chapter V, however, does suggest how we might extend this work to produce some useful results. Let us consider the general ε-coupled system of Kokotovic [8] in the partioned form

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_1 & \varepsilon A_{12} \\ \varepsilon A_{21} & A_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 & \varepsilon B_{12} \\ \varepsilon B_{21} & B_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{6.1}$$

where x_1 and x_2 are n_1- and n_2-dimensional substates, $n_1 + n_2 = n$, and u_1 and u_2 are r_1- and r_2-dimensional subcontrols, $r_1 + r_2 = r$. Then, the discrete optimal full-state feedback control (3.6)-(3.14) will involve rn online multiplications (and r(n-1) online additions, which we will ignore). If we were merely to consider the systems

$$\dot{x}_{10} = A_1 x_{10} + B_1 u_{10} \tag{6.2}$$

$$\dot{x}_{20} = A_2 x_{20} + B_2 u_{20}, \tag{6.3}$$

we would find that the optimal control would involve only $r_1 n_1 + r_2 n_2$ multiplications, which is less than rn. We suggest that, using the assumptions and techniques of Chapter V, that for any Δ, the difference between the costs resulting from optimally controlling (6.1) or controlling (6.2)-(6.3) is a linear function of ε. Note, however, that unlike the example of Chapter V, this does not demonstrate the superiority of the suboptimal control over the optimal; the earlier example depended on the fact that the costs of the decoupled subsystems were monotonic functions of Δ. Thus, this approach will not be helpful in controlling (6.1) unless we know something of the behavior of (6.2) and (6.3), and are able to study system performance as a function of Δ.

In conclusion, it is clear that the use of digital computers in controlling large-scale systems will require the examination of issues such as those outlined in this thesis. It is hoped that our efforts will provide some useful insights for future researchers in this area.

Appendix A

We herein prove that

$$\frac{\partial K(\Delta)}{\partial \Delta} \geq 0, \tag{A.1}$$

for all $\Delta \geq 0$, where K is given in equation (2.14).

If we define

$$G(\Delta) = \frac{\Delta(\phi+1)(ra^2+qb^2)}{2b^2 a(\phi-1)}, \tag{A.2}$$

we see immediately that

$$G(\Delta) \geq 0, \tag{A.3}$$

for all a, and

$$\hat{K}(\Delta) = aG(\Delta) - \frac{q}{a} + \sqrt{G(\Delta)a(aG(\Delta) - \frac{q}{a})} . \tag{A.4}$$

Moreover, since the scalar system (2.4) is obviously observable, the solution (A.4) must be positive definite (see [10]) and thus the radical in (A.4) must be real, i.e.,

$$G(\Delta)a(aG(\Delta) - \frac{q}{a}) \geq 0, \text{ all a.} \tag{A.5}$$

Now we calculate

$$\frac{\partial \hat{K}(\Delta)}{\partial \Delta} = aG'(\Delta) + \frac{\frac{1}{2} G'(\Delta)\,[2a^2 G(\Delta) - q]}{\sqrt{a^2 G(\Delta)^2 - qG(\Delta)}} \tag{A.6}$$

This will have the same sign as

$$F(\Delta) = aG'(\Delta)\left[\sqrt{a^2G(\Delta)^2 - qG(\Delta)} + aG(\Delta) - \frac{q}{2a}\right] \tag{A.7}$$

where

$$G'(\Delta) = \frac{(ra^2+qb^2)(\phi^2-2a\Delta\phi-1)}{2b^2a(\phi-1)^2} \tag{A.8}$$

and $aG'(\Delta)$ will have the same sign as

$$L(a\Delta) = \phi^2 - 2a\Delta\phi - 1. \tag{A.9}$$

Then, we calculate

$$\frac{\partial L}{\partial \phi} = 2(\phi - \ell n\phi - 1), \tag{A.10}$$

$$\frac{\partial^2 L}{\partial \phi^2} = 2(1 - \frac{1}{\phi}), \tag{A.11}$$

and note that $\frac{\partial L}{\partial \phi} = 0$ at $a\Delta = 0$ $(\phi = 1)$. Since $\frac{\partial^2 L}{\partial \phi^2}$ takes on the same sign as $a\Delta$, $\frac{\partial L}{\partial \phi}$ has a minimum at $a\Delta = 0$, and thus $\frac{\partial L}{\partial \phi} \geq 0$, and $\frac{\partial L}{\partial a} \geq 0$. But since $L(0) = 0$, $L(a\Delta)$ will take the same sign as $a\Delta$. Then $aG'(\Delta)$ must have the same sign as $a\Delta$, and $G'(\Delta) \geq 0$.

Now, suppose $a \leq 0$; then

$$\sqrt{a^2G^2(\Delta) - qG(\Delta)} + aG(\Delta) - \frac{q}{2a}$$
$$= -aG(\Delta)\left[\sqrt{1 - q/a^2G(\Delta)} - 1 + \frac{q}{2a^2G(\Delta)}\right]. \tag{A.12}$$

If we denote

$$X(\Delta) = \frac{q}{a^2G(\Delta)} , \tag{A.13}$$

then (A.12) becomes

$$-aG(\Delta)\left[\sqrt{1-X} - 1 + \frac{X}{2}\right]. \tag{A.14}$$

Clearly $X(\Delta)$ will be at a maximum for the value of $\Delta \geq 0$ for which $G(\Delta)$ is minimal; since $G'(\Delta) \geq 0$, this point occurs at

$$G(0) = \frac{ra^2+qb^2}{a^2b^2}, \tag{A.15}$$

and thus

$$X(\Delta) \leq \frac{qb^2}{ra^2+qb^2} \leq 1 \quad \text{for all a, all } \Delta \geq 0. \tag{A.16}$$

Since $G(\Delta) \geq 0$ for all $a\Delta$,

$$X(0) \geq 0, \quad \text{all a,} \quad \text{all } \Delta \geq 0. \tag{A.17}$$

We can use the Taylor Series expansion

$$\sqrt{1-X} = 1 - \frac{X}{2} - \frac{X^2}{4(2!)} - \frac{3X^3}{8(3!)} - \frac{15X^4}{16(4!)} - \ldots \tag{A.18}$$

and note that, for all a and all $\Delta \geq 0$,

$$\sqrt{1-X} - 1 + \frac{X}{2} < 0. \tag{A.19}$$

Thus (A.12) and (A.14) are negative. Since for $a\Delta \leq 0$, $aG'(\Delta) \leq 0$, we can conclude that $F(\Delta) \geq 0$, for $a \leq 0$, and

$$\frac{\partial\hat{K}(\Delta)}{\partial\Delta} \geq 0, \tag{A.20}$$

for $a \leq 0$.

Now suppose $a > 0$. Then (A.12) becomes

$$aG(\Delta)\left[\sqrt{1-X} + 1 - \frac{X}{2}\right]. \tag{A.21}$$

Since $0 \leq \frac{X}{2} \leq \frac{1}{2}$, the expression in (A.21) is now positive. Since $aG'(\Delta) \geq 0$, we may then conclude

$$\frac{\partial \hat{K}(\Delta)}{\partial \Delta} \geq 0 \qquad \text{(A.22)}$$

for all a, and all $\Delta \geq 0$.

Appendix B

We fill in the gaps in the derivation of Chapter IV.

B.1 Evaluation of Λ.

If we define

$$\Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}' & \Lambda_{22}' \end{bmatrix} \tag{B.1.1}$$

then the Lyapunov equation of equation (4.34) is equivalent to

$$(A' + G^{*\prime}B')\Lambda_{11}(A + BG^*) + Q + G^{*\prime}RG^* + MG^* + G^{*\prime}M' = \Lambda_{11} \tag{B.1.2}$$

$$-(A' + G^{*\prime}B')\Lambda_{11}BG^*N + (A' + G^{*\prime}B')\Lambda_{12}F - G^{*\prime}RG^*N - MG^*N = \Lambda_{12} \tag{B.1.3}$$

$$N'G^{*\prime}B'\Lambda_{11}BG^*N - F'\Lambda_{12}'BG^*N - N'G^{*\prime}B'\Lambda_{12}F + F'\Lambda_{22}F + N'G^{*\prime}RG^*N = \Lambda_{22}. \tag{B.1.4}$$

Now, comparison of (B.1.2) with equation (3.13) shows that Λ_{11} is indeed the steady-state solution of the full-state output regulator equation. We will denote Λ_{11} as K. Then from (3.11) we see that

$$G^* = R^{-1}M' - (R + B'KB)^{-1}B'K[A - BR^{-1}M']. \tag{B.1.5}$$

Then, (B.1.3) may be rewritten, using (B.1.5)

$$\begin{aligned}\Lambda_{12} = &-(A'+G^{*\prime}B')KBG^{*}N + (A'+G^{*\prime}B')\Lambda_{12}F - MR^{-1}M'N \\ &- (A'-MR^{-1}B')KB(R+B'KB)^{-1}R\ (R+B'KB)^{-1}B'K(A-BR^{-1}M')N \\ &- M(R+B'KB)^{-1}B'K(A-BR^{-1}M')N - (A'-MR^{-1}B')KB(R+B'KB)^{-1}M'N \\ &+ MR^{-1}M'N + M[R+B'KB]^{-1}B'K(A-BR^{-1}M^{1})N, \qquad (B.1.6)\end{aligned}$$

which through judicious substitution of (B.1.5) and several cancellations simplifies to

$$\begin{aligned}\Lambda_{12} = &(-A'+MR^{-1}B')KBG^{*}N + (A'-MR^{-1}B')KB(R+B'KB)^{-1}B'KBG^{*}N \\ &+ (A+G^{*\prime}B)\Lambda_{12}F + (A'-MR^{-1}B')KB(R+B'KB)^{-1}RG^{*}N, \qquad (B.1.7)\end{aligned}$$

and if we recognize the pattern

$$\begin{aligned}&(A'-MR^{-1}B')K(-B+B(R+B'KB)^{-1}B'KB - B(R+B'KB)^{-1}R)G^{*}N \\ &= (A'-MR^{-1}B')K(-B+B(R+B'KB)^{-1}(B'KB+R))G^{*}N \\ &= 0, \qquad (B.1.8)\end{aligned}$$

then (B.1.7) becomes

$$\Lambda_{12} = (A'+G^{*\prime}B')\Lambda_{12}F. \qquad (B.1.9)$$

If we assume that H* is restricted to values for which the error e (4.16) is stable, then $||F|| < 1$; if G* is chosen for stability of the plant, then $||A'+G^{*\prime}B'|| < 1$; thus,

$$||\Lambda_{12}|| \leq ||A'+G^{*\prime}B'||\ ||F||\ ||\Lambda_{12}|| < ||\Lambda_{12}||, \qquad (B.1.10)$$

unless $||\Lambda_{12}|| = 0$; therefore, $||\Lambda_{12}|| = 0$.

Equation (B.1.4) may now be simplified to

$$N'G^{*\prime}(B'KB+R)G^*N + F'\Lambda_{22}F = \Lambda_{22}, \qquad \text{(B.1.11)}$$

giving us

$$\Lambda = \begin{bmatrix} K & 0 \\ 0 & \Lambda_{22} \end{bmatrix}, \qquad \text{(B.1.12)}$$

so in fact the total cost is

$$\mathrm{tr}(K\Sigma_0) + \mathrm{tr}(\Lambda_{22}E_0). \qquad \text{(B.1.13)}$$

B.2 Choice of H*.

Clearly, we want to choose H to minimize the added cost $J = \mathrm{tr}(\Lambda_{22}E_0)$. We set

$$\left.\frac{\partial J}{\partial H}\right|_{H^*} = 0, \qquad \text{(B.2.1)}$$

and noting that Λ_{22} is the solution of Lyapunov equation (B.1.11), we see

$$\Lambda_{22} = \sum_{i=0}^{\infty} F^{i\prime} N'G^{*\prime}(B'KB+R)G^*NF^i. \qquad \text{(B.2.2)}$$

If we let

$$M_0 = N'G^{*\prime}(B'KB+R)G^*N, \qquad \text{(B.2.3)}$$

then we may rewrite $\mathrm{tr}(\Lambda_{22}E_0)$ as

$$J(H) = \mathrm{tr}\{M_0 \sum_{i=0}^{\infty} F(H)^i E_0(H) F(H)^{i\prime}\}. \qquad \text{(B.2.4)}$$

Again, as in Chapter III, we will make use of Kleinman's lemma summarized in (3.25)-(3.26) to determine

$$E_0(H+\varepsilon\Delta H) = E_0(H) + \varepsilon\Delta H(\Sigma_{11}H'-\Sigma_{12}) + \varepsilon(\Sigma_{11}H'-\Sigma_{12})'\Delta H' \qquad (B.2.5)$$

$$F(H+\varepsilon\Delta H) = F(H)-\varepsilon\Delta HA_{12}. \qquad (B.2.6)$$

Since

$$J(H+\varepsilon\Delta H) = tr\{M_0 \sum_{i=0}^{\infty} (F-\varepsilon\Delta HA_{12})^i (E_0+\varepsilon\Delta H(\Sigma_{11}H'-\Sigma_{12}) + \varepsilon(\Sigma_{11}H'-\Sigma_{12})'\Delta H')(F-\varepsilon\Delta HA_{12})^{i'}\}, \qquad (B.2.7)$$

using the approximation

$$(F-\varepsilon\Delta HA_{12})^i \approx F^i - \varepsilon \sum_{j=0}^{i-1} F^{i-1-j}\Delta HA_{12}F^j \qquad (B.2.8)$$

when $i \neq 0$ (and is the identity when $i = 0$), we may rewrite

$$J(H+\varepsilon\Delta H) \approx tr\{M_0 \sum_{i=0}^{\infty} [(F^i-\varepsilon T_i)E_0(F^i)' - \varepsilon F^i E_0 T_i' + F^i\varepsilon(\Delta H(\Sigma_{11}H'-\Sigma_{12}) + (\Sigma_{11}H'-\Sigma_{12})'\Delta H')F^{i'}]\}, \qquad (B.2.9)$$

where

$$T_i = \sum_{j=0}^{i-1} F^{i-1-j}\Delta HA_{12}F^j = \sum_{j=0}^{i-1} F^j\Delta HA_{12}F^{i-1-j} \ , \quad i > 0 \qquad (B.2.10)$$

($T_i \equiv 0$ when $i = 0$).

Using trace properties we can write

$$J(H+\varepsilon\Delta H)-J(H) = tr\{\varepsilon M_0 \sum_{i=0}^{\infty} [-T_iE_0F^{i'}-F^iE_0T_i' + F^{i'}(\Delta H(\Sigma_{11}H'-\Sigma_{12}) + (\Sigma_{11}H'-\Sigma_{12})'\Delta H')F^{i'}]\}=$$

$$= \mathrm{tr}\{\varepsilon M_0 \sum_{i=0}^{\infty} F^i(\Delta H(\Sigma_{11}H'-\Sigma_{12}) + (\Sigma_{11}H'-\Sigma_{12})'\Delta H')F^{i'}$$

$$- \varepsilon M \sum_{i=1}^{\infty} \sum_{j=0}^{i-1} [(F^{(i-1-j)'}A_{12}'\Delta H'F^{j'})'E_0F^{i'}$$

$$+ F^iE_0(F^{(i-1-j)}\Delta HA_{12}F^j)']\}$$

$$= \varepsilon \mathrm{tr}\ \{\sum_{i=0}^{\infty} 2(\Sigma_{11}H'-\Sigma_{12})F^{i'}M_0F^i\Delta H$$

$$- \sum_{i=1}^{\infty} \sum_{j=0}^{i-1} [A_{12}F^jE_0F^{i'}M_0F^{(i-1-j)}\Delta H + A_{12}F^{(i-1-j)}E_0F^{i'}M_0F^j\ \Delta H]\}.$$

(B.2.11)

We may now apply the Kleinman lemma to get

$$\frac{d}{dH} J(H) = 2 \sum_{i=0}^{\infty} F^{i'}M_0F^i(\Sigma_{11}H'-\Sigma_{12})'$$

$$- \sum_{i=1}^{\infty} \sum_{j=0}^{i-1} [F^{j'}E_0F^iM_0F^{(i-1-j)'} + F^{(i-1-j)'}E_0F^iMF^{j'}]A_{12}'$$

$$= 2 \sum_{i=0}^{\infty} F^{i'}M_0F^i(\Sigma_{11}H'-\Sigma_{12})'$$

$$- 2 \sum_{i=1}^{\infty} \sum_{j=0}^{i-1} [F^{j'}M_0F^iE_0F^{(i-1-j)'}]A_{12}' \qquad \text{(B.2.12)}$$

Now we can make the manipulation

$$\sum_{i=1}^{\infty} \sum_{j=0}^{i-1} [F^{j'}M_0F^iE_0F^{(i-1-j)'}] = \sum_{j=0}^{\infty} \sum_{i=j+1}^{\infty} F^{j'}M_0F^iE_0F^{(i-1-j)'}$$

$$= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} F^{j'}M_0F^{j+k+1}E_0F^{k'} = (\sum_{j=0}^{\infty} F^{j'}M_0F^j)(\sum_{k=0}^{\infty} F^{k+1}E_0F^{k'}) \qquad \text{(B.2.13)}$$

If we let

$$W = \sum_{k=0}^{\infty} F^k E_0 F^{k'} \qquad \text{(B.2.14)}$$

so that

$$FWF' = W - E_0, \qquad \text{(B.2.15)}$$

and set $\frac{d}{dH} J(H^*) = 0$, (4.63) becomes

$$0 = (2 \sum_{l=0}^{\infty} F^l M F^{l'})[(\Sigma_{11}H^{*'}-\Sigma_{12})' - FWA_{12}'], \qquad \text{(B.2.16)}$$

so in fact

$$H^*\Sigma_{11}-\Sigma_{12}' = FWA_{12}' = (A_{22}-H^*A_{12})WA_{12}'. \qquad \text{(B.2.17)}$$

Solving this for H*, we see

$$H^* = (\Sigma_{12}' + A_{22}WA_{12}')(\Sigma_{11} + A_{12}WA_{12}')^{-1}. \qquad \text{(B.2.18)}$$

We may then rewrite (B.2.18), using (B.2.15), to obtain

$$\begin{aligned} W &= A_{22}WA_{22}' - H^*A_{12}WA_{22}' + H^*A_{12}'WA_{12}'H^{*'} - A_{22}WA_{12}'H^{*'} \\ &\quad + H^*\Sigma_{11}H^{*'} - \Sigma_{22} + H^*\Sigma_{12} + \Sigma_{12}'H^{*'} \\ &= A_{22}WA_{22}' - H^*(A_{12}WA_{22}' + \Sigma_{12}) + H^*(A_{12}WA_{12}'H^{*'} + \Sigma_{11})H^{*'} \\ &\quad - (A_{22}WA_{12}' + \Sigma_{12}')H^{*'} + \Sigma_{22} \end{aligned} \qquad \text{(B.2.19)}$$

so

$$\begin{aligned} W &= A_{22}WA_{22}' - (\Sigma_{12}' + A_{22}WA_{12}')(\Sigma_{11} + A_{12}WA_{12}')^{-1}(A_{12}WA_{22}' + \Sigma_{12}) \\ &\quad + \Sigma_{22} \end{aligned} \qquad \text{(B.2.20)}$$

and we get the results of equations (4.51)-(4.55).

Appendix C

Reduction of Results of Chapter III to the Full-State Solution

Here, we let $C = I$ in equations (3.42)-(3.46) and show the results agree exactly with the work of Levis [10].

Equation (3.42) becomes

$$F = (R + D'KD)^{-1}(M' + D'K\phi) \tag{C.1}$$

which implies

$$\begin{aligned}(R + D'KD)F &= M' + D'K\phi \\ &= M' + D'KDR^{-1}M' + D'K\phi - D'KDR^{-1}M' \\ &= (R + D'KD)R^{-1}M' + D'K(\phi - DR^{-1}M'),\end{aligned} \tag{C.2}$$

so

$$F = R^{-1}M' + (R + D'KD)^{-1}D'K(\phi - DR^{-1}M'). \tag{C.3}$$

Moreover, equation (3.37) becomes

$$\begin{aligned}K &= \phi_0{}'K\phi_0 + Q_0 = (\phi - DF)'K(\phi - DF) + Q_0 \\ &= (\phi - DR^{-1}M')'K(\phi - DR^{-1}M') \\ &- ((\phi - DR^{-1}M')'KD(R + D'KD)^{-1}D')K(\phi - DR^{-1}M') \\ &- (\phi - DR^{-1}M')'K(D(R + D'KD)^{-1}D'K(\phi - DR^{-1}M')) + Q_0 \\ &+ ((\phi - DR^{-1}M')'KD(R + D'KD)^{-1}D'KD(R + D'KD)^{-1}D'K(\phi - DR^{-1}M')) \\ &= Q_0 + (\phi - DR^{-1}M')'[K - 2KD(R + D'KD)^{-1}D'K \\ &+ KD(R + D'KD)^{-1}D'KD(R + D'KD)^{-1}D'K](\phi - DR^{-1}M')\end{aligned} \tag{C.4}$$

Now Q_0 is

$$
\begin{aligned}
Q_0 &= Q - MF - F'M' + F'RF \\
&= Q - MR^{-1}M' - M(R + D'KD)^{-1}D'K(\phi - DR^{-1}M') \\
&- (\phi - DR^{-1}M')'KD(R + D'KD)^{-1}M' - MR^{-1}M' \\
&+ MR^{-1}M' + M(R + D'KD)^{-1}D'K(\phi - DR^{-1}M') \\
&+ (\phi - DR^{-1}M')'KD(R + D'KD)^{-1}M' \\
&+ (\phi - DR^{-1}M')'KD(R + D'KD)^{-1}R(R + D'KD)^{-1}D'K(\phi - DR^{-1}M') \\
&= Q - MR^{-1}M' + (\phi - DR^{-1}M')'KD(R + D'KD)^{-1}R(R + D'KD)^{-1}D'K \\
&\cdot(\phi - DR^{-1}M'), \qquad (C.5)
\end{aligned}
$$

Now, to combine Q_0 with the third term in brackets in (C.4), we note

$$
\begin{aligned}
&KD(R + D'KD)^{-1}R(R + D'KD)^{-1}D'K \\
&\quad + KD(R + D'KD)^{-1}D'KD(R + D'KD)^{-1}D'K \\
&\quad = KD(R + D'KD)^{-1}(R + D'KD)(R + D'KD)^{-1}D'K \\
&\quad = KD(R + D'KD)^{-1}D'K \qquad (C.6)
\end{aligned}
$$

which we may combine with the remainder of (C.4) to show

$$
\begin{aligned}
K &= (\phi - DR^{-1}M')'[K - KD(R + D'KD)^{-1}D'K](\phi - DR^{-1}M') \\
&\quad + (Q - MR^{-1}M'). \qquad (C.7)
\end{aligned}
$$

Equations (C.3) and (C.7) agree exactly with the results of Levis.

Bibliography

[1] Athans, M. and P.L. Falb, Optimal Control: An Introduction to the Theory and Its Applications, McGraw-Hill, Inc., New York, 1966.

[2] Blanvillain, P.J.-P., "Optimal Compensator Structure for Linear Time-Invariant Plant with Inaccessible States", Report ESL-R-546, Electronic Systems Laboratory, M.I.T., Cambridge, Mass., May 1974.

[3] Kalman, R.E., "Contributions to the Theory of Optimal Control", Soc. Mat. Mex. Bol., Vol. 5, Ser. 2, Num. 1, pp. 102-119, April 1960.

[4] Kalman, R.E. and R. Bucy, "New Results in Linear Filtering and Prediction", J. Basic. Eng. (ASME), Vol. 83D, pp. 95-108, March 1961.

[5] Kleinman, D.L., "On the Linear Regulator Problem and the Matrix Riccati Equation", Report ESL-R-271, Electronic Systems Laboratory, M.I.T., Cambridge, Mass., 1966.

[6] Kleinman, D.L. and M. Athans, "The Design of Suboptimal Linear Time-Varying Systems", IEEE Trans. on Automatic Control, Vol. AC-13, pp. 150-159, April 1968.

[7] Kleinman, D.L., T. Fortmann, and M. Athans, "On the Design of Linear Systems with Piecewise Constant Feedback Gains", Preprints of Ninth Joint Automatic Control Conference, pp. 698-710, June 1968.

[8] Kokotovic, P.V., "Feedback Design of Large Linear Systems", from Feedback Systems (ed., J.B. Cruz, Jr.), pp. 99-137, McGraw-Hill, Inc., New York, 1972.

[9] Levine, W.S., "Optimal Output-Feedback Controllers for Linear Systems", Ph.D. Thesis, Dept. of Electrical Eng., M.I.T., Cambridge, Mass., January 1969.

[10] Levis, A.H., "On the Optimal Sampled-Data Control of Linear Processes", Sc.D Thesis, Dept. of Mechanical Eng., M.I.T., Cambridge, Mass., June 1968.

[11] Levis, A.H. and M. Athans, "On the Optimal Sampled Data Control of Strings of Vehicles", Transportation Sci., Vol. 2, No. 4, Nov. 1968.

[12] Luenberger, D.G., "An Introduction to Observers", IEEE Trans. on Automatic Control, Vol. AC-16, No. 6, pp. 596-602, Dec. 1971.

[13] Miller, R.A., "Specific Optimal Control of the Linear Regulator Using a Minimal-Order Observer", Int. J. Control, Vol. 18, No. 1, pp. 139-159, 1973.

[14] Pontryagin, L.S., Ordinary Differential Equations, Chapter 4, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1962.

[15] Wonham, W.M., "On Pole Assignment in Multi-input Controllable Linear Systems", IEEE Trans. on Automatic Control, Vol. AC-12, pp. 660-665, December 1967.