AD-A200 257

NASA Contractor Report 181705

ICASE REPORT NO. 88-48

(2)

# ICASE

## APPROXIMATION THEORY FOR LQG OPTIMAL CONTROL OF FLEXIBLE STRUCTURES

DTIC
ELECTE
S OCT 1 2 1988 D
GMD

J. S. Gibson

A. Adamian

Contract Nos. NAS1-17070 and NAS1-18107
August 1988

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING
NASA Langley Research Center, Hampton, Virginia 23665

Operated by the Universities Space Research Association

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

88 1011 051

# APPROXIMATION THEORY FOR LQG OPTIMAL CONTROL OF FLEXIBLE STRUCTURES

J. S. Gibson and A. Adamian
Mechanical, Aerospace and Nuclear Engineering
University of California, Los Angeles, CA 90024

## ABSTRACT

This paper presents approximation theory for the linear-quadratic-Gaussian optimal control problem for flexible structures whose distributed models have bounded input and output operators. The main purpose of the theory is to guide the design of finite dimensional compensators that approximate closely the optimal compensator separates into an optimal linear-quadratic control problem lies in the solution to an infinite dimensional Riccati operator equation. The approximation scheme in the paper approximates the infinite dimensional LQG problem with a sequence of finite dimensional LQG problems defined for a sequence of finite dimensional, usually finite element or modal, approximations of the distributed model of the structure. Two Riccati matrix equations determine the solution to each approximating problem.

The finite dimensional equations for numerical approximation are developed, including formulas for converting matrix control and estimator gains to their functional representation to allow comparison of gains based on different orders of approximation. Convergence of the approximating control and estimator gains and of the corresponding finite dimensional compensators is studied. Also, convergence and stability of the closed-loop systems produced with the finite dimensional compensators are discussed. The convergence theory is based on the convergence of the solutions of the finite dimensional Riccati equations to the solutions of the infinite dimensional Riccati equations. A numerical example with a flexible beam, a rotating rigid body, and a lumped mass is given.

This research was supported by the National Aeronautics and Space Administration under NASA Contract Nos. NAS1-17070 and NAS1-18107 while the first author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23665. Also, research for both authors was supported by the United States Air Force under AFOSR Grant 840309 and by the Jet Propulsion Laboratory, Pasadena, CA under Grant 957114.

CONTENTS


Abstract

1. Introduction ..... 1

2. The Control System ..... 12
   2.1 The Energy Spaces and the First-order Form of the System ..... 13
   The Elastic-strain-energy Space V and the Total-energy Space E ..... 13
   The Damping Functional and Operator ..... 15
   The Semigroup Generator ..... 16
   2.2 The Adjoint of $\tilde{A}$ ..... 19
   2.3 Exponential Stability ..... 19

3. The Optimal Control Problem ..... 22
   3.1 The Generic Optimal Regulator Problem ..... 22
   3.2 Application to Optimal Control of Flexible Structures ..... 28

4. The Approximation Scheme ..... 32
   4.1 Approximation of the Open-loop System ..... 32
   4.2 The Approximating Optimal Control Problems ..... 39

5. Convergence ..... 45
   5.1 Generic Approximation Results ..... 45
   5.2 Convergence of the Approximating Optimal Control Problems of Section 4.2 ..... 48

6. Example ..... 53
   6.1 The Control System ..... 53
   6.2 The Optimal Control Problem ..... 58
   6.3 Approximation ..... 59
   6.4 Numerical Results ..... 61



DTIC COPY INSPECTED 4

Accesion For
NTIS CRA&I
DTIC TAB
Unannounced
Justification
By
Distribution /
Availability Codes
Dist Avail and / or Special
A-1

7. The Optimal Infinite Dimensional Estimator, Compensator and Closed-loop System ..... 71
7.1 The Generic Problem ..... 71
7.2 Application to Structures ..... 76

8. Approximation of the Infinite Dimensional Estimator .. 79
8.1 The Approximating Finite Dimensional Estimators . 79
8.2 Stochastic Interpretation of the Approximating Estimators ..... 83
8.3 The Approximating Functional Estimator Gains .... 85
8.4 Convergence ..... 87

9. The Finite Dimensional Compensators and Realizable Closed-loop Systems ..... 90
9.1 Closing the Loop ..... 90
9.2 Convergence of the Closed-loop Systems ..... 92
9.3 Convergence of the Compensator Transfer Functions ..... 96

10. Closing the Loop in the Example ..... 99
10.1 The Estimator Problem ..... 99
10.2 Approximation of the Optimal Compensator ..... 107
10.3 Comments on the Structure and Dimension of the Implementable Compensators ..... 109

11. Conclusions ..... 111

References ..... 113

Appendix: Errata for [G1] ..... 116

## 1. Introduction

Recent years have seen increasing research in active control of flexible structures. The primary motivation for this research is control of large flexible aerospace structures, which are becoming larger and more flexible at the same time that their performance requirements are becoming more stringent. For example, in tracking and other applications, satellites with large antennae, solar collectors and other flexible components must perform fast slew maneuvers while maintaining tight control over the vibrations of their flexible elements. Both of these conflicting objectives can be achieved only with a sophisticated controller. There are applications also to control of robotic manipulators with flexible links, and possibly to stabilization of large civil engineering structures such as long bridges and tall buildings.

The first question that must be answered when designing a controller for a flexible structure is whether a finite dimensional model is sufficient as a basis for a controller that will produce the required performance, or is a distributed model necessary? While some structures can be modeled well by a fixed number of dominant modes, there are structures whose flexible character can be captured sufficiently for precise control only by a distributed model. Still others -- perhaps most of the aerospace structures of the future -- can be modeled sufficiently for control purposes by some finite dimensional approximation, but an adequate approximation may be impossible to determine before design of the controller, or compensator. This paper deals with structures that are flexible enough to require a distributed model in the design of an optimal LQG compensator.

The linear-quadratic-gaussian optimal control problem for distributed, or infinite dimensional, systems is a generalization to Hilbert space of the now classical LQG problem for finite dimensional systems. The solution to the infinite dimensional problem yields an infinite dimensional state-estimator-based compensator, which is optimal in the context of this paper. By a separation principle [B1, C4], the problem reduces to a deterministic linear-quadratic optimal control problem and an optimal estimation, or filtering, problem with gaussian white noise. In infinite dimensions, the control system dynamics are represented by a semigroup of bounded linear operators instead of the matrix exponential operators in finite dimensions, and the plant noise process may be an infinite dimensional random process. The solutions to both the control and filtering problems involve Riccati operator equations, which are generalizations of the Riccati matrix equations in the finite dimensional case. Current results on the infinite dimensional LQG problem are most complete for problems where the input and measurement operators are bounded, as this paper requires throughout. This boundedness also permits the strongest approximation results here. For related control problems with unbounded input and measurement, see [C3, C5, L1, L2].

Our primary objective in this paper is to approximate the optimal infinite dimensional LQG compensator for a distributed model of a flexible structure with finite dimensional compensators based on approximations to the structure, and to have these finite dimensional compensators produce near optimal performance of the closed-loop system. We discuss how the gains that determine the finite dimensional compensators converge to the gains that determine the infinite dimensional compensator, and we examine the sense in

which the finite dimensional compensators converge to the infinite dimensional compensator. With this analysis, we can predict the performance of the closed-loop system consisting of the distributed plant and a finite dimensional compensator that approximates the infinite dimensional compensator.

Our design philosophy is to let the convergence of the finite dimensional compensators indicate the order of the compensator that is required to produce the desired performance of the structure. The two main factors that govern rate of convergence are the desired performance (e.g., fast response) and the structural damping. We should note that any one of our compensators whose order is not sufficient to approximate the infinite dimensional compensator closely will not in general be the optimal compensator of that fixed order; i.e., the optimal fixed-order compensator that would be constructed with the design philosophy in [B7, B8]. But as we increase the order of approximation to obtain convergence, our finite dimensional compensators become essentially identical to the compensator that is optimal over compensators of all orders.

An important question, of course, is how large a finite dimensional compensator we must use to approximate the infinite dimensional compensator. In [G6, G7, G8, M1], we have found that our complete design strategy yields compensators of reasonable size for distributed models of complex space structures. This strategy in general requires two steps to obtain an implementable compensator that is essentially identical to the optimal infinite dimensional compensator: the first step determines the optimal compensator by letting the finite dimensional compensators converge to it; the second step reduces, if possible, the order of a large (converged) approximation to the optimal compensator. The first step, which is the one involving control theory and

approximation theory for distributed systems, is the subject of this paper. For the second step, a simple modal truncation of the large compensator sometimes is sufficient, but there are more sophisticated methods in finite dimensional control theory for order reduction. For example [G8, M1], we have found balanced realizations [M2] to work well for reducing large compensators.

The approximation theory in this paper follows from the application of approximation results in [B6, G3, G4] to a sequence of finite dimensional optimal LQG problems based on a Ritz-Galerkin approximation of the flexible structure. For the optimal linear-quadratic control problem, the approximation theory here is a substantial improvement over that in [G1] because here we allow rigid-body modes, more general structural damping (including damping in the boundary), and much more general finite element approximations. These generalizations are necessary to accommodate common features of complex space structures and the most useful finite element schemes. For example, we write the equations for constructing the approximating control and estimator gains and finite dimensional compensators in terms of matrices that are built directly from typical mass, stiffness and damping matrices for flexible structures, along with actuator influence matrices and measurement matrices.

For the estimator problem, this paper presents the first rigorous approximation theory. (We have used less complete versions of the results in previous research [G6, G7, G8, M1].) As in the finite dimensional case, the infinite dimensional optimal estimation problem is the dual of the infinite dimensional optimal control problem, and the solutions to both problems have the same structure. Because we exploit this duality to obtain the approximation theory for the estimation problem from the approximation theory for the

optimal control problem, the analysis in this paper is almost entirely deterministic. We discuss the stochastic interpretation of the estimation problem and the approximating state estimators briefly, but we are concerned mainly with deterministic questions about the structure and convergence of approximations to an infinite dimensional compensator and the performance -- especially stability -- of the closed-loop systems produced by the approximating compensators.

Next, an outline of the paper should help. The paper has two main parts, which correspond roughly to the separation of the optimal LQG problem into an optimal linear-quadratic regulator problem and an optimal state estimation problem. The first half, Sections 2 through 6, deal with the control system and the optimal regulator problem. Sections 7 through 10 treat the state estimator and the compensator that is formed by applying the control law of the first half of the paper to the output of the estimator.

Section 2 defines the abstract model of a flexible structure and the energy spaces to be used throughout the paper. We assume a finite number of actuators, since this is the case in all applications, and we assume that the actuator influence operator is bounded. Section 2 also establishes certain mathematical properties of the open-loop system that are useful in control and approximation. To our knowledge, the exponential stability theorem in Section 2.3 is a new result, and we find it interesting that such a simple Lyapunov functional accommodates such a general damping model.

Section 3 discusses the linear-quadratic optimal control problem for the distributed model of the structure and establishes some estimates involving

bounds on solutions to infinite dimensional Riccati equations and open-loop and closed-loop decay rates. We need these estimates for the subsequent approximation theory. To get the approximation theory for the estimation problem, we have to give certain results on the control problem in a more general form than would be necessary were we interested only in the control problem for flexible structures. Therefore, in Section 3, as in Sections 5 and 7, we first give some generic results applicable to the LQG problem for a variety of distributed systems and then apply the generic results to the control of flexible structures.

Because we assume a finite number of actuators and a bounded input operator, the optimal feedback control law consists of a finite number of bounded linear functionals on the state space, which is a Hilbert space. This means that that the feedback law can be represented in terms of a finite number of vectors, which we call functional control gains, whose inner products with the generalized displacement and velocity vectors define the control law. For any finite-rank, bounded linear feedback law for a control system on a Hilbert space, the existence of such gains is obvious and well known. A functional control gain for a flexible structure will have one or more distributed components, or kernels, corresponding to each distributed component of the structure and scalar components corresponding to each rigid component of the structure.

We introduce the functional control gains at the end of Section 3, and we introduce analogous functional estimator gains in Section 7. The functional gains play a prominent roll in our analysis. They give a concrete representation of the infinite dimensional compensator and provide a criterion for con-

vergence of the approximating finite dimensional compensators.

We develop the approximation scheme for the control problem in Section 4. The idea is to solve a finite dimensional linear-quadratic regulator problem for each of a sequence of Ritz-Galerkin approximations to the structure. We develop the approximation of the structure in Section 4.1 and prove convergence of the approximating open-loop systems. The approximation scheme includes most finite element approximations of flexible structures. For convergence, we use the Trotter-Kato semigroup approximation theorem, which was used in optimal open-loop control of hereditary systems in [B5] and has been used in optimal feedback control of hereditary, hyperbolic and parabolic systems in [B6, G1, G3] and other papers. The usual way to invoke Trotter-Kato is to prove that the resolvents of the approximating semigroup generators converge strongly. To prove this, we introduce an inner product that involves both the strain-energy inner product and the damping functional, and show that the resolvent of each finite dimensional semigroup generator is the projection, with respect to this special inner product, of the resolvent of the original semigroup generator onto the approximation subspace. The idea works as well for the adjoints of the resolvents, and when the open-loop semigroup generator has compact resolvent, it follows from our projection that the approximating resolvent operators converge in norm.

In Section 4.2, we define the sequence of finite dimensional optimal control problems, whose solutions approximate the solution to the infinite dimensional problem of Section 3. The solution to each finite dimensional problem is based on the solution to a Riccati matrix equation, and we give formulas for using the solution to the Riccati matrix equation to compute approxima-

tions to the functional control gains as linear combinations of the basis vectors.

Section 5.1 summarizes some generic convergence results from [B6, G3, G4] on approximation of solutions to infinite dimensional Riccati equations. Section 5.2 applies these generic results to obtain sufficient conditions for convergence -- and nonconvergence -- of the solutions of the approximating optimal control problems in Section 4.2. A main sufficient condition for convergence is that the structure have damping, however small, that makes all elastic vibrations of the open-loop system exponentially stable. This is a necessary condition if the state weighting operator in the control problem is coercive.

In Section 6, we present an example in which the structure consists of an Euler-Bernoulli beam attached on one end to a rotating rigid hub and on the other end to a lumped mass. We emphasize the fact that we do not solve, or even write down, the coupled partial and ordinary differential equations of motion. For both the definition and numerical solution of the problem, only the kinetic and strain energy functionals and a dissipation functional for the damping are required. We show the approximating functional control gains obtained by using a standard finite element approximation of the beam, and we discuss the effect on convergence of structural damping and of the ratio of state weighting to control weighting in the performance index. As suggested by a theorem in Section 5, the functional gains do not converge when no structural damping is modeled.

In Section 7, we begin the theory for closing the loop on the control system. We assume a finite number of bounded linear measurements and

construct the optimal state estimator, which is infinite dimensional in general. The gains for this estimator are obtained from the solution to an infinite dimensional Riccati equation that has the same form as the infinite dimensional Riccati equation in the control problem. We call these gains functional estimator gains because they are vectors in the state space.

Since the approximation issues that this paper treats are fundamentally deterministic, we make the paper self contained by defining the infinite dimensional estimator as an observer, although the only justification for calling this estimator and the corresponding compensator optimal is their interpretation in the context of stochastic estimation and control. We discuss the stochastic interpretation but do not use it. We say estimator and observer interchangeably to emphasize the deterministic definition of the estimator here.

With the optimal control law of Section 3 and the optimal estimator of Section 7, we construct the optimal compensator, which also is infinite dimensional in general. The transfer function of this compensator is irrational, but it is still an m(number of actuators) p(number of sensors) matrix function of a complex variable, as in finite dimensional control theory. The optimal closed-loop system consists of the distributed model of the structure controlled by the optimal compensator.

Approximation of the optimal compensator is based on approximating the infinite dimensional estimator with the sequence of finite dimensional estimators defined in Section 8.1. The gains for the approximating estimators are given in terms of the solutions to finite dimensional Riccati equations that approximate the infinite dimensional Riccati equation in Section 7. Although

defined as observers, these finite dimensional estimators can be interpreted as Kalman filters, as shown in Section 8.2. In Section 8.3, we give formulas for finite dimensional functional estimator gains that approximate the functional estimator gains of Section 7. These approximating estimator gains indicate how closely the finite dimensional estimators approximate the infinite dimensional estimator. In Section 8.4, we apply the Riccati equation approximation theory of Section 5 to describe the convergence of the finite dimensional estimators.

Most of the results in Section 8 are analogous to results for the control problem and follow from the same basic approximation theory, but certain differences require careful analysis. There is an important difference in the way that the Riccati matrices to be computed are defined in terms of the finite dimensional Riccati operators. Indeed, the Riccati matrix equations to be solved numerically might seem incorrect at first. To demonstrate that the finite dimensional estimators that we define in Section 8.1 are natural approximations to the optimal infinite dimensional estimator, we show in Section 8.2 that each finite dimensional estimator is a Kalman filter for the corresponding finite element approximation of the flexible structure. The brief discussion in Section 8.2 is the only place in the paper where stochastic estimation theory is necessary, and none of the analysis in the rest of the paper depends on this discussion.

In Section 9.1, we apply the nth control law of Section 4 to the output of the nth estimator to form the nth compensator. (The order of approximation is n.) The nth closed-loop system consists of the distributed model of the structure controlled by the nth compensator. Since each finite dimensional

estimator is realizable, the nth compensator and the nth closed-loop system are realizable. In Section 9.2, we discuss how the sequence of realizable closed-loop systems approximates the optimal closed-loop system. Probably the most important question here is whether exponential stability of the optimal closed-loop system implies exponential stability of the nth closed-loop system for n sufficiently large. We have been able to prove this only when the approximation basis vectors are the natural modes of undamped free vibration and these modes are not coupled by structural damping. That this stability result can be generalized is suggested by the results in Section 9.3, which describe how the transfer functions of the finite dimensional compensators approximate the transfer function of the optimal compensator.

In Section 10, we complete the compensator design for the example in Section 6. Assuming that white noise corrupts the single measurement and that distributed white noise disturbs the structure, we compute the gains for the finite dimensional estimators and show the functional estimator gains. As in the control problem, the functional gains do not converge when no damping is modeled. We apply the control laws computed in Section 6 to the output of the estimators in Section 10 to construct the finite dimensional compensators, and we show the frequency response of these compensators. As predicted by Section 9.3, the frequency response of the nth compensator converges to the frequency response of the optimal infinite dimensional compensator as n increases. In Section 10.3, we discuss the structure and dimension of the finite dimensional compensator that should be implemented.

We conclude in Section 11 by discussing where the approximation theory presented in this paper is most complete and what further results would be most important.

## 2. The Control System

We consider the system

$$\ddot{x}(t) + D_0\dot{x}(t) + A_0x(t) = B_0u(t), \quad t > 0, \tag{2.1}$$

where $x(t)$ is in a real Hilbert space $H$ and $u(t)$ is in $R^m$ for some finite $m$. The linear stiffness operator $A_0$ is densely defined and selfadjoint with compact resolvent and at most a finite number of negative eigenvalues. We will postpone discussion of the damping operator $D_0$ momentarily, except to say that it is symmetric and nonnegative. The input operator $B_0$ is a linear operator from $R^m$ to $H$, hence bounded.

By natural modes, we will mean the eigenvectors $\phi_j$ of the eigenvalue problem

$$\lambda_j\phi_j = A_0\phi_j . \tag{2.2}$$

From our hypotheses on $A_0$, we know that these eigenvalues form an infinitely increasing sequence of real numbers, of which all but a finite number are positive. Also, the corresponding eigenvectors are complete in $H$ and satisfy

$$\langle\phi_i, \phi_j\rangle_H = \langle A_0\phi_i, \phi_j\rangle_H = 0, \quad i \neq j. \tag{2.3}$$

(These properties of the eigenvalue problem (2.2) are standard. See, for example, [B1], [K1].) For $\lambda_j > 0$, $\omega_j = \sqrt{\lambda_j}$ is a natural frequency.

Remark 2.1. Our analysis includes the system

$$M_0\ddot{x}(t) + D_0\dot{x}(t) + A_0x(t) = B_0u(t), \quad t > 0, \tag{2.1'}$$

where the mass operator $M_0$ is a selfadjoint, bounded and coercive linear operator on a real Hilbert space $H_0$. The operators $A_0$, $B_0$ and $D_o$ in (2.1') have the same properties with respect to $H_0$ that the corresponding operators in (2.1) have with respect to H. To include (2.1') in our analysis, we need only take H to be $H_0$ with the norm-equivalent inner product $\langle \cdot, \cdot \rangle_H = \langle M_0 \cdot, \cdot \rangle_{H_0}$, and multiply (2.1') on the left by $M_0^{-1}$. In H, the operator $M_0^{-1}A_0$ is selfadjoint with compact resolvent, and $M_0^{-1}D_0$ is symmetric and nonnegative. With no loss of generality then, we will refer henceforth only to (2.1) and assume that the H-inner product accounts for the the mass distribution. □

## 2.1 The Energy Spaces and the First-order Form of the System

### The Elastic-Strain-Energy Space V and Total-Energy Space E

We choose a bounded, selfadjoint linear operator $A_1$ on H such that $\tilde{A}_0 = A_0 + A_1$ is coercive; i.e., there exists $\rho > 0$ for which

$$\langle \tilde{A}_0x, x \rangle_H \geq \rho \|x\|_H^2, \qquad x \in D(\tilde{A}_0) = D(A_0). \tag{2.4}$$

Since $A_0$ is bounded from below, there will be infinitely many such $A_1$'s. In applications like our example in Section 6, it is natural to select for $A_1$ an operator whose null space is the orthogonal complement (in H) of the eigenspace of $A_0$ corresponding to nonpositive eigenvalues. Obviously, any $A_1$ that makes $\tilde{A}_0$ coercive must be positive definite on the nonpositive eigenspace of $A_0$.

With $A_1$ chosen, we define the Hilbert space V to be the completion of $D(A_0)$ with respect to the inner product $\langle v_1, v_2 \rangle_V = \langle \tilde{A}_0 v_1, v_2 \rangle_H$, $v_1$ and $v_2 \in D(A_0)$. Note that $V = D(\tilde{A}_0^{1/2})$ and $\langle v_1, v_2 \rangle_V = \langle \tilde{A}_0^{1/2} v_1, \tilde{A}_0^{1/2} v_2 \rangle_H$. (Since $A_1$ is a bounded operator on H, different choices of $A_1$ yield V's with equivalent norms, thus containing the same elements).

In the usual way, we will use the imbedding

$$V \subset H = H' \subset V',$$

where the injections from V into H and from H into V' are continuous with dense ranges. We denote by $\Lambda_V$ the Riesz map from V onto its dual V'; i.e.,

$$\langle v, v_1 \rangle_V = (\Lambda_V v_1) v, \qquad v_1, v \in V. \tag{2.5}$$

Then $\tilde{A}_0$ is the restriction of $\Lambda_V$ to $D(A_0)$ in the sense that

$$(\Lambda_V v_1) v = \langle v, \tilde{A}_0 v_1 \rangle_H, \quad v_1 \in D(A_0), \quad v \in V. \tag{2.6}$$

Now we define the <u>total energy space</u> $E = V \times H$, noting that when $A_0$ is coercive and x(t) is the solution to (2.1), then $\|(x(t), \dot{x}(t))\|_E^2$ is twice the total energy (kinetic plus potential) in the system. We want to write (2.1) as a first-order evolution equation on E. To to this, we must determine the appropriate semigroup generator for the open-loop system. We will derive this generator by constructing its inverse explicitly, and then we will try to convince the reader that we do have the appropriate open-loop semigroup generator. The approach seems mathematically efficient, and we will need the inverse of the generator for the approximation scheme. First, we must state our precise hypotheses on damping and discuss its representation.

The Damping Functional and Operator

Actually, we do not require an operator $D_0$ defined from some subset of H into H. Rather, we assume only that there exists a damping functional

$$d_0(v_1,v_2): \quad V \times V \to R \tag{2.7}$$

such that $d_0$ is bilinear, symmetric, continuous (on $V \times V$) and nonnegative.

If we have a symmetric, nonnegative damping operator $D_0$ defined on $D(A_0)$ such that $D_0$ is bounded relative to $A_0$, then $\langle D_0 v_1, v_2 \rangle_H$ defines a bilinear, symmetric, bounded, nonnegative functional on a dense subset of $V \times V$. In this case, the unique extension of this functional to $V \times V$ is $d_0$. (That $D_0$ being $A_0$-bounded implies continuity of $\langle D_0 \cdot, \cdot \rangle_H$ with respect to the V-norm follows from [K1, Theorem 4.12, p. 292].)

Under our hypotheses on $d_0$, there is a unique bounded linear operator $\Lambda_D$ from V into $V'$ such that

$$d_0(v,v_1) = (\Lambda_D v_1)v, \quad v_1, v \in V. \tag{2.8}$$

The operator $(\Lambda_V^{-1}\Lambda_D)$ is then a bounded linear operator from V to V, and $(\Lambda_V^{-1}\Lambda_D)$ is selfadjoint (on V) because $d_0$ is symmetric. Also

$$d_0(v,v_1) = \langle v, \Lambda_V^{-1}\Lambda_D v_1 \rangle_V = \langle \Lambda_V^{-1}\Lambda_D v, v_1 \rangle_V, \quad v_1, v \in V. \tag{2.9}$$

Remark 2.2. We chose to begin our description of the control system model with (2.1) because its form is familiar in the context of flexible structures. The stiffness operator $A_0$, for example, is the infinite dimensional analogue of the stiffness matrix in finite dimensional structural analysis. In applications like the example in Section 6, though, it is often easier to begin

with a strain-energy functional from which the correct strain-energy inner product for V is obvious. The stiffness operator is defined then in terms of the Riesz map for V (see [S3] for this approach), rather than V being defined in terms of the stiffness operator; specifically, $\tilde{A}_0$ is defined by (2.6) with $D(\tilde{A}_0) = \Lambda_V^{-1}H$. Either way, the relationship between $A_0$ and V is the same. But the only thing that needs to be computed in applications is the V-inner product; an explicit $A_0$ need not be written down. □

The Semigroup Generator

We define $\tilde{A}^{-1} \in L(E,E)$ by

$$\tilde{A}^{-1} = \begin{bmatrix} -\Lambda_V^{-1}\Lambda_D & -\tilde{A}_0^{-1} \\ I & 0 \end{bmatrix}. \tag{2.10}$$

This operator is clearly one-to-one, and its range is dense, since V is dense in H and $D(A_0)$ is dense in V. Now, we take

$$\tilde{A} = (\tilde{A}^{-1})^{-1}. \tag{2.11}$$

Direct calculation of the inner product shows

$$\langle \tilde{A}^{-1} \begin{matrix} v \\ h \end{matrix}, \begin{matrix} v \\ h \end{matrix} \rangle_E = -d_0(h,h), \tag{2.12}$$

so that $\tilde{A}$ is dissipative with dense domain. Also, since $D(\tilde{A}^{-1}) = E$, $\tilde{A}$ is maximal dissipative by [G1, Theorem 2.1]. Therefore, $\tilde{A}$ generates a $C_0$-contraction semigroup on E.

Finally, the open-loop semigroup generator is

$$A = \tilde{A} + \begin{bmatrix} 0 & 0 \\ A_1 & 0 \end{bmatrix} \qquad D(A) = D(\tilde{A}), \tag{2.13}$$

where $A_1$ is the bounded linear operator discussed above. With

$$B = \begin{bmatrix} 0 \\ B_0 \end{bmatrix} \in L(R^m, E), \tag{2.14}$$

the first-order form of (2.1) is

$$\dot{z}(t) = Az(t) + Bu(t), \qquad t>0 \tag{2.15}$$

where $z = (x, \dot{x}) \in E$.

To see that A is indeed the appropriate open-loop semigroup generator, suppose that $A_0$ is coercive (so that $A_1 = 0$) and that we have a symmetric, nonnegative $A_0$-bounded damping operator $D_0$. Then the appropriate generator should be a maximal dissipative extension of the operator

$$\overset{\circ}{A} = \begin{bmatrix} 0 & I \\ -A_0 & -D_0 \end{bmatrix}, \qquad D(\overset{\circ}{A}) = D(A_0) \times D(A_0). \tag{2.16}$$

It is shown in [G1, Section 2] that $\overset{\circ}{A}$ has a unique maximal dissipative extension, and it can be shown easily that the A defined above is an extension of $\overset{\circ}{A}$ after noting that, in the present case,

$$(\Lambda_V^{-1}\Lambda_D)|_{D(A_0)} = A_0^{-1}D_0. \tag{2.17}$$

We should note that Showalter [S3, Chapter VI] elegantly derives a semigroup generator for a class of second-order systems that includes the flexible-structure model here. The presentation here is most useful for our approximation theory because of the explicit construction of the inverse of

the semigroup generator. For the purposes of this paper, we do not need to characterize the operator A itself more explicitly, but we should make the following points.

First, from $\tilde{A}^{-1}$ we see

$$D(A)=\{(x,\dot{x}):\ \dot{x} \in V,\ x + \Lambda_V^{-1}\Lambda_D\dot{x} \in D(A_0)\}. \tag{2.18}$$

In many applications, especially those involving beams, the "natural boundary conditions" can be determined from (2.18) and the boundary conditions included in the definition of $DA_0$). In the case of a damping operator that is bounded relative to $\tilde{A}_0^{1/2}$, $D(A) = D(A_0) \times V$. If the damping operator is bounded relative to $A_0^{\mu}$ for $\mu < 1$, then A has compact resolvent.

In many structural applications, the open-loop semigroup is analytic, although this has been proved only for certain important cases. Showalter obtains an analytic semigroup when the damping functional is V-coercive; for example, when there exists a damping operator $D_0$ that is both $A_0$-bounded and as strong as $A_0$. Such a damping operator results from the Kelvin-Voigt viscoelastic material model. Also, it can be shown that the semigroup is analytic for a damping operator equal to $c_0 A_0^{\mu}$ for $1/2 \leq \mu \leq 1$ and $c_0$ a positive scalar. The case $\mu = 1/2$, which produces the same damping ratio in all modes, is especially common in structural models, and Chen and Russell [C1] have shown that the semigroup is analytic for a more general class of damping operators involving $A_0^{1/2}$.

Finally, we can guarantee that the open-loop semigroup generator is a spectral operator (i.e., its eigenvectors are complete in E) only for a damping operator that is a linear combination of an H-bounded operator and a

fractional power of $A_0$. However, nowhere do we use or assume anything about the eigenvectors of either the open-loop or the closed-loop semigroup generator. The natural modes -- of undamped free vibration -- in (2.2) are always complete in both H and V.

## 2.2 The Adjoint of $\tilde{A}$

Since $(\Lambda_V^{-1}\Lambda_D)$ is selfadjoint on V, direct calculation shows that $\tilde{A}^{-*} = (\tilde{A}^{-1})^*$ -- the adjoint of $\tilde{A}^{-1}$ with respect to the E-inner product -- is

$$\tilde{A}^{-*} = \begin{bmatrix} -\Lambda_V^{-1}\Lambda_D & \tilde{A}_0^{-1} \\ -I & 0 \end{bmatrix} . \tag{2.19}$$

Then $\tilde{A}^* = (\tilde{A}^{-*})^{-1}$. Having $\tilde{A}^{-*}$ explicitly facilitates proving strong convergence for approximating adjoint semigroups.

## 2.3 Exponential Stability

The following theorem says that, if there are no rigid-body modes and if the damping is coercive (basically, all structural components have positive damping), then the open-loop system is uniformly exponentially stable. That the decay rate given depends only on the lower bound for the stiffness operator and the upper and lower bounds for the damping functional is essential for convergence results for the approximating optimal control problems of subsequent sections. The theorem is a generalization of Theorem 6.1 of [G1] to allow more general damping, but the proof is entirely different and much nicer. The current proof uses an explicit Lyapunov functional for the homogeneous part of the system in (2.15). Recall that $T(\cdot)$ is the open-loop semigroup, with generator A, and E is the total energy space $V \times H$.

**Theorem 2.3.** Suppose that $A_0$ and $d_0$ are H-coercive. Let $\rho$ be the positive constant in (2.4), and let $\delta_0$ and $\delta_1$ be positive constants such that

$$\delta_0 ||v||_H^2 \leq d_0(v,v) \leq \delta_1 ||v||_V^2, \qquad v \in V. \tag{2.20}$$

Then

$$||T(t)|| \leq (1 + \frac{\delta_0}{\sqrt{\rho}} + \delta_1\delta_0/2)^{1/2} \exp[-t/(\frac{2}{\delta_0} + \frac{2}{\sqrt{\rho}} + \delta_1)] \quad , t \geq 0. \tag{2.21}$$

**Proof.** For $\gamma > \max\{\frac{1}{\sqrt{\rho}}, \frac{2}{\delta_0}\}$, define $Q \in L(E)$ as

$$Q = \begin{bmatrix} (\gamma I + \Lambda_V^{-1}\Lambda_D) & A_0^{-1} \\ I & \gamma I \end{bmatrix}. \tag{2.22}$$

Since $(\Lambda_V^{-1}\Lambda_D)$ is selfadjoint and nonnegative on V, Q is selfadjoint and coercive on E. Define the functional $p(\cdot)$ on E by

$$p(z) = \langle Qz, z\rangle_E = \gamma(||x||_V^2 + ||\dot{x}||_H^2) + 2\langle \dot{x}, x\rangle_H + d_0(x,x), \tag{2.23}$$

where $z = (x, \dot{x})$. From (2.4), we have

$$2|\langle x, \dot{x}\rangle_H| \leq \frac{2}{\sqrt{\rho}} ||x||_V ||\dot{x}||_H \leq \frac{1}{\sqrt{\rho}} ||z||_E^2, \tag{2.24}$$

so that

$$(\gamma - \frac{1}{\sqrt{\rho}}) ||z||_E^2 \leq p(z) \leq \beta^{-1} ||z||_E^2 \tag{2.25}$$

with

$$\beta = (\gamma + \frac{1}{\sqrt{\rho}} + \delta_1)^{-1}. \tag{2.26}$$

Next take $z = (x,\dot{x}) \in D(A)$ and set $A(x,\dot{x}) = (y,\dot{y}) \in E$, or

$$(x,\dot{x}) = A^{-1}(y,\dot{y}) = (-\Lambda_V^{-1}\Lambda_D y - A_0^{-1}\dot{y}, y) . \tag{2.27}$$

Note $\dot{x} = y \in V$. Now,

$$\langle QAz,z\rangle_E = \langle Q(y,\dot{y}), A^{-1}(y,\dot{y})\rangle_E =$$

$$- \gamma\langle y,\Lambda_V^{-1}\Lambda_D y\rangle_V - \|\Lambda_V^{-1}\Lambda_D y + A_0^{-1}\dot{y}\|_V^2 + \|y\|_H^2$$

$$= - [\|x\|_V^2 + \gamma d_0(\dot{x},\dot{x}) - \|\dot{x}\|_H^2]. \tag{2.28}$$

From (2.20) then,

$$\langle QAz,z\rangle_E \leq - [\|x\|_V^2 + (\gamma\delta_0 - 1)\|\dot{x}\|_H^2] \leq -\|z\|_E^2. \tag{2.29}$$

Therefore $p(\cdot)$ is a Lyapunov functional, and the theorem follows from (2.25), (2.26) and (2.29), with $\gamma = \frac{1}{\sqrt{p}} + \frac{2}{\delta_0}$ . □

## 3. The Optimal Control Problem

Subsection 3.1 presents some preliminary definitions and results for the optimal linear-quadratic regulator problem on an arbitrary real Hilbert space. These results are generic in the sense that the Hilbert space E is not necessarily the energy space of Section 2, and the operators A, B, etc., do not necessarily represent an abstract flexible structure as in Section 2. In the second half of the paper, having such generic results will allow us to obtain the approximation theory for the infinite dimensional state estimator from the analogous results for the control problem. Subsection 3.2 gives some important implications of the general results for the case where the control system is that defined in Section 2.

### 3.1 The Generic Optimal Regulator Problem

Let a linear operator A generate a $C_0$-semigroup T(t) on a real Hilbert space E, and suppose $B \in L(R^m,E)$, $Q \in L(E,E)$ and $R \in L(R^m)$, with Q nonnegative and selfadjoint and R positive definite and symmetric. The optimal control problem on E is to choose the control $u \in L_2(0,\infty;R^m)$ to minimize the cost functional

$$J(z(0),u) = \int_0^\infty (\langle Qz(t),z(t)\rangle_E + \langle Ru(t),u(t)\rangle_{R^m})dt, \tag{3.1}$$

where the state z(t) is given by

$$z(t) = T(t)z(0) + \int_0^t T(t-\eta)Bu(\eta)d\eta, \quad t \geq 0. \tag{3.2}$$

Definition 3.1. A function $u \in L_2(0,\infty;U)$ is an admissible control for the

initial state z, or simply an admissible control for z, if J(z,u) is finite; i.e., if the state z(t) corresponding to the control u(t) and the initial condition z(0) = z is in $L_2(0,\infty;E)$.

Definition 3.2 Let the operators A, B, Q, and R be as defined above. An operator $\Pi$ in L(E) is a solution of the Riccati algebraic equation if $\Pi$ maps the domain of A into the domain of $A^*$ and satisfies the Riccati algebraic equation

$$A^*\Pi + \Pi A - \Pi BR^{-1}B^*\Pi + Q = 0. \tag{3.3}$$

Theorem 3.3 (Theorems 4.6 and 4.11 of [G4]). There exists a nonnegative selfadjoint solution of the Riccati algebraic equation if and only if, for each $z \in E$, there is an admissible control for the initial state z. If $\Pi$ is the minimal nonnegative selfadjoint solution of (3.3), then the unique control u( ) which minimizes J(z,u) and the corresponding optimal trajectory z( ) are given by

$$u(t) = -R^{-1}B^*\Pi z(t) \tag{3.4}$$

and

$$z(t) = S(t)z, \tag{3.5}$$

where S(·) is the semigroup generated by $A-BR^{-1}B^*\Pi$. Also,

$$J(z,u) = \min_v J(z,v) = \langle \Pi z, z\rangle_E \tag{3.6}$$

If, for each initial state and admissible control.

$$\lim_{t\to\infty} ||z(t)|| = 0, \tag{3.7}$$

there exists at most one nonnegative selfadjoint solution of (3.3). If Q is coercive, (3.7) holds for each initial state and admissible control and $S(\cdot)$ is uniformly exponentially stable. □

We will refer to $T(\cdot)$ as the open-loop semigroup and to $S(\cdot)$ as the optimal closed-loop semigroup.

To prepare for the convergence analysis in Sections 5 and 9, we must present now some rather arcane estimates for the decay rate of the closed-loop system in the optimal control problem.

Theorem 3.4 Suppose that the open-loop semigroup $T(\cdot)$ satisfies

$$||T(t)|| \leq M_1 e^{\alpha_1 t}, \quad t \geq 0, \tag{3.8}$$

for positive constants $M_1$ and $\alpha_1$, that $\Pi$ is the minimal nonnegative selfadjoint solution to (3.3), and that $S(t)$ is the optimal closed-loop semigroup in Theorem 3.3. If there exists a constant $M_0$ such that, for each $z \in E$,

$$\int_0^\infty ||S(t)z||^2 dt \leq M_0(\langle \Pi z, z\rangle_E + ||z||^2) \tag{3.9}$$

and a constant $M_0'$ such that

$$||\Pi|| \leq M_0', \tag{3.10}$$

then there exist positive constants $M_2$ and $\alpha_2$, which are functions of $M_0$, $M_0'$, $M_1$, and $\alpha_1$ only, such that

$$||S(t)|| \leq M_2 e^{-a_2 t}, \quad t \geq 0. \tag{3.11}$$

Proof. This follows easily from Theorems 2.2 and 4.7 of [ ]. □

Lemma 3.5. Suppose that there exist positive constants M and $\alpha$ such that

$$||T(t)|| \leq Me^{-\alpha t}, \quad t \geq 0. \tag{3.12}$$

If $z(0) \in E$, $h \in L_2(0,\infty;E)$, and $z(t) = T(t)z(0) + \int_0^t T(t-s)h(s)ds$, then

$$\int_0^\infty ||z(t)||^2 \, dt \leq \left[\frac{M}{\sqrt{2\alpha}}||z(0)|| + \frac{M}{\alpha}||h||_{L_2}\right]^2. \tag{3.13}$$

Proof. The result follows from (3.12), the convolution theorem [D1, page 951] and the triangle inequality. □

Lemma 3.6. Suppose that E is finite dimensional and that the pair (Q,A) is observable (in the usual finite dimensional sense). Then there exists a constant M, which is a function of A, B and Q only, such that

$$\int_0^\infty ||z(t)||^2 dt \leq M\left(\int_0^\infty (\langle Qz(t),z(t)\rangle_E + ||u(t)||^2)dt, \right. \tag{3.14}$$

where z(t) is given by (3.2).

Proof. The proof, which is at most a mild challenge, is based on the fact that the observability grammian

$$W(t) = \int_0^t e^{A^T t} Q e^{At} dt \tag{3.15}$$

is coercive for any positive t. □

The next theorem says, among other things, that if the open-loop control system decouples into a finite dimensional part that is stabilizable (in the usual finite dimensional sense) and an infinite dimensional part that is uniformly exponentially stable, then the entire system is uniformly exponentially stabilizable, so that (3.3) has a nonnegative selfadjoint solution.

Theorem 3.7. Suppose that there exists a finite dimensional subspace $E_0 \subset D(A)$ such that $E_0$ and $E_0^{\perp}$ reduce A (and T(t)), and write

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}, \quad Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^* & Q_{22}^* \end{bmatrix}, \tag{3.16}$$

where $A_{11}$ and $A_{22}$ are the restrictions of A to $E_0$ and $D(A) \times E_0^{\perp}$, respectively. Similarly,

$$T(t) = \begin{bmatrix} T_{11}(t) & 0 \\ 0 & T_{22}(t) \end{bmatrix}. \tag{3.17}$$

Also, suppose that the pair $(A_{11}, B_{11})$ is stabilizable and that there exist positive constants $M_1'$, $\alpha_1'$ and $\beta$ such that

$$||T_{22}(t)|| \leq M_1' e^{-\alpha_1' t}, \quad t \geq 0 \tag{3.18}$$

and

$$\max\{||B||, ||Q||\} \leq \beta. \tag{3.19}$$

i) Then there exists $F \in L(E, R^m)$ such that A-BF generates a uniformly exponentially stable semigroup on E. Also, (3.3) has a nonnegative selfadjoint solution, and the minimal such solution satisfies (3.10) with $M_0'$ a function of $A_{11}$, $B_{11}$, R, $M_1'$, $\alpha_1'$ and $\beta$ only.

ii) If $Q_{12} = 0$ and the pair $(Q_{11}, A_{11})$ is observable, then there exists a unique nonnegative selfadjoint solution $\Pi$ to (3.3), and there exist positive constants $M_2$ and $\alpha_2$ -- which depend on $A_{11}$, $B_{11}$, $Q_{11}$, R, $M_1'$, $\alpha_1'$ and $\beta$ only -- such that the optimal closed-loop semigroup satisfies

$$||S(t)|| \leq M_2 e^{-\alpha_2 t}, \quad t \geq 0. \tag{3.20}$$

Proof. i) To say that $(A_{11}, B_{11})$ is stabilizable means that there exists a linear operator $F_{11}$ from $E_0$ to $R^m$ such that each eigenvalue of $A_{11}-B_{11}F_{11}$ has negative real part. Hence A-BF generates a uniformly exponentially stable semigroup if $F = [F_{11}\ 0]$, so that there exists an admissible control for each initial condition in E.

It is easy to write down an upper bound for the performance index in (3.1) in terms of R, $\beta$, $M_1'$, $\alpha_1'$ and the decay rate of $\exp([A_{11} - B_{11} F_{11}]t)$. That the $M_0'$ in (3.10) depends only on $A_{11}$, $B_{11}$, R, $M_1'$, $\alpha_1'$ and $\beta$ follows then from the fact that $F_{11}$ is a function of $A_{11}$ and $B_{11}$.

ii) Clearly, (3.8) holds with $M_1$ and $\alpha_1$ depending only on $A_{11}$, $B_{11}$, $M_1'$, $\alpha_1'$ and $F_{11}$. Therefore, we have (3.8) and (3.10) with the bounds depending only on $A_{11}$, $B_{11}$, $Q_{11}$, R, $M_1'$, $\alpha_1'$ and $\beta$. Finally, the existence of an $M_0$ for

(3.9) which depends only on these parameters follows from using (3.1) and (3.6) in applying Lemma 3.6 to the part of the system on $E_0$ and then Lemma 3.5 to the part of the system on $E_0^{\perp}$. Part ii) of this theorem then follows from Theorem 3.4. □

Remark 3.8 When we say in Theorem 3.4 that $M_2$ and $\alpha_2$ are functions of $M_0$, $M_0'$, $M_1$, and $\alpha_1$ only, we mean, for example, that for two optimal control problems on different spaces E, with different operators A, B, etc., if the same constants $M_0$, $M_0'$, $M_1$, and $\alpha_1$ work in (3.8)-(3.10) for both problems, then the same constants $M_2$ and $\alpha_2$ will work in (3.11) for both problems. Similarly, in Theorem 3.7 ii), as long as $E_0$, $A_{11}$, $B_{11}$, $Q_{11}$, R, $M_1'$, $\alpha_1'$ and $\beta$ remain the same, the same $M_2$ and $\alpha_2$ will work in (3.20) even if $E_0^{\perp}$, $A_{22}$, $B_{21}$ and $Q_{22}$ change. □

## 3.2 Application to Optimal Control of Flexible Structures

For the rest of this section, $A_0$, $A_1$, A, T(t), $B_0$ and B are the operators defined in Section 2.1, and $E = V \times H$ is the energy space defined there.

Remark 3.9. Theorem 3.7 is useful mainly when all but a finite number of modes have coercive damping in the open-loop system and the damped and undamped parts of the open-loop system remain orthogonal. This is the case, for example, with modal damping. The next theorem does not require orthogonality of the damped and undamped parts of the system, but it does require an independent actuator for each undamped mode. The situation of Theorem 3.10 is typical in aerospace structures: Any elastic component should have some structural damping, but rigid-body modes are common; for a structure to be controllable, an actuator is required for each rigid-body mode. □

Theorem 3.10. i) Suppose that $A_1 = B_0 B_0^*$ and that $\tilde{A}_0 = A_0 + A_1$ and $\tilde{d}_0 = d_0 + A_1$ are H-coercive, so that there exist positive constants $\rho$, $\gamma$ and $\beta$ such that, for all $v \in V$,

$$||v||_V^2 \geq \rho ||v||_H^2, \tag{3.21}$$

$$\tilde{d}_o(v,v) \geq \rho ||v||_H^2, \tag{3.22}$$

$$\tilde{d}_o(v,v) \leq \gamma ||v||_V^2, \tag{3.23}$$

and

$$\max\{||B_o||, ||Q||, ||R||\} \leq \beta \tag{3.24}$$

(The V-continuity of $d_0$ implies (3.23).) Then (3.3) has a minimal nonnegative selfadjoint solution $\Pi$, which satisfies (3.10) with $M_0'$ a function of $\rho$, $\gamma$ and $\beta$ only.

ii) Suppose also that

$$\langle Qz, z\rangle_E \geq \rho ||z||_E^2, \quad z \in E. \tag{3.25}$$

Then the optimal closed-loop semigroup satisfies

$$||S(t)|| \leq M_2 e^{-\alpha_2 t}, \quad t \geq 0, \tag{3.26}$$

where $M_2$ and $\alpha_2$ are positive constants depending on $\rho$, $\gamma$ and $\beta$ only.

Proof. i) The suboptimal control

$$u(t) = -B_o^*[x(t) + \dot{x}(t)] \tag{3.27}$$

produces a closed-loop system with exponential decay at least as fast as that in Theorem 2.1. The required upper bound in (3.10) follows then from (3.1), (3.6) and (3.24)

ii) In this case, the $M_0$ in (3.9) is $\rho$, and Theorem 3.4 yields the result. □

Now we will consider the structure of the optimal control law in more detail. Since $\Pi \in L(E,E)$ and $E = V \times H$, we can write

$$\Pi = \begin{bmatrix} \Pi_0 & \Pi_1 \\ \Pi_1^* & \Pi_2 \end{bmatrix}, \tag{3.28}$$

where $\Pi_0 \in L(V,V)$, $\Pi_1 \in L(H,V)$, $\Pi_2 \in L(H,H)$, and $\Pi_0$ and $\Pi_2$ are nonnegative and selfadjoint. With $z = (x,\dot{x})$, as in Section 2, (3.4) becomes

$$u(t) = -R^{-1}B_0^*[\Pi_1^* x(t) + \Pi_2 \dot{x}(t)]. \tag{3.29}$$

Since $B_0 \in L(R^m,H)$, we must have vectors $b_i \in H$, $1 \leq i \leq m$, such that

$$B_0 u = \sum_{i=1}^{m} b_i u_i \tag{3.30}$$

for

$$u = [u_1 u_2 \cdots u_m]^T \in R^m. \tag{3.31}$$

Also, for $h \in H$,

$$B_o^* h = [\langle b_1,h\rangle_H \ \langle b_2,h\rangle_H \cdots \langle b_m,h\rangle_H]^T. \tag{3.32}$$

Since $\Pi_1^* x(t)$ and $\Pi_2 \dot{x}(t)$ are elements of H, we see from (3.29) and

(3.32) that the components of the optimal control have the feedback form

$$u_i(t) = -\langle f_i, x(t)\rangle_V - \langle g_i, \dot{x}(t)\rangle_H, \quad i = 1,\ldots,m, \tag{3.33}$$

where $f_i \in V$ and $g_i \in H$ are given by

$$f_i = \sum_{j=1}^{m} (R^{-1})_{ij} \Pi_1 b_j, \tag{3.34a}$$

$$g_i = \sum_{j=1}^{m} (R^{-1})_{ij} \Pi_2 b_j, \quad i = 1,\ldots,m. \tag{3.34b}$$

We call $f_i$ and $g_i$ <u>functional gains</u>.

## 4. The Approximation Scheme

### 4.1 Approximation of the Open-loop System

Hypothesis 4.1. There exists a sequence of finite dimensional subspaces $V_n$ of V such that the sequence of orthogonal projections $P_{Vn}$ converges V-strongly to the identity, where $P_{Vn}$ is the V-projection onto $V_n$. Also, each $V_n$ is the span of n linearly independent vectors $e_j$.

Since it should cause no confusion, we will omit the subscript n and write just $e_j$, keeping in mind that the basis vectors may change from one $V_n$ to another, as in most finite element schemes. Also, we will refer to the Hilbert space $E_n = V_n \times V_n$, which has the same inner product as $E = V \times H$.

For $n \geq 1$, we approximate x(t) by

$$x_n(t) = \sum_{j=1}^{n} \xi_j(t) e_j, \tag{4.1}$$

where $\xi(t) = (\xi_1(t), \xi_2(t), \ldots \xi_n(t))^T$ satisfies

$$M^n \ddot{\xi}(t) + D^n \dot{\xi}(t) + K^n \xi(t) = B_0^n u(t), \tag{4.2}$$

and the mass matrix $M^n$, damping matrix $D^n$, stiffness matrix $K^n$, and actuator influence matrix $B_0^n$ are given by

$$M^n = [\langle e_i, e_j \rangle_H], \qquad D^n = [d_0(e_i, e_j)],$$

$$K^n = [A_0^{1/2} e_i, A_0^{1/2} e_j \rangle_H] = [\langle e_i, e_j \rangle_V] - [\langle A_1 e_i, e_j \rangle_H], \tag{4.3}$$

$$B_0^n = [\langle e_i, b_j \rangle_H].$$

Of course, (4.2) can be written as

$$\dot{\eta} = A^n \eta + B^n u \tag{4.2'}$$

where

$$\eta = [\xi^T, \dot{\xi}^T]^T \tag{4.4}$$

and

$$A^n = \begin{bmatrix} 0 & I \\ -M^{-n}K^n & -M^{-n}D^n \end{bmatrix}, \quad B^n = \begin{bmatrix} 0 \\ M^{-n}B_0^n \end{bmatrix}. \tag{4.5}$$

Note: Throughout this paper, we use the superscript n in the designation of matrices in the $n^{th}$ approximating system and control problem, like $A^n$, $B^n$, $M^n$, etc. Hence the superscript n indicates the order of approximation -- and it never indicates a power of the matrix. By $M^{-n}$, we denote the inverse of the mass matrix $M^n$.

In the designation of a linear operator in the $n^{th}$ approximation, we use the the subscript n. For example, $A_n$ and $B_n$ are the operators whose matrix representations are $A^n$ and $B^n$, respectively.

For convergence analysis, it is useful to note that (4.1) and (4.2) or (4.2') are equivalent to

$$\dot{z}_n(t) = A_n z_n(t) + B_n u(t), \tag{4.6}$$

where $z_n = (x_n, \dot{x}_n) \in E_n$, and $A_n \in L(E_n)$ and $B_n \in L(R^m, E_n)$ are the operators whose matrix representations are given in (4.5). Also, for any real $\lambda$,

$$(\lambda - A_n) \begin{pmatrix} v_n^1 \\ v_n^2 \end{pmatrix} = \begin{pmatrix} h_n^1 \\ h_n^2 \end{pmatrix} \tag{4.7}$$

is equivalent to

$$(\lambda^2 M^n + \lambda D^n + K^n)\alpha^1 = (\lambda M^n + D^n)\beta^1 + M^n\beta^2 \tag{4.8}$$

and

$$\alpha^2 = \lambda\alpha^1 - \beta^1 \tag{4.9}$$

if

$$v_n^j = \sum_{i=1}^{n} \alpha_i^j e_i \quad \text{and} \quad h_n^j = \sum_{i=1}^{n} \beta_i^j e_i, \quad j = 1,2. \tag{4.10}$$

(Substituting $A^n$ and (4.10) into (4.7) yields (4.8) and (4.9)).

Next, we will prepare to invoke the Trotter-Kato semigroup approximation theorem to show how (4.2), (4.2') and (4.6) approximate (2.1) and (2.15). First, we will treat the case in which $A_0$ is coercive (no rigid-body modes), so that $A_1=0$ and $\tilde{A}_0=A_0$; the general case is a straight-forward extension. For $A_0$ coercive, the open-loop semigroup generator A is maximal dissipative. Also, for each n, $A_n$ is dissipative on $E_n$. The main idea here is to project $(\lambda-A)^{-1}$

onto $V_n$ in a certain inner product and observe that the result is exactly $(\lambda - A_n)^{-1}$, where $A_n$ is the operator on $V_n$ in (4.6) and (4.7). Of course, we need only do this for real $\lambda > 0$.

For real $\lambda > 0$ then, define an innerproduct on V by

$$\langle \cdot, \cdot \rangle_\lambda = \lambda^2 \langle \cdot, \cdot \rangle_H + \lambda d_0(\cdot, \cdot) + \langle \cdot, \cdot \rangle_V. \tag{4.11}$$

Under the hypotheses in Section 2 on $d_0$, $\langle \cdot, \cdot \rangle_\lambda$ is norm-equivalent to $\langle \cdot, \cdot \rangle_V$. For $n \geq 1$, let $P_n(\lambda)$ be the projection of V onto $V_n$ in the inner product $\langle \cdot, \cdot \rangle_\lambda$. Now let $h_1, h_2 \in H$ and note that

$$(\lambda - A) \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = \begin{pmatrix} h^1 \\ h^2 \end{pmatrix} \tag{4.12}$$

is equivalent to

$$\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = A^{-1} \left[ \lambda \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} - \begin{pmatrix} h^1 \\ h^2 \end{pmatrix} \right] . \tag{4.13}$$

With $A^{-1}$ from (2.10), (4.13) is equivalent to

$$(I + \lambda \Lambda_V^{-1} \Lambda_D + \lambda^2 A_0^{-1}) v^1 = (\lambda A_0^{-1} + \Lambda_V^{-1} \Lambda_D) h^1 + A_0^{-1} h^2 \tag{4.14}$$

and

$$v_2 = \lambda v^1 - h^1. \tag{4.15}$$

If

$$v_n^1 = P_n(\lambda) v^1 \quad \text{and} \quad v_n^2 = P_n(\lambda) v^2, \tag{4.16}$$

it follows from (4.11) and (4.14) that

$$\langle e_i, v_n^1\rangle_\lambda = \langle e_i, v^1\rangle_\lambda$$

$$= \lambda^2\langle e_i, A_0 A_0^{-1} v^1\rangle_H + \lambda\langle e_i, \Lambda_V^{-1}\Lambda_D v^1\rangle_V + \langle e_i, v^1\rangle_V \qquad (4.17)$$

$$= \langle e_i, (\lambda^2 A_0^{-1} + \lambda\Lambda_V^{-1}\Lambda_D + I)v^1\rangle_V$$

$$= \langle e_i, (\lambda A_0^{-1} + \Lambda_V^{-1}\Lambda_D)h^1 + A_0^{-1}h^2\rangle_V,$$

and from (4.15) that

$$\langle e_i, v_n^2\rangle_\lambda = \langle e_i, v^2\rangle_\lambda = \lambda\langle e_i, v^1\rangle_\lambda - \langle e_i, h^1\rangle_\lambda \ . \qquad (4.18)$$

Now, for $h^1 = h_n^1 \in V_n$, $h^2 = h_n^2 \in V_n$, and $v_n^1$, $v_n^2$, $h_n^1$ and $h_n^2$ written as in (4.10), (4.17) and (4.18) yield (4.8) and (4.9) again.

This shows that

$$\begin{bmatrix} P_n(\lambda) & 0 \\ 0 & P_n(\lambda) \end{bmatrix} (\lambda - A)^{-1}\big|_{E_n} = (\lambda - A_n)^{-1}, \qquad (4.19)$$

which yields

$$\begin{bmatrix} P_n(\lambda) & 0 \\ 0 & P_n(\lambda) \end{bmatrix} (\lambda - A)^{-1} P_{En} = (\lambda - A_n)^{-1} P_{En} \qquad (4.20)$$

where $P_{En}$ is the E-projection of E onto $E_n$. The projection $P_{En}$ can be written

$$P_{En} = \begin{bmatrix} P_{Vn} & 0 \\ 0 & P_{Hn} \end{bmatrix}, \qquad (4.21)$$

where $P_{Vn}$ is the V-projection onto $V_n$, as before, and $P_{Hn}$ is the H-projection onto $V_n$. Since the V-norm is stronger than the H-norm and the norm induced by

the $\lambda$-innerproduct is equivalent to the V-norm, it follows from Hypothesis 4.1 that $(\lambda-A_n)^{-1} P_{En}$ converges E-strongly to $(\lambda-A)^{-1}$ as $n\to\infty$. Now, with $A_n$ extended to $E_n^{\perp}$ as, say, $n(P_{En}-I)$, Trotter-Kato [K1, page 504, Theorem 2.16] yields the following.

Theorem 4.2. For $A_0$ coercive, let $T_n(\cdot)$ be the (contraction) semigroup generated on $E_n$ by $A_n$. Then, for each $t \geq 0$, $T_n(t)P_{En}$ converges strongly to $T(t)$, uniformly in $t$ for $t$ in bounded intervals.

In the general case, when $A_0$ is not coercive, the open-loop generator $A$ is obtained from the dissipative $\tilde{A}$ by the bounded perturbation in (2.13), so that [G3, Theorem 6.6] yields the following generalization of Theorem 4.2.

Corollary 4.3. Let $T_n(\cdot)$ be the semigroup generated on $E_n$ by $A_n$. Then, for each $t \geq 0$, $T_n(t)P_{En}$ converges strongly to $T(t)$, uniformly in $t$ for $t$ in bounded intervals.

Theorem 4.4. When $A$ has compact resolvent, $(\lambda-A_n)^{-1}P_{En}$ converges in $L(E)$ to $(\lambda-A)^{-1}$.

Proof. This follows from (4.20) and a standard result that the projections of a compact linear operator onto a sequence of subspaces converge in norm if the projections converge strongly to the identity, as do $P_{En}$ and $P_n(\lambda)$. □

That the adjoint semigroups also converge strongly follows from an argument entirely analogous to the proof of Theorem 4.2. In particular, equations like (4.11)-(4.17) are used to show that

$$\begin{bmatrix} P_n(\lambda) & 0 \\ 0 & P_n(\lambda) \end{bmatrix} (\lambda - A^*)^{-1} P_{En} = (\lambda - A_n^*)^{-1} P_{En} \tag{4.22}$$

In showing this, $A^{-*}$ is used as $A^{-1}$ was used above. Also, care must be taken to calculate $A_n^*$ with respect to the E-inner product. The result is

Theorem 4.5. Let $T_n(\cdot)$ be the sequence of semigroups in Corollary 4.3. Then, for each $t \geq 0$, $T_n^*(t)P_{En}$ converges strongly to $T^*(t)$, uniformly in t for t in bounded intervals.

Finally, for the approximation to the actuator influence operator $B \in L(R^m, E)$, recall $B_n \in L(R^m, E_n)$, the operator whose matrix representation is the matrix $B^n$ in (4.5). From (4.3), it follows that

$$B_n = P_{En}B. \tag{4.23}$$

Since B has finite rank m, $B_n$ and $B_n^*$ converge in norm to B and $B^*$, respectively.

## 4.2 The Approximating Optimal Control Problems

The $n^{th}$ optimal control problem is: given $z_n(0) = (x_n(0), \dot{x}_n(0)) \in E_n$, choose $u \in L_2(0,\infty;R^m)$ to minimize

$$J_n(z_n(0),u) = \int_0^\infty (\langle Q_n z_n(t), z_n(t)\rangle_E + \langle Ru(t), u(t)\rangle_{R^m})dt, \tag{4.24}$$

where $Q_n = P_{En}Q|_{En}$. We assume:

Hypothesis 4.6. For each $n \geq 1$ and $z_n(0) \in E_n$, there exists an admissible control (Definition 3.1) for (4.6) and (4.24). □

A sufficient condition for Hypothesis 4.6 is that, for each $n$, the system $(A_n, B_n)$ be stabilizable.

By Theorem 3.1, the optimal control $u_n(t)$ has the feedback form

$$u_n(t) = -R^{-1}B_n^* \Pi_n z_n(t) \tag{4.25}$$

where $\Pi_n$ is a linear operator on $E_n$, $\Pi_n$ is nonnegative and selfadjoint, and $\Pi_n$ satisfies the Riccati equation

$$A_n^* \Pi_n + \Pi_n A_n - \Pi_n B_n R^{-1} B_n^* \Pi_n + Q_n = 0. \tag{4.26}$$

As a result of Hypothesis 4.6, (4.26) has at least one nonnegative, selfadjoint solution. The minimal such solution is the correct $\Pi_n$ here. If the system $(A_n, Q_n)$ is observable, then $\Pi_n$ is the unique nonnegative, selfadjoint solution to (4.26), and is positive definite. If we write $\Pi_n$ as

$$\Pi_n = \begin{bmatrix} \Pi_{0n} & \Pi_{1n} \\ \Pi_{1n}^* & \Pi_{2n} \end{bmatrix}, \tag{4.27}$$

then (4.25) becomes

$$u_n(t) = -R^{-1}B_0^*[\Pi_{1n}^* x_n(t) + \Pi_{2n}\dot{x}_n(t)]. \tag{4.28}$$

The feedback law (4.28) can be written in functional-feedback form, just as in Section 3. We have

$$u_n(t) = [u_{1n}(t) \quad u_{2n}(t) \ldots u_{mn}(t)]^T, \tag{4.29}$$

where

$$u_{in}(t) = -\langle f_{in}, x_n(t)\rangle_V - \langle g_{in}, \dot{x}_n(t)\rangle_H, \quad 1 \leq i \leq m, \tag{4.30}$$

and

$$f_{in} = \sum_{j=1}^{m} (R^{-1})_{ij} \Pi_{1n} P_{Hn} b_j, \qquad 1 \leq i \leq m, \tag{4.31a}$$

$$g_{in} = \sum_{j-1}^{m} (R^{-1})_{ij} \Pi_{2n} P_{Hn} b_j, \qquad 1 \leq i \leq m. \tag{4.31b}$$

Of course, $f_{in}$ and $g_{in}$ are the $n^{th}$ approximations to the functional gains $f_i$ and $g_i$ in (3.34).

In Section 5, (4.25)-(4.31) will be useful for studying how the solution to the $n^{th}$ optimal control problem converges to the solution to the original problem of Section 3, but for numerical solution of the $n^{th}$ problem, we need the matrix representations of these equations.

We will need the following grammian matrices:

$$\tilde{K}^n = [\langle e_i, e_j \rangle_V] = K^n + [\langle A_1 e_i, e_j \rangle_H] \tag{4.32}$$

and

$$W^n = \begin{bmatrix} \tilde{K}^n & 0 \\ 0 & M^n \end{bmatrix}. \tag{4.33}$$

Note: The matrix $W^{-n}$ will be the inverse of $W^n$. The superscript n on any matrix indicates the order of approximation, not a power of the matrix. Also, recall the note following (4.5).

Now recall $Q_n = P_{En} Q|_{En}$. Since $Q = Q^* \in L(E)$ and $E = V \times H$, we can write

$$Q = \begin{bmatrix} Q_0 & Q_1 \\ Q_1^* & Q_2 \end{bmatrix} \tag{4.34}$$

where $Q_0 = Q_0^* \in L(V)$, $Q_1 \in L(H,V)$, and $Q_2 = Q_2^* \in L(H)$. Straightforward calculation shows that

$$Q^n = W^{-n} \tilde{Q}^n, \tag{4.35}$$

where $Q^n$ is the matrix representation of $Q_n$ and $\tilde{Q}^n$ is the nonnegative, symmetric matrix

$$\tilde{Q}^n = \begin{bmatrix} \tilde{Q}_0^n & \tilde{Q}_1^n \\ \tilde{Q}_1^{nT} & \tilde{Q}_2^n \end{bmatrix} \tag{4.36}$$

with

$$\tilde{Q}_0^n = [\langle e_i, Q_0 e_j \rangle_V],$$

$$\tilde{Q}_1^n = [\langle e_i, Q_1 e_j \rangle_V], \qquad (4.37)$$

$$\tilde{Q}_2^n = [\langle e_i, Q_2 e_j \rangle_H].$$

Also, recall that $A_n$ and $B_n$ are the operators whose matrix representations are given by (4.5), and note that the matrix representations of $A_n^*$ and $B_n^*$ are $W^{-n}(A^n)^T W^n$ and $(B^n)^T W^n$, respectively.

With the matrix representation of $\Pi_n$ denoted by $\Pi^n$, the Riccati operator equation (4.26) is equivalent to the Riccati matrix equation

$$W^{-n}(A^n)^T W^n \Pi^n + \Pi^n A^n - \Pi^n B^n R^{-1} (B^n)^T W^n \Pi^n + Q^n = 0. \qquad (4.38)$$

While $\Pi_n$ is selfadjoint, $\Pi^n$ in general is not symmetric, but the matrix

$$\tilde{\Pi}^n = W^n \Pi^n \qquad (4.39)$$

is symmetric and nonnegative, and positive definite if $\Pi_n$ is. Premultiplying (4.39) by $W^n$, we obtain

$$(A^n)^T \tilde{\Pi}^n + \tilde{\Pi}^n A^n - \tilde{\Pi}^n B^n R^{-1} (B^n)^T \tilde{\Pi}^n + \tilde{Q}^n = 0, \qquad (4.40)$$

which is the Riccati matrix equation to be solved numerically.

Now we need one more set of matrix equations for the numerical solution of the $n^{th}$ optimal control problem. Since the functional gains $f_{in}$ and $g_{in}$ are elements of $V_n$, they can be written as

$$f_{in} = \sum_{j=1}^{n} \beta_j^{f_i} e_j \quad \text{and} \quad g_{in} = \sum_{j=1}^{n} \beta_j^{g_i} e_j, \quad i = 1,\ldots,m, \qquad (4.41)$$

where $\beta^{f_i}, \beta^{g_i} \in R^n$. We need equations for $\beta^{f_i}$ and $\beta^{g_i}$ in terms of $\tilde{\Pi}^n$. One

way to get these equations is to partition $\Pi^n$ (obtained from (4.39)) and then work out the matrix representation of (4.31). However, another approach is more instructive because it relates the present Hilbert space methods to the standard finite dimensional solution of the $n^{th}$ optimal control problem.

The $n^{th}$ optimal control problem can be stated equivalently as: given $\eta(0) = [\xi(0)^T, \dot{\xi}(0)^T] \in R^{2n}$ choose $u \in L_2(0,\infty;R^m)$ to minimize

$$J_n(\eta(0),u) = \int_0^\infty (\eta(t)^T \tilde{Q}^n \eta(t) + u(t)^T R u(t))dt, \tag{4.42}$$

where $\eta(t) = [\xi(t)^T, \dot{\xi}(t)^T]^T$ satisfies (4.2'). For (4.2') and (4.42), the optimal control law is

$$u_n = -R^{-1} B^n \tilde{\Pi}^n \eta(t) \tag{4.43}$$

where $\tilde{\Pi}^n$ is the minimal nonnegative, symmetric solution to (4.40).

Since $\xi$ is related to $x_n$ by (4.1), the optimal control $u_n$ in (4.43) must be equal to the optimal control $u_n$ in (4.29)-(4.31). Substituting (4.41) into (4.30) yields

$$u_{in} = -(\beta^{f_i})^T \tilde{K}^n \xi - (\beta^{g_i})^T M^n \dot{\xi} = -[(\beta^{f_i})^T (\beta^{g_i})^T] W^n \eta. \tag{4.44}$$

Then, using (4.44) and equating (4.29) and (4.43) yields

$$\begin{bmatrix} \beta^{f_1} & \beta^{f_2} \cdots & \beta^{f_m} \\ \beta^{g_1} & \beta^{g_2} \cdots & \beta^{g_m} \end{bmatrix} = W^{-n} \tilde{\Pi}^n B^n R^{-1} . \tag{4.45}$$

We now have the complete solution to the $n^{th}$ optimal control problem:

The Riccati matrix equation (4.40) is solved for $\tilde{\Pi}^n$; then the optimal control is given by (4.43), and equivalently by (4.29)-(4.30) with the functional gains $f_{in}$ and $g_{in}$ given by (4.41) and (4.45). In the next section, we will give sufficient conditions for the solution to the $n^{th}$ optimal control problem to converge to the solution to the optimal control problem in Section 3 for the original infinite dimensional system.

## 5. Convergence

As in Section 3, subsection 5.1 will state some results for the optimal linear regulator problem involving generic linear operators A, B, Q, etc., on an arbitrary real Hilbert space E, and subsection 5.2 will expand upon these results for the particular class of control problems treated in this paper.

### 5.1 Generic Approximation Results

Let the Hilbert space E and the linear operators A, $T(\cdot)$, B, Q and R be as in Section 3. Suppose that there is a sequence of finite dimensional subspaces $E_n$, with the projection of E onto $E_n$ denoted by $P_{En}$, such that $P_{En}$ converges strongly to the identity as $n \to \infty$, and suppose that there exist sequences of operators $A_n \in L(E_n)$, $B_n \in L(R^m, E_n)$, $Q_n = Q_n^* \in L(E_n)$, $Q_n \geq 0$, such that we have the following strong convergence. For all $z \in E$ and $t \geq 0$,

$$\exp(A_n t)P_{En}z \to T(t)z \tag{5.1}$$

and

$$\exp(A_n^* t)P_{En}z \to T^*(t)z \tag{5.2}$$

as $n \to \infty$, uniformly in t for t in bounded intervals; for each $u \in R^m$,

$$B_n u \to Bu; \tag{5.3}$$

for each $z \in E$,

$$Q_n P_{En} z \to Qz. \tag{5.4}$$

Theorem 5.1. Suppose that for each n there is a nonnegative, selfadjoint linear operator $\Pi_n$ on $E_n$ which satisfies the Riccati algebraic equation

$$A_n^* \Pi_n + \Pi_n A_n - \Pi_n B_n R^{-1} B_n^* \Pi_n + Q_n = 0. \tag{5.5}$$

If there exist positive constants M and $\beta$, independent of n, such that

$$||\exp([A_n - B_n R^{-1} B_n^* \Pi_n]t)|| \leq Me^{-\beta t}, \quad t \geq 0, \tag{5.6}$$

and if $||\Pi_n||$ is bounded uniformly in n, then the Riccati algebraic equation (3.3) has a nonnegative selfadjoint solution $\Pi$, and, for each $z \in E$,

$$\Pi_n P_{En} z \to \Pi z \tag{5.7}$$

and

$$\exp([A_n - B_n R^{-1} B_n^* \Pi_n]t) P_{En} z \to S(t)z \tag{5.8}$$

uniformly in $t \geq 0$, where $S(\cdot)$ is the semigroup generated by $A - BR^{-1}B^*\Pi$. If there exists a positive constant $\delta$, independent of n, such that

$$Q_n \geq \delta, \tag{5.9}$$

then $||\Pi_n||$ being bounded uniformly in n guarantees the existence of positive constants M and $\beta$ for which (5.6) holds for all n.

Proof. The theorem follows from Theorem 5.3 of [G4] when the operators $A_n$, $Q_n$ and $\Pi_n$ are extended to all of E by defining them appropriately on $E_n^{\perp}$. For the details of this procedure, see Section 4 of [G1]. Or better, Banks and Kunisch [B6] have modified Theorem 5.3 of [G4] to obtain essentially the present theorem without using the artificial, and rather clumsy, extensions to $E_n^{\perp}$ in the proof. □

**Theorem 5.2** The strong convergence in (5.7) implies uniform norm convergence of the optimal feedback laws:

$$\|B_n^* \Pi_n P_{E_n} - B^* \Pi\| \to 0 \quad \text{as} \quad n \to \infty. \tag{5.10}$$

**Proof.** This follows from the selfadjointness of $\Pi_n$ and $P_{En}$ and the finite dimensionality of the control space $R^m$. See equations (4.23) and (4.24) of [G1]. □

**Theorem 5.3** Assume the hypotheses of Theorem 5.1 but do not assume (5.6) or (5.9). If $\|\Pi_n\|$ is bounded uniformly in n, then the Riccati algebraic equation (3.3) has a nonnegative selfadjoint solution $\Pi$, and, for each $z \in E$, $\Pi_n P_{En} z$ converges weakly to $\Pi z$.

**Proof.** This is Theorem 6.7 of [G3], whose proof is valid under the hypotheses here. □

The main shortcoming of the weak convergence in Theorem 5.3 is that it does not yield uniform norm convergence of the feedback control laws.

## 5.2 Convergence of the Approximating Optimal Control Problems of Section 4.2

For the rest of this section, $A_0$, $A_1$, A, T(t), $B_0$ and B are the operators defined in Section 2. The operators $A_n$, $B_n$, $Q_n$ and $\Pi_n$ are the operators in the approximation scheme of Section 4. In particular, $\Pi \in L(E_n, E_n)$ is the minimal nonnegative, self-adjoint solution of the Riccati operator equation (4.26). According to Corollary 4.3 and Theorem 4.5, the Ritz-Galerkin approximation scheme presented in Section 4.1 converges as required in (5.1) and (5.2); (5.3) and (5.4) follow from (4.23) and the definition $Q_n = P_{En} Q|_{En}$ in Section 4.2. Also, Hypothesis 4.6 guarantees for each n the existence of the required solution of the Riccati equation (5.5) in Theorem 5.1.

Since $\Pi_n$ is nonnegative and self-adjoint, its eigenvalues, which are also the eigenvalues of its matrix representation, are real and nonnegative, and its norm is equal to its maximum eigenvalue.

Theorem 5.4 If Q is E-coercive and $d_0 = 0$ (i.e., there is no open-loop damping), then there is no nonnegative self-adjoint solution of the Riccati operator equation (3.3), and

$$\| \Pi_n \| \to \infty \quad \text{as} \quad n \to \infty. \tag{5.11}$$

Proof. Recall the operator $\tilde{A}$ in Section 2.1. By Theorem 1 of [G2], there can be no compact operator $C \in L(E,E)$ such that $\tilde{A} + C$ generates a uniformly exponentially stable semigroup. Therefore, since a compact linear perturbation of $\tilde{A}$ yields A, there can be no compact linear C such that A + C generates a uniformly exponentially stable semigroup.

Now, unless (5.11) holds, there exists a subsequence such that $\| \Pi_{n_j} \|$ is bounded in $n_j$, so that Theorem 5.3 says that there exists a nonnegative self-adjoint solution $\Pi$ of (3.3). Since Q is coercive, Theorem 3.3 then says that the semigroup generated by $A - BR^{-1}B^* \Pi$ is uniformly exponentially stable. But this is impossible -- $BR^{-1}B^* \Pi$ is compact because its rank is not greater than m. □

Theorem 5.5. Suppose that $A_0$ and $d_0(\cdot,\cdot)$ are both H-coercive. Then there exist positive constants $M_1$ and $\alpha_1$, independent of n, such that

$$||\exp[A_n t]|| \leq M_1 e^{-\alpha_1 t}, \quad t \geq 0. \tag{5.12}$$

Proof. First, we define $A_{0n}$ and $D_{0n} \in L(V_n, V_n)$ to be the operators whose matrix representations are $M^{-n}K^n$ and $M^{-n}D^n$, respectively. (See Section 4.1.) The operator $A_n$ is then

$$A_n = \begin{bmatrix} 0 & I \\ -A_{0n} & -D_{0n} \end{bmatrix} . \tag{5.13}$$

Since $A_0$ and $d_0$ are H-coercive, there exists a positive constant $\rho$, independent of n, such that

$$\langle A_{0n} h, h \rangle_H \geq \rho ||h||_H^2 \tag{5.14}$$

and

$$\langle D_{0n} h, h \rangle_H \geq \rho ||h||_H^2 \tag{5.15}$$

for all $h \in V_n$. Since $d_0$ is continuous on $V \times V$, there exists a positive

constant $\gamma$, independent of n, such that

$$\langle D_{0n}h,h\rangle_H \leq \gamma\|h\|_V^2 \tag{5.16}$$

for all $h \in V_n$. The theorem follows then from Theorem 2.1. □

Theorem 5.6. Suppose that $A_0$ has an invariant subspace $V_0$ which is also invariant under the damping map $\Lambda_V^{-1}\Lambda_D$, that $E_0 = V_0 \times V_0$ is a stabilizable subspace for the control system, and that the restrictions of $A_0$ and $d_0(\cdot,\cdot)$ to $V_0^\perp$ are both H-coercive. Also, suppose that $V_0$ has finite dimension $n_0$ and that, for each $n \geq n_0$ in the approximation scheme, the first $n_0$ $e_i$'s span $V_0$ and the rest are orthogonal to $V_0$ in both V and H.

i) Then (3.3) has a nonnegative solution $\Pi$, and for each $n \geq n_0$, (5.5) has a nonnegative self-adjoint solution $\Pi_n$. Also, $\Pi_n$ is bounded uniformly in n, so that $\Pi_n$ converges to $\Pi$ weakly, as in Theorem 5.3.

ii) If $E_0$ and $E_0^\perp$ (the E-orthogonal complement of $E_0$) are invariant under Q, and if the part of the open-loop system on $E_0$ is observable with the measurement Qz, then (5.6)-(5.8) hold as in Theorem 5.1.

Proof. We will invoke Theorem 3.7 to establish the existence of the uniform bounds and decay rates needed in Theorem 5.1. In the approximating optimal control problems, the part of the control system on $E_0$ is the same for each n; the approximation of the control system takes place on $E_0^\perp$. We can write $E_n = E_0 \oplus E_{0n}^\perp$, where $E_{0n}^\perp$ is the orthogonal complement of $E_0$ in $E_n$, and $E_0$ and $E_{0n}^\perp$ clearly reduce the open-loop semigroup for each $n \geq n_0$.

For the part of the open-loop system on $E_{0n}$, Theorem 5.5 establishes

positive $M_1'$ and $\alpha_1'$, independent of n, for (3.17). Also, we have a $\beta$ independent of n for (3.18) because $B_n = P_{En}B$ and $Q_n = P_{En}Q|_{En}$. Therefore, i) follows from Theorem 3.7 i) and Theorem 5.3.

The definition of $Q_n$ and the requirement on where the various basis vectors must lie imply that $E_0$ and $E_{0n}^{\perp}$ reduce $Q_n$ if $E_0$ and $E_0^{\perp}$ reduce Q and that the restriction of Q to $E_0$ is the same for all n. Therefore, ii) follows from Theorem 3.7 ii) and Theorem 5.1. □

Remark 5.7. In applications, the subspace $V_0$ in Theorem 5.6 usually contains rigid-body modes. The theorem includes the case where both $A_0$ and $d_0$ are H-coercive on all of V (no rigid-body modes and all modes damped). In this case, $V_0$ is the trivial subspace. □

Remark 5.8. Otherwise, for applications to flexible structures, Theorem 5.6 usually requires two things: first, modal damping must be modeled for the structure, so that the natural modes remain uncoupled in the open-loop system; second, the natural mode shapes must be used for the basis functions in the approximating optimal control problems. Although these requirements may seem restrictive from a mathematical standpoint, such modeling and approximation predominate in engineering practice. Also, we get our strongest convergence results under these conditions. For applications where the basis vectors are not the natural mode shapes, the following theorem is useful. □

Theorem 5.9. Suppose that $A_0 + B_0B_0^*$ and $d_0 + B_0B_0^*$ are H-coercive. Then (3.3) has nonnegative solution $\Pi$, for each n (5.5) has a nonnegative self-adjoint solution $\Pi_n$, and $\|\Pi_n\|$ is bounded uniformly in n. Hence Theorem 5.3 applies. Furthermore, if Q is E-coercive, then (5.6)-(5.9) hold in Theorem 5.1.

Proof. The required bounds follow from Theorem 3.10 and the proof of Theorem 5.5. Although we took $A_1 = B_0 B_0^*$ in Theorem 3.10, this is not necessary in the final result, since all bounded self-adjoint operators $A_1$ on H that make $A_0 + A_1$ coercive yield equivalent norms for V. □

Theorem 5.10. If (5.7) holds for each $z \in E$, then

$$\|f_{in} - f_i\|_V \to 0, \tag{5.17a}$$

$$\|g_{in} - g_i\|_H \to 0, \quad \text{as} \quad n \to \infty, \tag{5.17b}$$

where $f_i$ and $g_i$ are the functional gains in (3.16), and $f_{in}$ and $g_{in}$ are the approximating functional gains in (4.31) and (4.41).

Proof. The result follows from (4.31). □

Note that (5.10) and (5.17) are equivalent.

# 6. Example

## 6.1 The Control System

One end of the uniform Euler–Bernoulli beam in Figure 6.1 is attached rigidly (cantilevered) to a rigid hub (disc) which is free to rotate about its center, point 0, which is fixed. Also, a point mass $m_1$ is attached to the other end of the beam. The control is a torque u applied to the disc, and all motion is in the plane.

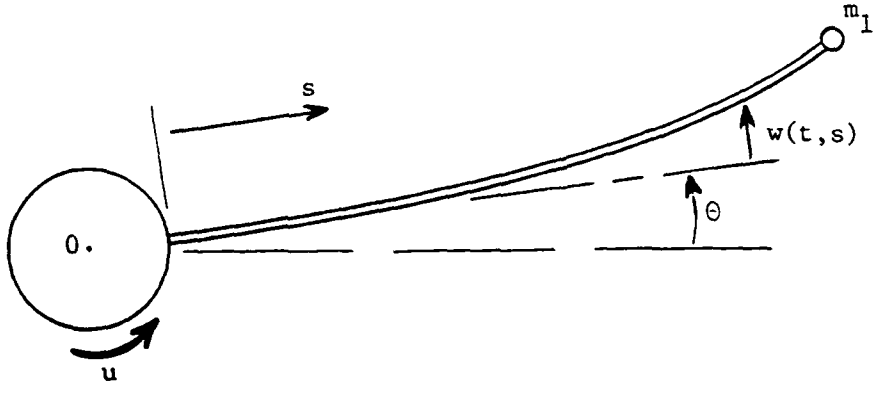



Figure 6.1. Control System

| | |
|---|---|
| $r$ = hub radius | 10 in |
| $\ell$ = beam length | 100 in |
| $I_0$ = hub moment of inertia about axis perpendicular to page through 0 | 100 slug $in^2$ |
| $m_b$ = beam mass per unit length | .01 slug/in |
| $m_1$ = tip mass | 1 slug |
| EI = product of elastic modulus and second moment of cross section for beam | 13,333 slg $in^3/sec^2$ |
| fundamental frequency of undamped structure | .9672 rad/sec |

Table 6.1 Structural Data

The angle $\theta$ represents the rotation of the disc (the rigid-body mode), $w(t,s)$ is the elastic deflection of the beam from the rigid-body position, and $w_1(t)$ is the displacement of $m_1$ from the rigid-body position. For technical reasons, we do not yet impose the condition $w_1(t) = w(t,\ell)$; more on this later.

The control problem is to stabilize rigid-body motions and linear (small) transverse elastic vibrations about the state $\theta = 0$ and $w = 0$. Our linear model assumes not only that the elastic deflection of the beam is linear but also that the axial inertial force produced by the rigid-body angular velocity has negligible effect on the bending stiffness of the beam. The rigid-body angle need not be small.

For this example, it is a straight forward exercise to derive the three coupled differential equations of motion in $\theta$, $w$ and $w_1$, and they do have the

form (2.1'). However, to emphasize the fact that we do not use the explicit partial differential equations, we will not write these equations here. Rather, we will write only what is normally needed in applications: the kinetic and strain-energy functionals, the damping functional and the actuator influence operator.

Remark 2.1 applies to this example, and to most examples with complex structures. The generalized displacement vector is

$$x = (\theta, w, w_1) \in H_0 = R \; L_2(0,\ell) \times R. \tag{6.1}$$

The kinetic energy in the system is

$$\text{Kinetic Energy} = \tfrac{1}{2} \langle \dot{x}, \dot{x} \rangle_H \tag{6.2}$$

where H is $H_0$ with the inner product

$$\langle x, \hat{x} \rangle_H = m_b \int_0^{\ell} [w+(r+s)\theta][\hat{w}+(r+s)\hat{\theta}]ds \tag{6.3}$$

$$+ \; I_0 \theta\hat{\theta} \; + \; m_1[w_1+(r+\ell)\theta][\hat{w}_1+(r+\ell)\hat{\theta}].$$

As in most applications, we need not write the mass operator explicitly, but there exists a unique selfadjoint linear operator $M_0$ on $H_0$ such that

$$\langle x, \hat{x} \rangle_H = \langle M_0 x, \hat{x} \rangle_{H_0} . \tag{6.4}$$

It is easy to see that $M_0$ is bounded and coercive. Hence $H_0$ and H have equivalent norms.

The input operator for (2.1') (which maps R to $H_0$) is

$$B_0 = (1,0,0). \tag{6.5}$$

Since we multiply (2.1') by $M_0^{-1}$ to get (2.1), the input operator for (2.1) is $(M_0^{-1}B_0)$. Note that

$$(M_0^{-1}B_0)^{*H} = B_0^*, \tag{6.6}$$

where $(M_0^{-1}B_0)^{*H}$ is the H-adjoint of $(M_0^{-1}B_0)$ and $B_0^*$ is the $H_0$-adjoint of $B_0$.

Remark 2.2 also applies here. The only strain energy is in the beam and is given by

$$\text{Strain Energy} = \tfrac{1}{2}\, a(x,x) \tag{6.7}$$

with

$$a(x,\hat{x}) = EI\int_0^{\ell} w''\hat{w}''ds, \tag{6.8}$$

where $(\cdot)'' = \partial^2(\cdot)/\partial s^2(\cdot)$. To make $a(\cdot,\cdot)$ into an inner product, we must account for rigid-body rotation. Thus we set

$$\langle x,\hat{x}\rangle_V = a(x,\hat{x}) + \theta\hat{\theta} \tag{6.9}$$

and define

$$V = \{x = (\theta,\phi,\phi(1)):\ \phi \in H^2(0,\ell),\ \phi(0) = \phi'(0) = 0\ \}. \tag{6.10}$$

Also, we have

$$\langle x,\hat{x}\rangle_V = a(x,\hat{x}) + \langle B_0B_0^*x,\hat{x}\rangle_{H_0} \tag{6.11}$$

$$= a(x,\hat{x}) + \langle (M_0^{-1}B_0)(M_0^{-1}B_0)^{*H}x,\hat{x}\rangle_H,$$

so that $A_1 = B_0 B_0^*$, or $(M_0^{-1}B_0)(M_0^{-1}B_0)^{*H}$, depending on whether the $H_0$ or the H-inner product is used in computing the V-inner product. But we need neither $A_1$ nor $A_0$ explicitly. We need only (6.8) and (6.9), along with (6.3), to compute the required inner products.

As mentioned in Remark 2.2, the operator $\tilde{A}_0$ can be defined now by (2.6), and the stiffness operator is $A_0 = \tilde{A}_0 - A_1$. Using the $H_0$-inner product in (2.6) yields the $A_0$ for (2.1'), and using the H-inner product yields the $A_0$ for (2.1), which is $M_0^{-1}$ following the $A_0$ for (2.1'). The $A_0$ for (2.1') is quite simple, and the reader might write it out. We will not, so that no one will think that we use it. We will point out that $D(A_0)$ requires both the geometric boundary conditions in V and the natural boundary condition $w''(t,\ell) = 0$; i.e., zero moment on the right end.

Remark 6.1. That the geometric boundary conditions

$$w(t,0) = w'(t,0) = 0 \tag{6.12}$$

and

$$w(t,\ell) = w_1(t) \tag{6.13}$$

are imposed in V but not in H -- i.e., on the generalized displacement but not on the generalized velocity -- is common in distributed models of flexible structures. The natural norm for expressing the kinetic energy of distributed components is the $L_2$-norm, which cannot preserve constraints on sets of zero measure. Because the strain energy involves spacial derivatives, the stronger strain-energy norm can preserve the geometric boundary conditions (although, as for the boundary slope of an elastic plate, the V-norm may impose some of

these boundary conditions in an $L_2$ rather than a pointwise sense). The strain-energy norm is based on the material model of the distributed components of the system, and it should not be surprising that such a norm is required to connect the various structural components. □

We assume that the beam has Voigt-Kelvin viscoelastic damping [C2], so that the damping operator in (2.1) is

$$D_0 = c_0 A_0 \tag{6.14}$$

where $c_0$ is a constant. This means that the damping functional is

$$d_0(x,\hat{x}) = c_0\, a(x,\hat{x}), \quad x,\hat{x} \in V. \tag{6.15}$$

## 6.2 The Optimal Control Problem

We take Q = I in the performance index in (3.1). This means that the state weighting term $\langle Qz,z\rangle_E$ is twice the total energy in the structure plus the square of the rigid-body rotation. Since there is one input, the control weighting R is a scalar.

According to (3.33), the optimal control has the feedback form

$$u(t) = -\langle f,x(t)\rangle_V - \langle g,\dot{x}(t)\rangle_H \tag{6.16}$$

where x(t) has the form (6.1), and

$$f = (a_f,\phi_f,\beta_f) = R^{-1}\Pi_1 B_0 \in V, \tag{6.17a}$$

$$g = (\alpha_g, \phi_g, \beta_g) = R^{-1} \Pi_2 B_0 \in H. \tag{6.17b}$$

Note that $\beta_f = \phi_f(\ell)$ is not used in the control law--recall (6.8) and (6.9).

## 6.3 Approximation

Our approximation of the distributed model of the structure is based on a finite element approximation of the beam which uses Hermite cubic splines as basis functions ([S1,S4]). These are the basis functions most commonly used in engineering finite element approximations of beams. The splines and their first derivatives are continuous at the nodes. Because the basis vectors $e_j$ in the approximation scheme in Section 4 must be in the space V defined in (6.10), we write them as

$$e_1 = (1,0,0), \tag{6.18a}$$

$$e_j = (0,\phi_j,\phi_j(1)), \quad j = 2, 3, \ldots, \tag{6.18b}$$

where the $\phi_j$'s are the cubic splines. When we use $n_e$ elements to approximate the beam, there are $2n_e$ linearly independent splines. Thus, with the rigid-body mode, the order of approximation is $n = 2n_e + 1$.

For the numerical solution to the optimal control problem, we have only to plug into the formulas of Section 4. The matrices in (4.3) are calculated according to (6.3), (6.8) and (6.9), with $B_0$ given by (6.5). In particular,

$$K^n = [a(e_i,e_j)], \quad D^n = c_0 K^n, \quad M^n = [\langle e_i, e_j \rangle_H], \tag{6.19}$$

$$B_0^n = [1\ 0\ 0\ \cdots\ 0]^T = [\langle e_i, M_0^{-1}(1,0,0)\rangle_H] = [\langle e_i,(1,0,0)\rangle_{H_0}] \tag{6.20}$$

Note that the first row and column of $K^n$ are zero. The matrix $\tilde{K}^n$ in (4.32) is $K^n$ with 1 added to the first element. The matrices $A^n$ and $B^n$ are given by (4.5) and, since $Q = I$, the matrix $Q^n$ is the $W^n$ in (4.33). With these matrices, we solve the Riccati equation (4.40) and use (4.41) and (4.45) to compute the approximations to the functional gains, which are

$$f_n = (\alpha_{fn}, \phi_{fn}, \beta_{fn}), \tag{6.21a}$$

$$g_n = (\alpha_{gn}, \phi_{gn}, \beta_{gn}). \tag{6.21b}$$

For convergence, we satisfy all the hypotheses of Theorem 5.9. In particular, since Q is the identity on E, it is coercive. Theorem 5.9 implies that the solutions to the finite dimensional Riccati equations converge as in Theorem 5.1 and that the functional control gains converge as in theorem 5.10.

Remark 6.2. It might appear that the hypotheses of Theorem 5.6 hold with $n_0 = 1$, but not so. For $j \geq 2$, $e_j$ is orthogonal to $e_1$ in $H_0$ and V not in H. Recall (6.1)-(6.3), (6.9)-(6.11) and (6.18). □

## 6.4 Numerical Results

Figures 6.2a and 6.2b show the computed functional gain kernals $\phi''_{fn}$ and $\phi_{gn}$ for the damping coefficient $c_0 = 10^{-4}$, the control weighting R = 1, and $n_e$ = 2, 3, 4, 5 and 8 beam elements. Table 6.2 lists the corresponding scalar components of the gains. For $c_0 = 10^{-4}$ and R = 05, the convergence is slower, as discussed below. To show the complete story of convergence, Figures 6.3a and 6.3b and Table 6.3 show the results for $n_e$ = 2,3,4,5,8, and Figures 6.4a and 6.4b and Table 6.4 show the results for $n_e$ = 4,6,8,10.

We have plotted $\phi''_{fn}$ because the second derivative appears in the strain-energy inner product in (6.8) and (6.9) and $\phi_{fn}$ converges in $H^2(0,\ell)$. Note that, since the Hermite cubic splines have discontinuous second derivatives at the nodes, the approximations to $\phi''_f$ are discontinuous at the nodes. Although $H^2$-convergence guarantees only $L_2$-convergence for $\phi''_{fn}$, it can be shown that $\phi''_{fn}$ converges uniformly on $[0,\ell]$ for this problem.

The tables omit $\beta_{fn}$ to emphasize the fact that it does not appear in the feedback law and the fact that the convergence of $\beta_{fn}$ is not an independent piece of information about the convergence of the control gains; since $\phi_{fn}(0) = \phi'_{fn}(0) = 0$, the convergence of $\phi''_{fn}$ implies the convergence of $\beta_{fn} = \phi_{fn}(\ell)$. On the other hand, although $\beta_{gn} = \phi_{gn}(\ell)$ for each n, the H-norm convergence of $g_n$ does not enforce this condition in the limit, as the V-norm convergence of $f_n$ enforces $\beta_f = \phi_f(\ell)$. Hence, as far as we can tell from our results in Sections 3.5, $\beta_{fn}$ is an independent indicator of the convergence of the control gains, as well as being used in the control law in (6.16). However, the behavior of $\phi_{gn}$ in Figures 6.2b, 6.3b and 6.4b suggests that $g_n$ converges in V. Stronger results on the continuity of $\phi_g$ and the convergence of $\phi_{gn}(\ell)$

should follow from a theorem stating that, because the open-loop semigroup generator A is analytic, the solution to the infinite dimensional Riccati equation maps all of E into $D(A^*)$. The fact that $\phi''_{fn}(\ell)$ converges to zero in Figure 6.2a also suggests such a theorem, but we have not proved it.

With the state weighting Q fixed, the two factors that determine the rate of convergence are $c_0$ and R. Although we have used splines to approximate the beam, the relation between the convergence rate and $c_0$ and R probably can be interpreted best in terms of the number of natural modes of the structure that the optimal infinite dimensional controller really controls. Strictly speaking, the controller controls all modes, but the functional gains lie essentially in the span of some finite number of modes. This would be the number of modes required for convergence of the gains if we used the natural modes as the basis vectors in the approximation. The rest of the modes are practically (but not exactly) orthogonal to the functional gains, so that the optimal feedback law essentially ignores them. In general, the lighter the damping, the more modes that will be controlled for given Q and R; the cheaper the control, the more modes that will be controlled for given Q and $c_0$. The question of the convergence of the finite element approximation to the functional gains becomes then a question of how many modes the optimal control law really wants and how many elements it takes to approximate those modes.

Numerical experience with optimal control of flexible structures has shown this modal interpretation of the convergence of the approximating control laws to be very useful, and that it is difficult to improve upon the natural modes as basis vectors for the approximation scheme (see [G5]). However, whether the natural modes are always or almost always the best basis

vectors is an open question. We use the cubic splines here to demonstrate that a standard finite element approximation works quite well. Also, to use the natural modes as basis vectors here, we first would have to compute them using a finite element approximation -- as in most real problems -- and we do not know in advance which or how many modes are needed. On the other hand, if the most important natural modes are determined from experiment, then modal approximation should be best.

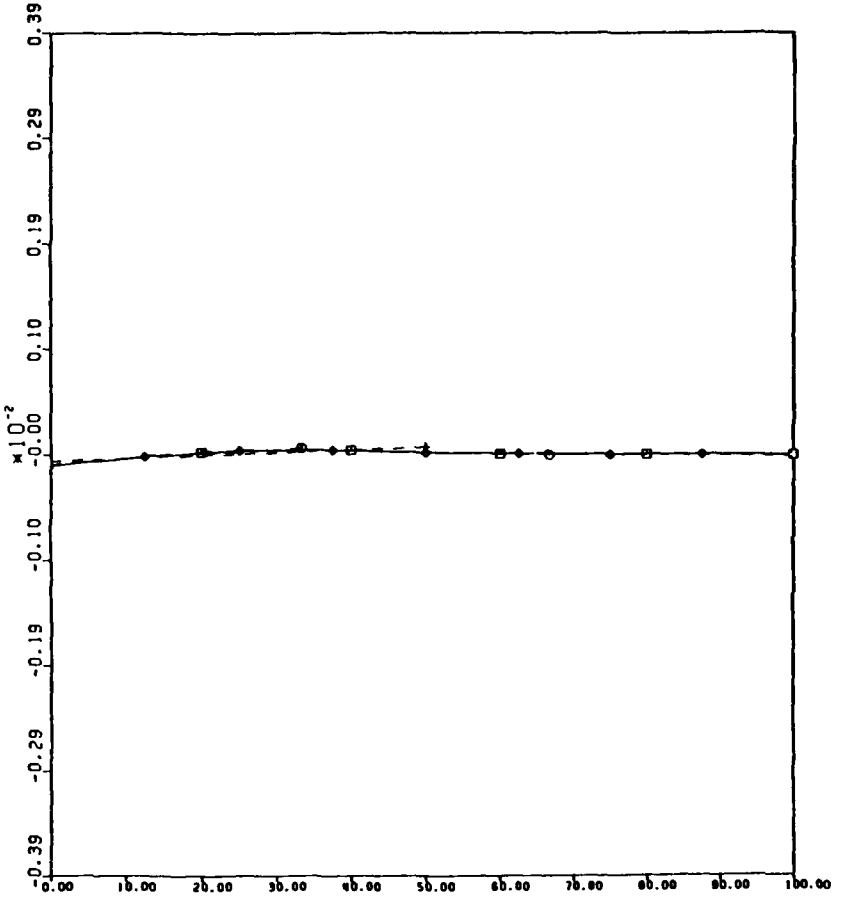



Figure 6.2a. Functional Control Gain Component $\phi_{fn}$''

Damping coefficient $c_0 = 10^{-4}$; control weighting R = 1

number of elements $n_e$ = 2, 3, 4, 5, 8

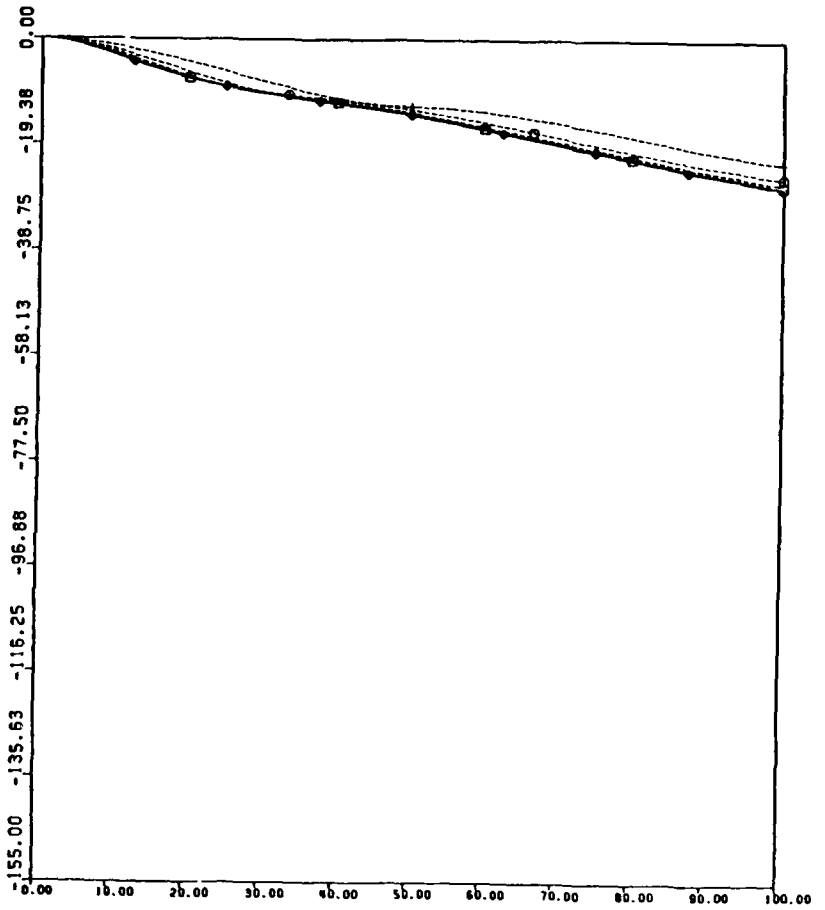



Figure 6.2b. Functional Control Garin Component $\phi_{gn}$

Damping coefficient $c_0 = 10^{-4}$; control weighting R = 1

number elements $n_e$ = 2, 3, 4, 5, 8

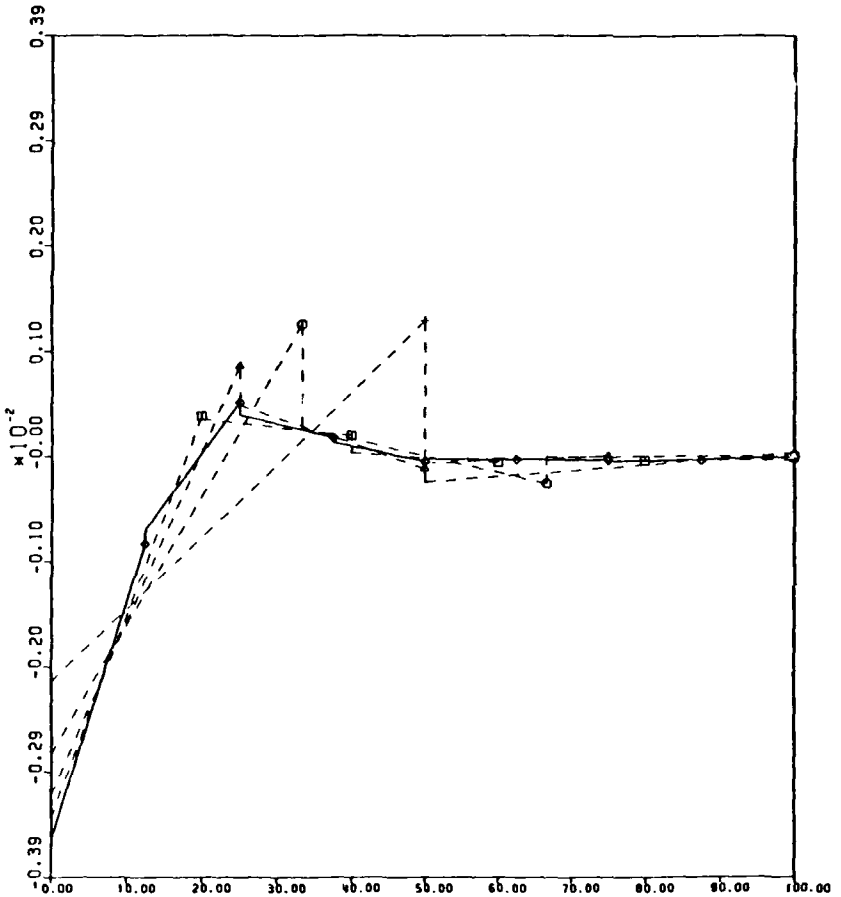



Figure 6.3a. Functional Control Gain Component $\phi''_{fn}$

Damping coefficient $c_0 = 10^{-4}$; control weighting R = .05

number of elements $n_e$ = 2, 3, 4, 5, 8

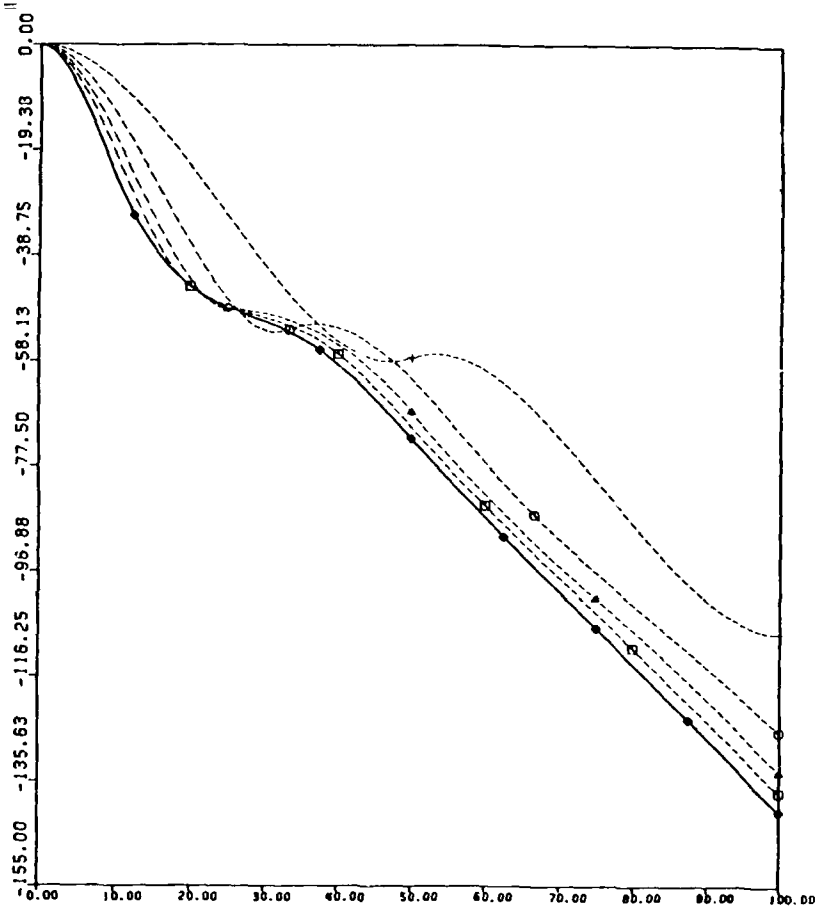



Figure 6.3b. Functional Control Gain Component $\phi_{gn}$

Damping coefficient $c_0 = 10^{-4}$; control weighting R = .05

number of elements $n_e$ = 2, 3, 4, 5, 8

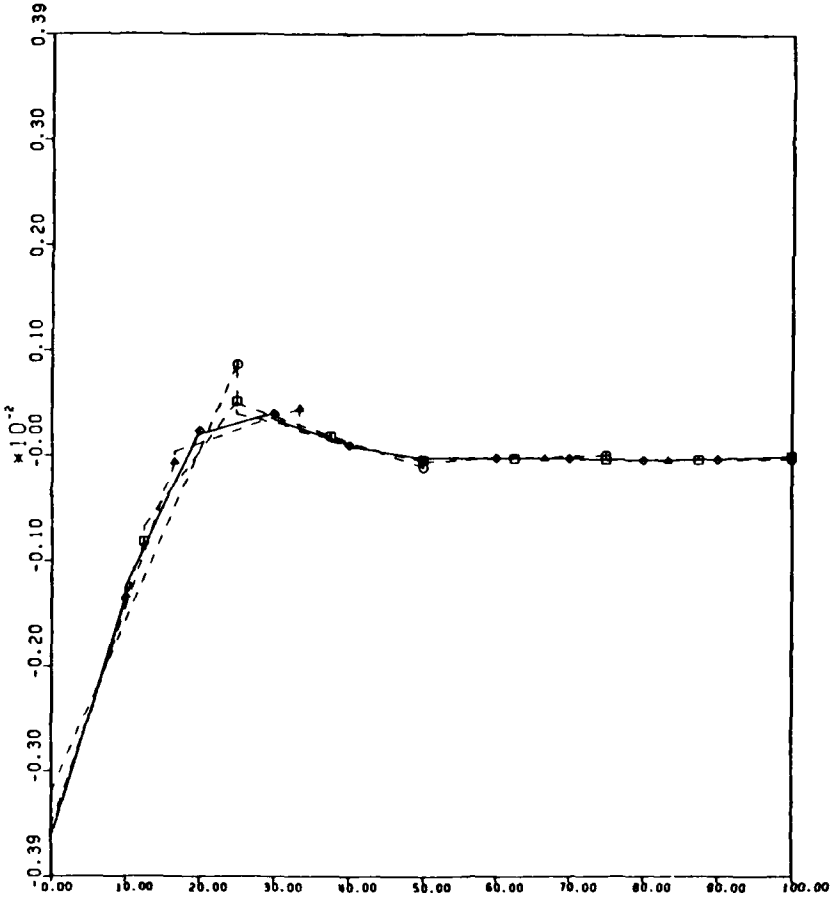



Figure 6.4a. Functional Control Gain Component $\phi_{fn}$''

Damping coefficient $c_0 = 10^{-4}$; control weighting R = .05

number of elements $n_e$ = 4, 6, 8, 10

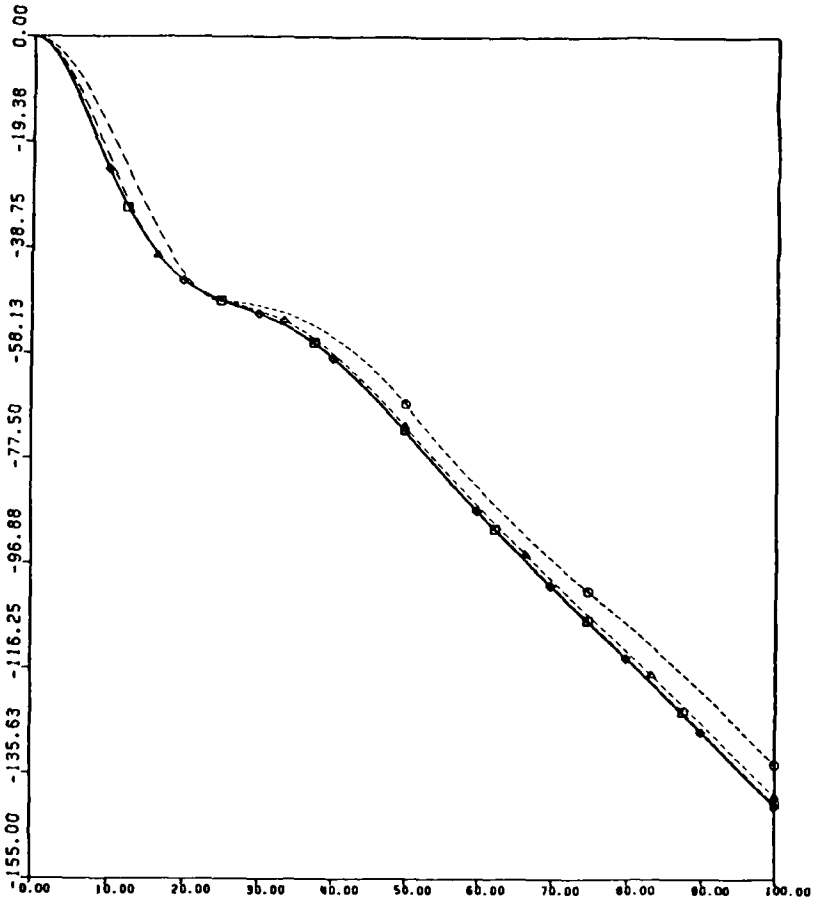



Figure 6.4b. Functional Control Gain Component $\phi_{gn}$

Damping coefficient $c_0 = 10^{-4}$; control weighting R = .05

number of elements $n_e$ = 4, 6, 8, 10

| $n_e$ | $\alpha_{fn}$ | $\alpha_{gn}$ | $\beta_{gn}$ |
|---|---|---|---|
| 2 | 1.000 | .2141 | -22.459 |
| 3 | 1.000 | .2396 | -25.221 |
| 4 | 1.000 | .2496 | -26.331 |
| 5 | 1.000 | .2534 | -26.786 |
| 8 | 1.000 | .2561 | -27.041 |

Table 6.2. Scalar Components of Functional Control Gains

Damping coefficient $c_0 = 10^{-4}$; control weighting R = 1

number of elements $n_e$ = 2, 3, 4, 5, 8

| $n_e$ | $\alpha_{fn}$ | $\alpha_{gn}$ | $\beta_{gn}$ |
|---|---|---|---|
| 2 | 4.4721 | 1.0136 | -108.27 |
| 3 | 4.4721 | 1.1770 | -126.40 |
| 4 | 4.4721 | 1.2440 | -133.87 |
| 5 | 4.4721 | 1.2781 | -137.57 |
| 8 | 4.4721 | 1.3106 | -141.15 |

Table 6.3. Scalar Components of Functional Control Gains

Damping coefficient $c_0 = 10^{-4}$; control weighting R = .05

number of elements $n_e$ = 2, 3, 4, 5, 8

| $n_e$ | $\alpha_{fn}$ | $\alpha_{gn}$ | $\beta_{gn}$ |
|---|---|---|---|
| 4 | 4.4721 | 1.2440 | -133.87 |
| 6 | 4.4721 | 1.2973 | -139.69 |
| 8 | 4.4721 | 1.3106 | -141.15 |
| 10 | 4.4721 | 1.3141 | -141.54 |

Table 6.4. Scalar Components of Functional Control Gains

Damping coefficient $c_0 = 10^{-4}$; control weighting R = .05

number of elements $n_e$ = 4, 6, 8, 10

Figures 6.5a and 6.5b and Table 6.5 represent attempts to compute an optimal control law for the structure when R = .05 but $c_0 = 0$. Since Q is the identity operator in E and hence coercive, Theorem 5.4 says that no optimal control law exists and that the norm of the solution to the finite dimensional Riccati equation grows without bound as the number of elements increases. This is reflected in the nonconvergence of $\alpha_{gn}$, $\phi_{gn}$ and $\beta_{gn}$, although $\alpha_{fn}$ converges and the convergence of $\phi''_{fn}$ is unclear.

In applications where the structural damping is not known, except that it is very light, it is tempting and not uncommon engineering practice to assume zero damping in the design of a control law for the first few modes, while trusting whatever damping is in the higher modes to take care of them. However, if high performance requirements (large Q) or coupling between modes in the closed-loop system necessitate a control law based on a more accurate approximation of the structure, Theorem 5.4 and the current example warn that the higher-order control laws are likely meaningless and rather strange if no damping is modeled.

We should note that we have seen similar problems [G9] where $\Pi_n$ remains bounded and the gains converge for zero damping but finite-rank Q. In such cases, Theorem 5.3 says that an optimal control law exists for the distributed model of the structure and that the finite dimensional control laws converge to an optimal infinite dimensional control law. Also, Balakrishnan [B2] has shown that an infinite dimensional optimal control law exists for no damping when $Q = BB^*$.

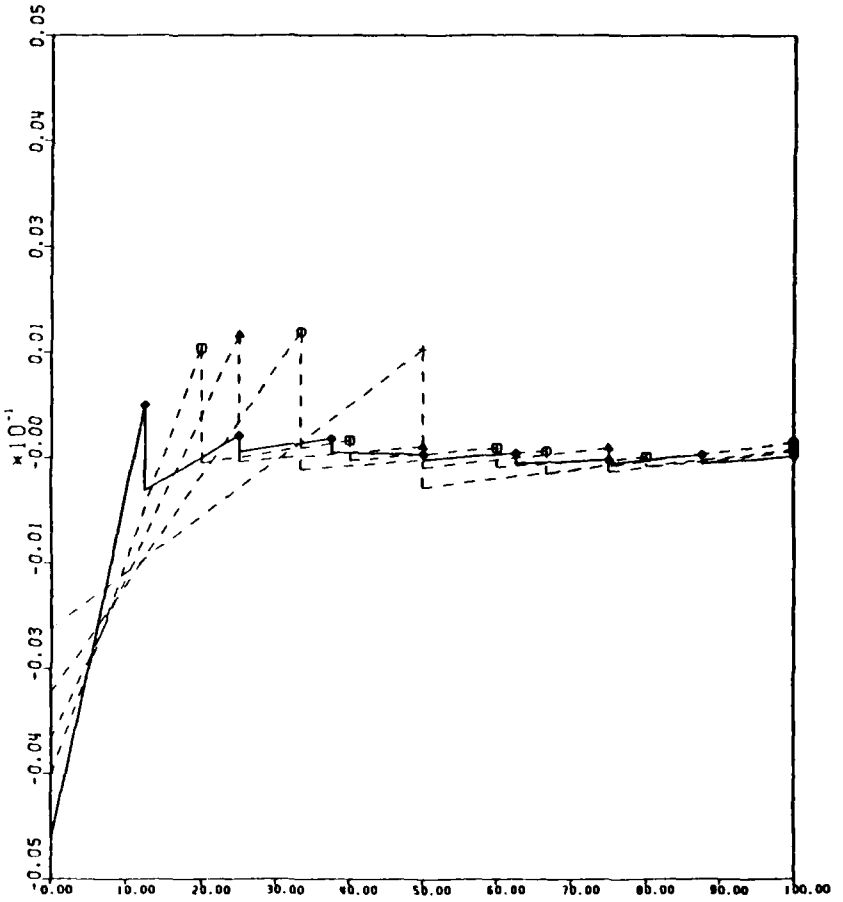



Figure 6.5a. Functional Control Gain Component $\phi_{fn}$''

Zero damping; control weighting R = .05

number of elements $n_e$ = 2, 3, 4, 5, 8

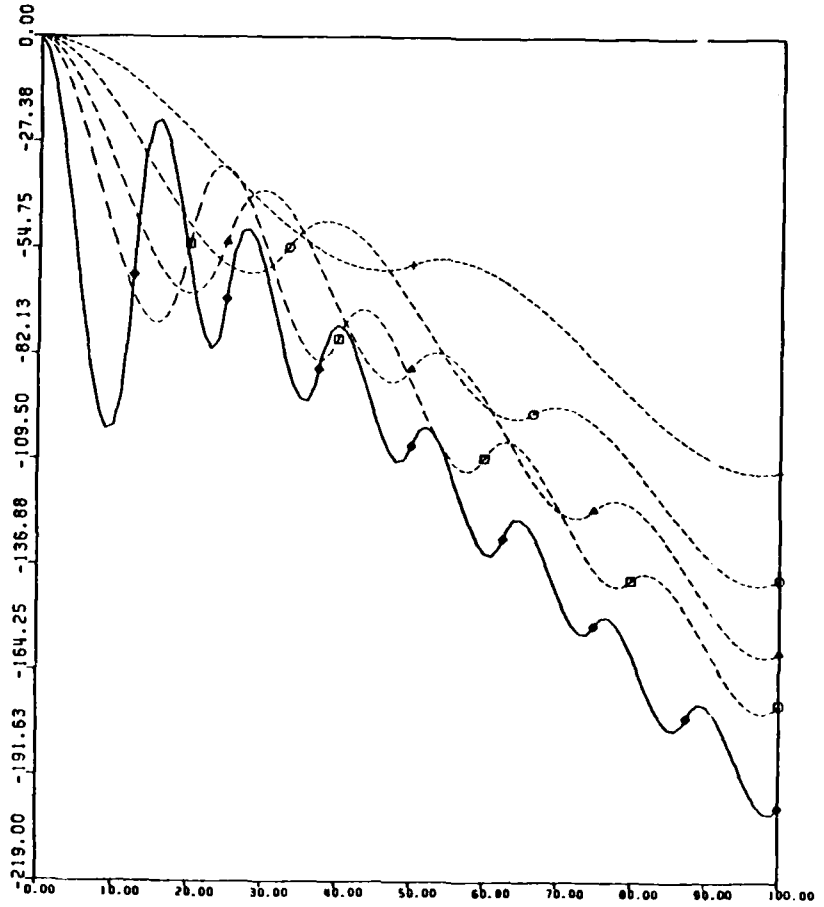



Figure 6.5b. Functional Control Gain Component $\phi_{gn}$

Zero damping; control weighting R = .05

number of elements $n_e$ = 2, 3, 4, 5, 8

| $n_e$ | $\alpha_{fn}$ | $\alpha_{gn}$ | $\beta_{gn}$ |
|---|---|---|---|
| 2 | 4.4721 | 1.0516 | -112.23 |
| 3 | 4.4721 | 1.3061 | -140.18 |
| 4 | 4.4721 | 1.4758 | -159.11 |
| 5 | 4.4721 | 1.5996 | -172.64 |
| 8 | 4.4721 | 1.8407 | -199.39 |

Table 6.5. Scalar Components of Functional Control Gains

Zero damping; control weighting R = .05

number of elements $n_e$ = 2, 3, 4, 5, 8

## 7. The Optimal Infinite Dimensional Estimator, Compensator and Closed-loop System

As in Sections 3 and 5, we will state some initial definitions and results for an arbitrary linear control system on a Hilbert space in Subsection 7.1, and then discuss implications for flexible-structure control in Subsection 7.2.

### 7.1 The Generic Problem

Let A, T(t) and B be as in Subsection 3.1, with E an arbitrary real Hilbert space. The differential equation corresponding to (3.2) is, of course,

$$\dot{z}(t) = Az(t) + Bu(t), \quad t > 0. \tag{7.1}$$

We assume that we have a p-dimensional measurement vector y(t) given by

$$y(t) = C_0u(t) + Cz(t), \tag{7.2}$$

where $C_0 \in L(R^m, R^p)$ and $C \in L(E, R^p)$ for some positive integer p.

Definition 7.1. For any $\hat{F} \in L(R^p, E)$, the system

$$\dot{\hat{z}}(t) = A\hat{z}(t) + Bu(t) + \hat{F}[y(t) - C_0u(t) - C\hat{z}(t)], \quad t > 0, \tag{7.3}$$

will be called an observer, estimator (we use the terms interchangeably), for the system (7.1)-(7.2). Let $\hat{S}(t)$ be the semigroup generated by $A-\hat{F}C$. The observer in (7.3) is strongly (uniformly exponentially) stable if $\hat{S}(t)$ is strongly (uniformly exponentially) stable. □

To justify this definition, we write

$$e(t) = z(t) - \hat{z}(t) \tag{7.4}$$

and, with (7.1)-(7.3), obtain

$$e(t) = \hat{S}(t)e(0), \quad t \geq 0. \tag{7.5}$$

Of course, an observer, or estimator, is necessary because the full state $z(t)$ will not be available for direct feedback, and the feedback control must be based on an estimate of $z(t)$. When, as in this paper, the desired control law has the form

$$u(t) = -Fz(t) \tag{7.6}$$

for some $F \in L(E,R^m)$, the observer in (7.3) can be used to construct $\hat{z}(t)$ from the measurement in (7.2) and then the control law in (7.6) can be applied to $\hat{z}(t)$. The control applied to the system is then

$$u(t) = -\hat{F}z(t), \tag{7.7}$$

and the resulting closed-loop system is

$$\begin{bmatrix} z(t) \\ \hat{z}(t) \end{bmatrix} = S_{\infty,\infty}(t)\begin{bmatrix} z(0) \\ \hat{z}(0) \end{bmatrix}, \quad t \geq 0, \tag{7.8}$$

where $S_{\infty,\infty}(t)$ is the semigroup generated on $E \times E$ by the operator

$$A_{\infty,\infty} = \begin{bmatrix} A & -BF \\ \hat{F}C & [A-BF-\hat{F}C] \end{bmatrix}, \quad D(A_{\infty,\infty}) = D(A) \times D(A). \tag{7.9}$$

With the estimator error $e(t)$ defined by (7.4), it is easy to show that (7.8) is equivalent to (7.5) and

$$\dot{z}(t) = (A-BF)z(t) + BFe(t), \quad t > 0, \tag{7.10}$$

where (A-BF) generates a semigroup S(t) on E. Also, it is easy to prove the following.

Theorem 7.2. Suppose that there exist positive constants $M_1$, $M_2$, $\alpha_1$ and $\alpha_2$ such that

$$\|S(t)\| \leq M_1 e^{-\alpha_1 t},$$

$$\|\hat{S}(t)\| \leq M_2 e^{-\alpha_2 t}, \quad t \geq 0. \tag{7.11}$$

Then, for each real $\alpha_3 < \min\{\alpha_1, \alpha_2\}$, there exists a constant $M_3$ such that

$$\| S_{\infty,\infty}(t)\| \leq M_3 e^{-\alpha_3 t}, \quad t \geq 0. \tag{7.12}$$

Also,

$$\sigma(A_{\infty,\infty}) = \sigma(A-BF) \cup \sigma(A-\hat{F}C), \tag{7.13}$$

where $\sigma(A_{\infty,\infty})$ is the spectrum of $A_{\infty,\infty}$. □

The observer in (7.3) and the control law in (7.7) constitute a compensator for the control system in (7.1) and (7.2). The transfer function of this compensator is

$$\Phi(s) = -F(sI-[A-BF + \hat{F}(C_0F-C)])^{-1}\hat{F}, \tag{7.14}$$

which is an m × p matrix function of the complex variable s. When E has infinite dimension, the compensator transfer function is irrational, except in

degenerate, usually unimportant cases.

The foregoing definitions of this section and Theorem 7.1 are straightforward generalizations to infinite dimensions of observer–controller results in finite dimensions. Balas [B3] and Schumacher [S2] have used similar extensions.

Now suppose that $\hat{F}$ is chosen as

$$\hat{F} = \hat{\Pi} C^* \hat{R}^{-1}, \tag{7.15}$$

where $\hat{\Pi} \in L(E,E)$ is the minimal nonnegative selfadjoint solution to the Riccati equation

$$A\hat{\Pi} + \hat{\Pi} A^* - \hat{\Pi} C^* \hat{R}^{-1} C \hat{\Pi} + \hat{Q} = 0, \tag{7.16}$$

with $\hat{Q} \in L(E,E)$ nonnegative and selfadjoint and $\hat{R} \in L(R^p, R^p)$ symmetric and positive definite. Theorem 3.3 (with A, B, Q, R, $\Pi$ and S(t) replaced by $A^*, C^*, \hat{Q}, \hat{R}, \hat{\Pi}$, and $\hat{S}^*(t)$) gives sufficient conditions for $\hat{\Pi}$ to exist and for the semigroup $\hat{S}^*(t)$ -- and equivalently its adjoint, the $\hat{S}(t)$ generated by $A - \hat{\Pi} C^* \hat{R}^{-1} C$ -- to be uniformly exponentially stable.

Definition 7.3. When the control gain operator is

$$F = R^{-1} B^* \Pi , \tag{7.17}$$

with $\Pi$ the solution to the Riccati equation (3.3), and the observer gain operator is given by (7.15) and (7.16), we will call the compensator consisting of the observer in (7.3) and the control law in (7.7) the optimal infinite dimensional compensator, and (7.8) the optimal closed-loop system. □

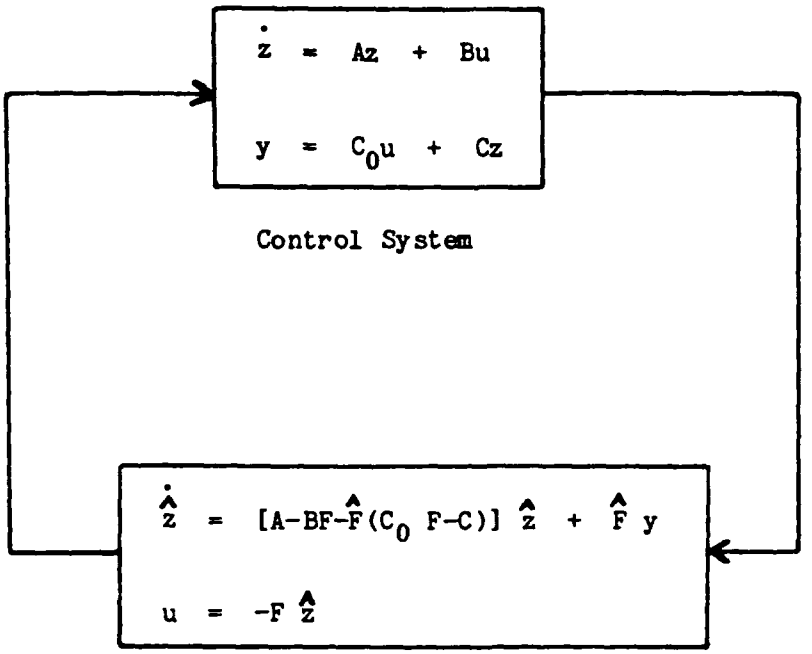



Figure 7.1. Optimal Closed-loop System

Remark 7.4. The infinite dimensional observer defined by (7.3), (7.15) and (7.16) is the optimal estimator for the stochastic version of (7.1) and (7.2) when (7.1) is disturbed by a stationary gaussian white noise process with zero mean and covariance operator $\hat{Q}$ and the measurement in (7.2) is contaminated by similar noise with covariance $\hat{R}$. For infinite dimensional stochastic estimation and control, see [B1, C4]. When the state weighting operator Q in (3.1) is trace class, the optimal infinite dimensional compensator minimizes the time-average of the expected steady-state value of the integrand in (3.1). Existing theory for stochastic control of infinite dimensional systems requires trace-class Q, but we have a well defined compensator for any bounded nonnegative selfadjoint Q and $\hat{Q}$, as long as the solutions to the Riccati

equations exist. As the next two sections show (without assuming trace-class $\hat{Q}$), the infinite dimensional compensator is the limit of a sequence of finite dimensional compensators, each of which can be interpreted as an optimal LQG compensator for a finite dimensional model of the structure. Therefore, we do not require trace-class Q in our definition of the optimal compensator, even though this compensator solves a precise optimization problem only when Q is trace class.

This paper is concerned primarily with how the finite dimensional compensators converge to the infinite dimensional compensator, and the analysis of this convergence requires only the theory of infinite dimensional Riccati equations for deterministic optimal control problems and the corresponding approximation theory. While the stochastic interpretation of the infinite dimensional compensator and, in Section 8.2, of the finite dimensional estimators should be motivational, nothing in the rest of the paper depends on a stochastic formulation. We assume that the operators Q, R, $\hat{Q}$ and $\hat{R}$ are determined by some design criteria. In many engineering applications, deterministic criteria such as the stability margin and robustness of the closed-loop system, rather than a stochastic performance index and an assumed noise model, govern the choice of Q, R, $\hat{Q}$ and $\hat{R}$.□

## 7.2 Application to Structures

For the rest of the paper, $E = V \times H$ as in Section 2, and A and B are the operators defined there.

The measurement operator C in (7.2) now must have the form

$$C = [C_1 \quad C_2] \tag{7.18}$$

where $C_1 \in L(V,R^p)$ and $C_2 \in L(H,R^p)$. Hence, if we denote by $(C(x,\dot{x}))_i$ the $i^{th}$ component of the p-vector $C(x,\dot{x})$, for $(x,\dot{x}) \in E$, then there must exist $c_{1i} \in V$ and $c_{2i} \in H$ such that

$$(C(x,\dot{x}))_i = \langle c_{1i},x\rangle_V + \langle c_{2i},\dot{x}\rangle_H, \quad i = 1,\ldots,p. \tag{7.19}$$

Also, the estimator gain operator F is given by

$$\hat{F}y = \sum_{i=1}^{p} (\hat{f}_i,\hat{g}_i)y_i \tag{7.20}$$

for $y = [y_1\ y_2 \ldots\ y_p]^T \in R^p$, where the <u>functional estimator gains</u> $\hat{f}_i$ and $\hat{g}_i$ are elements of V and H, respectively.

For the optimal estimator gains, we can partition $\hat{\Pi}$ as

$$\hat{\Pi} = \begin{bmatrix} \hat{\Pi}_0 & \hat{\Pi}_1 \\ \hat{\Pi}_1^* & \hat{\Pi}_2 \end{bmatrix} \tag{7.21}$$

and use (7.15) and (7.19) to get

$$\hat{f}_i = \sum_{j=1}^{p} (\hat{R}^{-1})_{ij}(\hat{\Pi}_0 c_{1j} + \hat{\Pi}_1 c_{2j}), \tag{7.22a}$$

$$\hat{g}_i = \sum_{j=1}^{p} (\hat{R}^{-1})_{ij}(\hat{\Pi}_1^* c_{1j} + \hat{\Pi}_2 c_{2j}), \quad i = 1,2,\ldots,p. \tag{7.22b}$$

Now let us partition $\hat{Q}$ as in (4.34):

$$\hat{Q} = \begin{bmatrix} \hat{Q}_0 & \hat{Q}_1 \\ \hat{Q}_1^* & \hat{Q}_2 \end{bmatrix} . \tag{7.23}$$

In the optimal control problem, we almost always have a nonzero $Q_0$ because this operator penalizes the generalized displacement. For the results in this paper, $\hat{Q}_0$ can be nonzero in the observer problem, and, as in the control problem, some of the strongest convergence results for finite dimensional approximations can be proved only for coercive $\hat{Q}$. However, if the observer is to be thought of as an optimal filter, then $\hat{Q}$ should be the covariance operator of the noise that disturbs (2.1). In this case, $\hat{Q}_0 = 0$ and $\hat{Q}_1 = 0$.

## 8. Approximation of the Infinite Dimensional Estimator

### 8.1 The Approximating Finite Dimensional Estimators

Here, the scheme for the approximation of the flexible structure is that in Section 4. We will construct on the subspace $E_n$ an estimator that approximates the optimal infinite dimensional estimator of Section 7, and this estimator will produce an $n^{th}$-order estimate $\hat{z}_n = (\hat{x}_n, \dot{\hat{x}}_n)$ of the infinite dimensional state vector $z = (x, \dot{x})$. In Section 9, the the $n^{th}$-order compensator that results from applying the $n^{th}$ approximation to the optimal control law (in Section 4.2) to $\hat{z}_n$ will approximate the optimal infinite dimensional compensator of Section 7.

Hypothesis 8.1. There exist a sequence $C_n \in L(E_n, R^p)$ such that

$$||C_n P_{En} - C|| \to 0 \quad \text{as} \quad n \to \infty \tag{8.1}$$

and a sequence $\hat{Q}_n \in L(E_n)$, $\hat{Q}_n^* = \hat{Q}_n \geq 0$, such that

$$\hat{Q}_n P_{En} \to \hat{Q} \quad \text{strongly as} \quad n \to \infty . \square \tag{8.2}$$

Hypothesis 8.2. For each n, the system $(A_n^*, C_n^*)$ is stabilizable. In particular, any unstable modes of the system $(C_n, A_n)$ are observable. $\square$

The $n^{th}$ observer, or $n^{th}$ estimator, is

$$\dot{\hat{z}}_n = A_n \hat{z}_n + B_n u + \hat{F}_n (y - C_0 u - C_n \hat{z}_n) \tag{8.3}$$

where the estimator gain $\hat{F}_n$ is

$$\hat{F}_n = \hat{\Pi}_n C_n^* \hat{R}^{-1} , \tag{8.4}$$

and $\hat{\Pi}_n$ is the nonnegative selfadjoint solution to the Riccati operator equation

$$A_n \hat{\Pi}_n + \hat{\Pi}_n A_n^* - \hat{\Pi}_n C_n^* \hat{R}^{-1} C_n \hat{\Pi}_n + \hat{Q}_n = 0. \tag{8.5}$$

Hypothesis 8.2 implies that such a solution exists and is unique.

This representation of the $n^{th}$ estimator as a system on $E_n$, with the estimator gain determined by the solution to a Riccati operator equation, is necessary for showing how the sequence of finite dimensional estimators approximate the infinite dimensional estimator. However, on-line computations will be based on the equivalent differential equation

$$\dot{\hat{\eta}} = A^n \hat{\eta} + B^n u + \hat{F}^n (y - C_0 u - C^n \hat{\eta}) \tag{8.6}$$

where $\hat{\eta}(t) \in R^{2n}$, $A^n$ and $B^n$ are the matrix representations of the operators $A_n$ and $B_n$, as in Section 4, and $C^n$ is the matrix representation of $C_n$. The $2n \times p$ gain matrix $\hat{F}^n$ is

$$\hat{F}^n = \hat{\Pi}^n W^{-n} (C^n)^T \hat{R}^{-1}, \tag{8.7}$$

where $W^n$ is the $2n \times 2n$ grammian matrix in (4.33) and $\hat{\Pi}^n$ satisfies

$$A^n \hat{\Pi}^n + \hat{\Pi}^n W^{-n} (A^n)^T W^n + \hat{\Pi}^n W^{-n} (C^n)^T \hat{R}^{-1} C^n \hat{\Pi}^n + \hat{Q}^n = 0, \tag{8.8}$$

with $\hat{Q}^n$ the matrix representation of $\hat{Q}_n$. The relationship between $\hat{z}_n = (\hat{x}_n, \dot{\hat{x}}_n)$ and $\hat{\eta}$ is, of course,

$$\hat{x}_n(t) = \sum_{i=1}^{n} \hat{\xi}_i(t) e_i \tag{8.9}$$

and

$$\hat{\eta} = [\hat{\xi}^T \quad \dot{\hat{\xi}}^T]^T. \tag{8.10}$$

Since the matrix representations of $A_n$ and $A_n^*$ are $A^n$ and $W^{-n}(A^n)^T W^n$, respectively, and the matrix representation of $C_n^*$ is $W^{-n}(C^n)^T$, (8.7) is the matrix representation of (8.4), and the $2n \times 2n$ Riccati matrix equation (8.8) is the matrix representation of (8.5), with $\hat{\Pi}^n$ the matrix representation of $\hat{\Pi}_n$. (Recall that $W^{-n}$ is the inverse of $W^n$.)

As in the control problem, we do not solve the matrix representation of the $n^{th}$ Riccati operator equation directly because the matrix representation of a selfadjoint operator in general is not symmetric. In the duality between the optimal control and estimator problems, (8.5) and (8.8) correspond to (4.26) and (4.38), respectively. In (4.39), we defined the symmetric matrix $\tilde{\Pi}^n = W^n \Pi^n$ and then obtained the Riccati equation (4.40) to solve for $\tilde{\Pi}^n$. We proceed in a similar fashion here, but with an interesting difference.

Since $\hat{\Pi}_n$ and $\hat{Q}_n$ are nonnegative selfadjoint operators on $E_n$ and $\hat{\Pi}^n$ and $\hat{Q}^n$ are their matrix representations, the matrices $W^n \hat{\Pi}^n$ and $W^n \hat{Q}^n$ are nonnegative and symmetric. Hence, the matrices

$$\tilde{\hat{\Pi}}^n = \hat{\Pi}^n W^{-n} \tag{8.11}$$

and

$$\tilde{\hat{Q}}^n = \hat{Q}^n W^{-n} \tag{8.12}$$

are nonnegative and symmetric. Substituting (8.11) and (8.12) into (8.7) and (8.8) yields

$$\hat{F}^n = \tilde{\hat{\Pi}}^n (C^n)^T \hat{R}^{-1} \tag{8.13}$$

and

$$A^n \tilde{\hat{\Pi}}^n + \tilde{\hat{\Pi}}^n (A^n)^T - \tilde{\hat{\Pi}}^n (C^n)^T \hat{R}^{-1} C^n \tilde{\hat{\Pi}}^n + \tilde{\hat{Q}}^n = 0, \tag{8.14}$$

the Riccati matrix equation to be solved numerically in the $n^{th}$ approximation to the infinite dimensional estimator. In view of the relationship between (8.5) and (8.8) and the relationship between (8.8) and (8.14), we see that Hypothesis 8.1 guarantees the existence of a unique nonnegative symmetric solution to (8.14).

To see the relationship between the matrices in (8.14) and the operators in (8.5) more clearly -- and the difference between the current approximation scheme and that used in Section 4.2 for the control problem -- suppose that we take $\hat{Q}_n = P_{En} \hat{Q}|_{En}$. Let $\tilde{\tilde{Q}}^n$ be defined as in (4.36) and (4.37) with $Q_0$, $Q_1$, and $Q_2$ replaced by $\hat{Q}_0$, $\hat{Q}_1$, and $\hat{Q}_2$. Then

$$\tilde{\hat{Q}}^n = W^{-n} \tilde{\tilde{Q}}^n W^{-n}. \tag{8.15}$$

For example, if Q in the control problem and $\hat{Q}$ in the estimator problem are both equal to the identity, then the $\tilde{Q}^n$ in (4.35) - (4.42) is $W^n$ and $\tilde{\tilde{Q}}^n = W^{-n}$. This may seem suspicious, but Subsection 8.2 should demonstrate that we are solving the appropriate estimator problem here.

The only thing missing now for numerical implementation of the $n^{th}$

estimator, or observer, is to give $C^n$, the matrix representation of $C_n$, explicitly. We write

$$C^n = [C_1^n \quad C_2^n] \tag{8.16}$$

where the $p \times n$ matrices $C_1^n$ and $C_2^n$ are, respectively, the matrix representations of the operators $C_1$ and $C_2$ in (7.18). We can cover virtually all applications by assuming $C_n = C|_{En}$, in which case the $i^{th}$ column of $C_1^n$ is the p-vector equal to $C_1 e_i$, and the $i^{th}$ column of $C_2^n$ is the p-vector equal to $C_2 e_i$.

We now have the complete set of equations for numerical implementation of the $n^{th}$ state estimator: For online computation, the $n^{th}$ estimator, or observer, is (8.6); the gain matrix $\hat{F}^n$ is given by (8.13) and the solution to the Riccati matrix equation (8.14). The matrices $\hat{Q}^n$ and $C^n$ are defined as above.

## 8.2 Stochastic Interpretation of the Approximating Estimators

As we have said, our approximation theory for the optimal estimator is based on approximation of the infinite dimensional Riccati equation, whose structure is the same for both control and estimator problems, and the stochastic properties of the optimal estimator problem never enter our approximation theory. Furthermore, using only the deterministic setting above, we will proceed, subsequently, to analyze the finite dimensional estimators and the compensators based upon them. Nontheless, we should consider momentarily the sequence of finite dimensional stochastic estimation problems whose solutions are given by the equations of the preceding subsection.

First, recall how the covariance operator of a Hilbert space-valued ran-

dom variable is defined. The covariance operator of an E-valued random variable $\omega$ is the operator Q for which

$$\text{expected value } \{\langle z,\omega\rangle_E \langle \tilde{z},\omega\rangle_E\} = \langle Qz,\tilde{z}\rangle_E, \quad z, \tilde{z} \in E. \tag{8.17}$$

(See [B1, C4].)

With $\hat{F}_n$ given by (8.4) and (8.5), (8.3) is the Kalman-Bucy filter for the system

$$\dot{z}_n = A_n z_n + B_n u + \omega_n, \tag{8.18}$$

$$y = C_0 u + C_n z_n + \omega_0, \tag{8.19}$$

where $\omega_n(t)$ is an $E_n$-valued white noise process with covariance operator $\hat{Q}_n$ and $\omega_0(t)$ is an $R^p$-valued white noise process with covariance operator (matrix) $\hat{R}$. Next, careful inspection will show that the filter defined by (8.6), (8.13) and (8.14) is the matrix representation of the filter defined by (8.3), (8.4) and (8.5).

With $z_n$ and $\eta$ related as in (4.1) and (4.4), (8.18) and (8.19) are equivalent to the system

$$\dot{\eta} = A^n \eta + B^n u + \gamma, \tag{8.20}$$

$$y = C_0 u + C^n \eta + \omega_0, \tag{8.21}$$

where $\gamma(t)$ is the $R^{2n}$-valued noise process related to $\omega_n(t)$ by

$$\omega_n(t) = \sum_{i=1}^{n} (\mathit{V}_i(t)e_i,\ \mathit{V}_{i+n}(t)e_i). \tag{8.22}$$

Certainly, a Kalman-Bucy filter for (8.20) and (8.21) has the form (8.6) with the filter gain given by (8.13) and (8.14). This particular filter is the matrix representation of the filter defined by (8.3), (8.4) and (8.5) if and only if the matrix $\tilde{\hat{Q}}^n$ defined by (8.12) is the covariance of the process $\mathit{V}(t)$. Since $\hat{Q}^n$ is the matrix representation of $\hat{Q}_n$, straightforward calculation using (8.12) and (8.17) shows that the $\tilde{\hat{Q}}^n$ in (8.12) is indeed the correct covariance matrix.

Of course, if $\omega_n(t)$ and $\mathit{V}(t)$ represent a physical disturbance to the structure, then $\omega_n(t)$ must have the form $(0,\omega_n^{(2)}(t))$ and the first n elements of $\mathit{V}(t)$ must be zero, but this is not necessary for our analysis.

The finite dimensional observers can be interpreted now as a sequence of filters designed for the sequence of finite dimensional approximations to the flexible structure, with the $n^{th}$ approximate system disturbed by the noise process $\omega_n(t)$, whose covariance operator is $\hat{Q}_n$. By Hypothesis 8.1, these covariance operators converge to the operator $\hat{Q}$ of Section 7. If we have a reliable model of a stationary, zero-mean gaussian disturbance for the structure, then we can take the covariance operator for this disturbance to be $\hat{Q}$ and think of the infinite dimensional observer as the optimal estimator. But, again, this interpretation is not necessary for the rest of our analysis.

## 8.3 The Approximating Functional Estimator Gains

The $n^{th}$ estimator gain operator in (8.4) has the same form as the infinite dimensional estimator gain in (7.15) and (7.20). We have

$$\hat{F}_n y = \sum_{i=1}^{p} (\hat{f}_{in}, \hat{g}_{in}) y_i \tag{8.23}$$

for $y = [y_1\ y_2\ \ldots\ y_p]^T \in R^p$, where the functional estimator gains $\hat{f}_{in}$ and $\hat{g}_{in}$ are elements of $V_n = H_n$. The matrix $\hat{F}^n$ in (8.7) and (8.13) is the matrix representation of $\hat{F}_n$, which means that, if we write

$$\hat{F}^n = \begin{bmatrix} {}^{f_1}\beta & {}^{f_2}\beta & \ldots & {}^{f_p}\beta \\ {}^{g_1}\beta & {}^{g_2}\beta & \ldots & {}^{g_p}\beta \end{bmatrix} \tag{8.24}$$

where the columns ${}^{f_i}\beta,\ {}^{g_i}\beta \in R^n$, then

$$\hat{f}_{in} = \sum_{j=1}^{n} {}^{f_i}\beta_j e_j, \qquad i = 1,\ldots,p, \tag{8.25a}$$

$$\hat{g}_{in} = \sum_{j=1}^{n} {}^{g_i}\beta_j e_j, \qquad i = 1,\ldots,p. \tag{8.25b}$$

For convergence analysis, it is useful to note that $\hat{f}_{in}$ and $\hat{g}_{in}$ are given also by equations corresponding to (7.22). With the measurement operator C written as in (7.19) and $C_n = C|_{E_n}$, we have

$$\hat{f}_{in} = \sum_{j=1}^{p} (\hat{R}^{-1})_{ij} (\hat{\Pi}_{0n} P_{Vn} c_{ij} + \hat{\Pi}_{1n} P_{Hn} c_{2j}), \tag{8.26a}$$

$$\hat{g}_{in} = \sum_{j=1}^{p} (\hat{R}^{-1})_{ij} (\hat{\Pi}^{*}_{1n} P_{Vn} c_{ij} + \hat{\Pi}_{2n} P_{Hn} c_{2j}), \tag{8.26b}$$

where

$$\hat{\Pi}_n = \begin{bmatrix} \hat{\Pi}_{0n} & \hat{\Pi}_{1n} \\ \hat{\Pi}^{*}_{1n} & \hat{\Pi}_{2n} \end{bmatrix}. \tag{8.27}$$

## 8.4 Convergence

Now we will indicate the sense in which the finite dimensional estimators/observers approximate the infinite dimensional estimator in Section 7. As we have said, implementation of the $n^{th}$ estimator is based on (8.6), (8.13) and (8.14), but convergence analysis is based on the equivalent system (8.3), (8.4) and (8.5). The question then is how the observer in (8.3), with gain given by (8.4) and (8.5), converges to the observer in (7.3) with gain given by (7.15) and (7.16).

Recall Hypothesis 8.1, and recall from Section 4 that the approximations to both the open-loop semigroup and its adjoint converge strongly. Also, recall that Hypothesis 8.2 guarantees a unique nonnegative selfadjoint solution to the Riccati equation (8.5) for each n. Replacing $A_n$ and $B_n$ with $A_n^*$ and $C_n^*$ in Theorems 5.1 and 5.3, we obtain

Theorem 8.3. i) If $\|\hat{\Pi}_n\|$ is bounded uniformly in n, then the Riccati algebraic equation (7.16) has a nonnegative selfadjoint solution $\hat{\Pi}$ and $\hat{\Pi}_n P_{En}$ converges weakly to $\hat{\Pi}$. ii) If there exist positive constants M and $\beta$, independent of n, such that

$$\|\exp([A_n - \hat{\Pi}_n C_n^* R^{-1} C_n]t)\| \leq Me^{-\beta t}, \quad t \geq 0, \tag{8.28}$$

then $\|\hat{\Pi}_n\|$ is bounded uniformly in n, $\hat{\Pi}_n P_{En}$ converges strongly to $\hat{\Pi}$ and $\exp([A_n - \hat{\Pi}_n C_n^* \hat{R}^{-1} C_n]t)P_{En}$ converges strongly to $\hat{S}(t)$, the semigroup generated by $A - \hat{\Pi} C^* \hat{R}^{-1} C$, the convergence uniform in $t \geq 0$. iii) If $\hat{Q}_n$ is bounded away from zero uniformly in n, then $\|\hat{\Pi}_n\|$ being bounded uniformly in n guarantees the existence of positive constants M and $\beta$ for which (8.28) holds for all n. □

The proof of the following theorem is practically identical to that of Theorem 5.4.

Theorem 8.4. If $\hat{Q}$ is E-coercive and $d_0 = 0$, then there is no nonnegative selfadjoint solution of the Riccati operator equation (7.16), and

$$\|\hat{\Pi}_n\| \to \infty \quad \text{as} \quad n \to \infty. \quad \square \tag{8.29}$$

Our purpose for bothering to state this obvious dual result is to point out the following question. Can Theorem 8.4 be modified to include the case where Q has the form (7.23) with $\hat{Q}_0 = 0$, $\hat{Q}_1 = 0$ and $\hat{Q}_2$ coercive on H?

Next, we have the dual result to Theorem 5.6:

Theorem 8.5 Suppose that $A_0$ has an invariant subspace $V_0$ which is also invariant under the damping map $\Lambda_V^{-1}\Lambda_D$, that $E_0 = V_0 \times V_0$ is an observable subspace, and that the restrictions of $A_0$ and $d_0(\cdot,\cdot)$ to $V_0^{\perp}$ are both H-coercive. Also, suppose that $V_0$ has finite dimension $n_0$ and that, for each $n \geq n_0$ in the approximation scheme, the first $n_0$ $e_i$'s span $V_0$ and the rest are orthogonal to $V_0$ in both V and H.

i) Then (7.16) has nonnegative solution $\hat{\Pi}$, and $\|\hat{\Pi}_n\|$ is bounded uniformly in n, so that $\hat{\Pi}_n P_{En}$ converges to $\hat{\Pi}$ weakly.

ii) If $E_0$ and $E_0^{\perp}$ (the E-orthogonal complement of $E_0$) are invariant under $\hat{Q}$, and if the $E_0$-part of the system $(A,\hat{Q})$ is controllable, then the hypothesis of Theorem 8.3 ii) holds.

Proof. The proof is practically identical to that of Theorem 5.6 with B replaced by $C^*$. For ii), note that, when we partition A and $\hat{Q}$ as in (3.16),

the finite dimensional system $(A_{11}, \hat{Q}_{11})$ is controllable if and only if the system $(\hat{Q}_{11}, A_{11}^*)$ is observable. □

Remark 8.6. Remarks 5.7 and 5.8 pertain to Theorem 8.5 as well as to Theorem 5.6; i.e., in most applications the theorem requires either that both $A_0$ and $d_0$ be coercive (so that $V_0 = \{0\}$) or that the natural mode shapes be the basis vectors and the damping not couple the natural modes. It seems unlikely that a finite number of observable rigid-body modes could change the nature of the convergence, but they greatly complicate the proofs. For applications where both rigid-body displacement and rigid-body velocity are measured, a result analogous to Theorem 5.9 can be obtained, but we will not bother here because it adds no significant insight and we cannot use it in the example in Sections 6 and 10. Also, see Remark 10.1. □

Theorem 8.7. If $\hat{\Pi}_n P_{En}$ converges strongly to $\hat{\Pi}$, then

$$\|\hat{f}_{in} - \hat{f}_i\|_V \to 0 \tag{8.30a}$$

$$\|\hat{g}_{in} - \hat{g}_i\|_H \to 0, \quad \text{as} \quad m \to \infty, \tag{8.30b}$$

where $\hat{f}_i$ and $\hat{g}_i$ are the functional estimator gains in (7.20) and $\hat{f}_{in}$ and $\hat{g}_{in}$ are the approximating functional gains in (8.25).

Proof. The result follows from (7.22) and (8.26). □

# 9. The Finite Dimensional Compensators and Realizable Closed-loop Systems

## 9.1 Closing the Loop

The $n^{th}$ compensator consists of the $n^{th}$ approximation to the optimal control law in Section 4, applied to the output of the $n^{th}$ estimator/observer in Section 8; i.e., the feedback control

$$u_n = - F_n \hat{z}_n \tag{9.1}$$

where

$$F_n = R^{-1} B_n^* \Pi_n \tag{9.2}$$

(recall (4.25)) and $\hat{z}_n(t)$ is the solution to (8.3). Equivalently, this compensator can be written as

$$u_n = - F^n \hat{\eta} \tag{9.3}$$

where

$$F^n = R^{-1} B^{n\,T} \tilde{\Pi}^n \tag{9.4}$$

(recall (4.43)) and the $2n$-vector $\hat{\eta}(t)$ is the solution to (8.6)). On-line computations will be based upon the latter representation, and the block diagram in Figure 9.1 shows the realizable closed-loop system that results from the $n^{th}$ compensator. We will refer to this system as the $n^{th}$ closed-loop system.

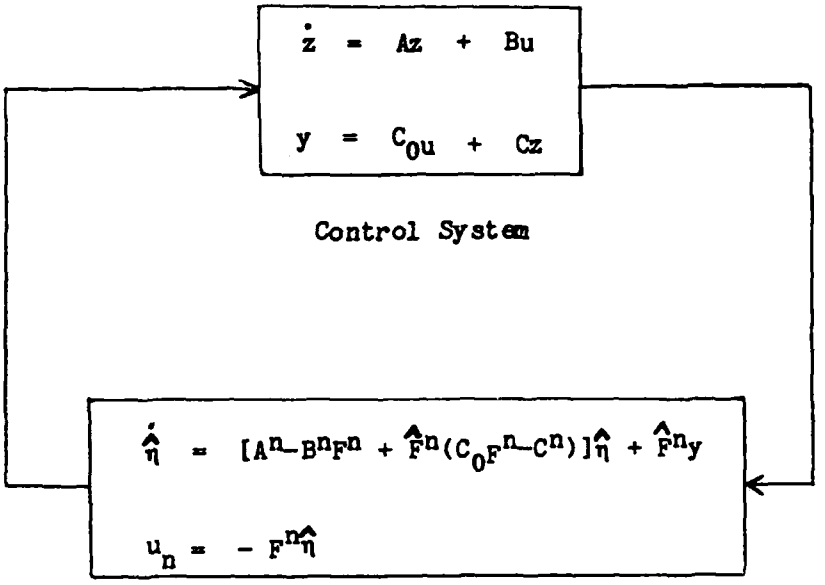



Figure 9.1 $n^{th}$ Closed-loop System

This closed-loop system is equivalent to

$$\begin{pmatrix} \dot{z} \\ \dot{\hat{z}}_n \end{pmatrix} = A_{\infty,n} \begin{pmatrix} z \\ \hat{z}_n \end{pmatrix}, \tag{9.5}$$

where the operator

$$A_{\infty,n} = \begin{bmatrix} A & -BF_n \\ \hat{F}_nC & [A_n - B_nF_n + \hat{F}_n C_n] \end{bmatrix}, \quad D(A_{\infty,n}) = D(A) \times E_n, \tag{9.6}$$

generates the closed-loop semigroup $S_{\infty,n}(t)$ on $E \times E_n$. The closed-loop response produced by the $n^{th}$ compensator -- i.e., the response of the $n^{th}$ closed-loop system -- can be written then as

$$\begin{pmatrix} z(t) \\ \hat{z}_n(t) \end{pmatrix} = S_{\infty,n}(t) \begin{pmatrix} z(0) \\ \hat{z}_n(0) \end{pmatrix}. \tag{9.7}$$

Note that $A_{\infty,n}$ has compact resolvent if and only if A does.

## 9.2 Convergence of the Closed-loop Systems

Now we will consider the sense in which the $n^{th}$ closed-loop system approximates the optimal closed-loop system in Section 7 (Definition 7.3). Recall from Sections 4.1 and 8.1 how the approximating open-loop semigroups $T_n(\cdot)$ and their adjoints converge strongly and how the input operators $B_n$, the measurement operators $C_n$ and their respective adjoints converge in norm. Sections 5 and 8 have given sufficient conditions for the approximating control and estimator gains to converge to the gains for the optimal infinite dimensional compensator. In this section, we will assume

Hypothesis 9.1. As $n \longrightarrow \infty$,

$$||F_n P_{En} - F|| \to 0, \tag{9.8}$$

$$||\hat{F}_n - \hat{F}|| \to 0. \quad \square \tag{9.9}$$

Remark 9.2. Of course, we are interested primarily in the case where the gains F and $\hat{F}$ are the optimal LQG gains in (7.15) and (7.17) and $F_n$ and $\hat{F}_n$ are the corresponding approximations in Sections 4 and 8 (i.e.,(9.2) and (8.4)). However, for the analysis of this section, we need only Hypothesis 9.1 for some $F \in L(E,R^m)$, $\hat{F} \in L(R^p,E)$ and approximating sequences $F_n$ and $\hat{F}_n$. Any such gain operators will yield closed-loop semigroup generators $A_{\infty,\infty}$ in (7.9) and $A_{\infty,n}$ in (9.6). $\square$

We denote the projection of $E \times E$ onto $E \times E_n$ by $P_{EEn}$.

Theorem 9.3. For $t \geq 0$, $S_{\infty,n}(t)P_{EEn}$ converges strongly to $S_{\infty,\infty}(t)$, and the convergence is uniform in t for t in bounded intervals.

Proof. This follows from the strong convergence of the open-loop semigroups and the uniform norm convergence of the control and estimator gains. □

We should expect at least Theorem 9.3, but we need more. We should require, for example, that if S(t) is uniformly exponentially stable, then $S_{\infty,n}(t)$ must be also for n sufficiently large. Although numerical results for numerous examples with various kinds of damping and approximations suggest that this is usually true, we have been unable to prove it in general. We do have the result for the following important case.

Suppose that the basis vectors $e_j$ of the approximation scheme are the natural modes of undamped free vibration and that the structural damping does not couple the modes. Then, for each each n, $E_n$ and $E_n^{\perp}$ reduce the open-loop semigroup T(t) and its generator A. For this case, we can extend $A_{\infty,n}$ to $D(A) \times D(A)$ as

$$\tilde{A}_{\infty,n} = \begin{bmatrix} A & -BF_nP_{En} \\ \hat{F}_nC & \tilde{A}_{nComp} \end{bmatrix} \tag{9.10}$$

where

$$\tilde{A}_{nComp} = [A_n - B_nF_n - \hat{F}_nC_n]P_{En} + A|_{D(A) \cap E_n^{\perp}}. \tag{9.11}$$

Note that $E_n^{\perp}$ is the span of the modes not represented in the $n^{th}$ compensator.

The operator $\tilde{A}_{\infty,n}$ generates a semigroup $\tilde{S}_{\infty,n}(t)$ on $E \times E$, $E \times E_n$ and

$\{0\}\times E_n^{\perp}$ reduce $\tilde{S}_{\infty,n}(t)$, and the restriction of $\tilde{S}_{\infty,n}(t)$ to $E \times E_n$ is $S_{\infty,n}(t)$. Hence $\tilde{S}_{\infty,n}(t)$ is uniformly exponentially stable if and only if both $S_{\infty,n}(t)$ and the part of the open-loop system on $E_n^{\perp}$ are uniformly exponentially stable.

Theorem 9.4. i) Suppose that the basis vectors of the approximation scheme are the natural modes of undamped free vibration and that the structural damping does not couple the modes. Then $\tilde{S}_{\infty,n}(t)$ converges in norm to $S_{\infty,\infty}(t)$, uniformly in bounded t-intervals.

ii) If, additionally, $S_{\infty,\infty}(t)$ is uniformly exponentially stable, then $S_{\infty,n}(t)$ is uniformly exponentially stable for n sufficiently large.

Proof. From (9.6), (9.10) and (9.11), we have

$$A_{\infty,\infty}-\tilde{A}_{\infty,n} = \begin{bmatrix} 0 & B[F_n P_{En}-F] \\ [\hat{F}-\hat{F}_n]C & \Delta_n \end{bmatrix} \tag{9.12}$$

where

$$\Delta_n = (B_n F_n P_{En}-BF) + (\hat{F}_n C_n P_{En}-\hat{F}C). \tag{9.13}$$

Therefore, $\|A_{\infty,\infty}-\tilde{A}_{\infty,n}\| \dashrightarrow 0$ as $n \dashrightarrow \infty$, and the theorem follows. □

This paper emphasizes using the convergence of the approximating control and estimator gain operators $F_n$ and $\hat{F}_n$, and the convergence of the functional gains that can be used to represent these operators, to determine the finite dimensional compensator that will produce essentially optimal closed-loop performance. However, close examination of the right sides of (9.12) and (9.13) reveals another important convergence question. While the gain convergence in (9.8) and (9.9) drives the off-diagonal blocks in (9.12) to zero, the norm convergence of the approximating input and output operators also is essential

in killing $\Delta_n$. Expanding the two terms in this block yields

$$B_n F_n P_{En} - BF = B_n(F_n P_{En} - F) + (B_n - B)F, \tag{9.14}$$

$$\hat{F}_n C_n P_{En} - \hat{F}C = (\hat{F}_n - \hat{F})C_n P_{En} + \hat{F}(C_n P_{En} - C). \tag{9.15}$$

The second term on the right side of each of these equations represents, respectively, control and observation spillover, which has been studied extensively by Balas. Together, the control spillover and observation spillover couple the modes modelled in the compensator with the modes not modelled in the compensator. The spillover must go to zero -- as it does when $B_n$ and $C_n$ converge -- for $A_{\infty,\infty} - \tilde{A}_{\infty,n}$ to go to zero.

We should ask then whether there exists a correlation between the convergence of $F_n$ and $\hat{F}_n$ and the elimination of spillover. The answer is yes if no modes lie in the null space of the state weighting operator $Q$ in the performance index and if the assumed process noise, whose covariance operator is $\hat{Q}$, excites all modes, but this correlation is difficult to quantify. As we discussed in Section 6.4, the two main factors that determine the convergence rates of the gains are the Q-to-R ratio and the damping, neither of which affects the convergence of $B_n$ and $C_n$. On the other hand, when either factor (small Q/R or large damping) causes the gains to converge fast, it generally also causes the magnitude of $F$ and $\hat{F}$ to be relatively small, thereby reducing the magnitude of the spillover terms in (9.14) and (9.15). Also, as $n$ increases, the increasing frequencies of the truncated modes usually reduce the coupling effect of spillover. This is well known, although it cannot be seen from the equations here. In examples that we have worked, we have found that when $n$ is large enough to produce convergence of the control and estimator gains, the effect of any remaining spillover is negligible. But this may not always be true, and spillover should be remembered.

## 9.3 Convergence of the Compensator Transfer Functions

The transfer function of the $n^{th}$ compensator (shown in the bottom block of Figure 9.1) is

$$\Phi_n(s) = -F_n(sI - [A_n - B_nF_n + \hat{F}_n(C_0F_n - C_n)])^{-1}\hat{F}_n, \tag{9.16}$$

which is an $m \times p$ matrix function of the complex variable $s$ for each $n$, as is the similar transfer function $\Phi(s)$ in (7.14) for the infinite dimensional compensator. We continue to assume Hypothesis 9.1.

We will denote the resolvent set of $[A-BF+\hat{F}(C_0F-C)]$ by $\rho([A-BF+\hat{F}(C_0F-C)])$.

Theorem 9.5. There exists a real number $a_1$ such that, if $Re(s) > a_1$, then $s \in \rho([A_n-B_n F_n+\hat{F}_n(C_0F_n-C_n)])$ for all $n$, and $\Phi_n(s)$ converges to $\Phi(s)$, uniformly in compact subsets of such $s$.

Proof. The operator $[A-BF+\hat{F}(C_0F-C)]$ is obtained from a contraction semigroup generator by perturbation with bounded operators, and the approximations to the perturbing operators are bounded in $n$, by strong convergence. In view of this, close examination of the basic approximation scheme in Section 4.1 will show that there exists a bound of the form $M_1\exp(a_1t)$, independent of $n$, for the semigroups generated by $[A_n-B_n F_n+\hat{F}_n(C_0F_n-C_n)]$. Also, these semigroups converge strongly to the semigroup generated by $[A-BF+\hat{F}(C_0F-C)]$, according to [G3, Theorem 6.6]. For $Re(s) > a_1$ then, the resolvent operator in $\Phi_n(s)$ converges strongly to that in $\Phi(s)$, uniformly in compact $s$-subsets, by [K1, page 504, Theorem 2.16, and page 427, Theorem 1.2]. □

This result leaves much to be desired. For example, it does not guarantee that any subset of the imaginary axis will lie in $\rho([A_n-B_nF_n+\hat{F}_n(C_0F_n- C_n)])$ for sufficiently large n, even if all of the imaginary axis lies in $\rho([A-BF+\hat{F}(C_0F-C)])$. As with the convergence of the closed-loop systems, we can get more for certain important cases.

Remark 9.6. If the open-loop semigroup $T(\cdot)$ (whose generator is A) is an analytic semigroup, then there exist real numbers a, $\theta$ and M, with $\theta$ and M positive, such that $\rho([A-BF+\hat{F}(C_0F-C)])$ contains the sector $\{s:|\arg(s-a)|< \frac{\pi}{2} + \theta\}$, and for each s in this sector,

$$\|(sI-[A-BF+\hat{F}(C_0F-C)])^{-1}\| \leq M/|s-a| \tag{9.17}$$

Theorem 9.7. i) If the basis vectors of the approximation scheme are the natural modes of undamped free vibration and the structural damping does not couple the modes, then each s in $\rho([A-BF+\hat{F}(C_0F-C)])$ is in $\rho([A_n-B_nF_n+\hat{F}_n(C_0F_n-C_n)])$ for n sufficiently large and $\Phi_n(s)$ converges to $\Phi(s)$ as $n \longrightarrow \infty$, uniformly in compact subsets of $\rho([A-BF+\hat{F}(C_0F-C)])$. ii) If, additionally, $T(\cdot)$ is an analytic semigroup, then $\Phi_n(s)$ converges to $\Phi(s)$ uniformly in the sector described in Remark 9.6.

Proof. i) In this case, we have also

$$\Phi_n(s) = F_nP_{En}(sI - \tilde{A}_{nComp})^{-1}\hat{F}_n, \tag{9.18}$$

where $\tilde{A}_{nComp}$ is the operator on D(A) defined by (9.11). The result follows from (9.8) and (9.9) and the fact that $\tilde{A}_{nComp}$ converges in norm to $[A-BF+\hat{F}(C_0F-C)]$. ii) The result follows from i) and a bound on $(sI-\tilde{A}_{nComp})^{-1}$ for large |s| that is obtained from the Neumann series in view

of (9.17) and the uniform-norm convergence of $\tilde{A}_{nComp}$.

Theorem 9.8. If A has compact resolvent, then $\Phi_n(s)$ converges to $\Phi(s)$ for each $s \in \rho([A-BF+\hat{F}(C_0F-C)])$, uniformly in compact subsets.

Proof. As a result of Theorem 4.4, the resolvent operator in $\Phi_n(s)$ converges in norm to the resolvent operator in $\Phi(s)$ for sufficiently large real s. After an artificial extension of $A_n$ to $E_n$ then, the present theorem follows from [K1, pages 206-207, Theorem 2.25].

## 10. Closing the Loop in the Example

As in Definition 7.3, the optimal closed-loop system is formed with the optimal infinite dimensional compensator, which consists of the optimal control law for the distributed model of the structure applied to the output of an optimal infinite dimensional state estimator. This optimal control law is the limit of the approximating finite dimensional control laws in Section 6. In this section, we first approximate the infinite dimensional estimator, as in Section 8, and then apply the approximating control laws in Section 6 to the approximating finite dimensional estimators to produce a sequence of finite dimensional compensators that approximate the optimal compensator.

### 10.1 The Estimator Problem

We assume that the only measurement is the rigid-body angle $\theta$ and that this measurement has zero-mean Gaussian white noise with variance $\hat{R} = 10^{-4}$. We model the process noise as a zero-mean Gaussian white disturbance that has a component distributed uniformly over the beam, as well as two concentrated components that exert a force on the tip mass and a moment on the hub. For this disturbance, the covariance operator $\hat{Q}$ has the form (7.23) with $\hat{Q}_0 = 0$, $\hat{Q}_1 = 0$ and $\hat{Q}_2 = I$.

We construct the approximating estimators as in Section 8.1. The gain for the $n^{th}$ estimator is given by (8.13) with the solution to the Riccati matrix equation (8.14). For the rigid-body measurement, the matrix $C^n$ is

$$C^n = [1\ 0\ 0\ 0 \ldots]. \tag{10.1}$$

According to (8.15), the matrix $\tilde{\tilde{Q}}^n$ is

$$\tilde{\tilde{Q}}^n = \begin{bmatrix} 0_* & 0_* \\ 0 & M^{-n} \end{bmatrix} , \tag{10.2}$$

since $W^n$ is the matrix in (4.33). (As always, $M^{-n}$ is the inverse of the mass matrix.) Recall from Section 6.3 that $n = 2n_e + 1$ where $n_e$ is the number of elements.

Our only use for the functional estimator gains is to measure the convergence of the finite dimensional estimators to the optimal infinite dimensional estimator. To see the convergence of the approximating estimator gains, we compute the approximating functional estimator gains as in Section 8.3. Like the functional control gains, the functional estimator gains have the form

$$\hat{f} = (\alpha_f, \phi_f, \beta_f) , \tag{10.3a}$$

$$\hat{g} = (\alpha_g, \phi_g, \beta_g) , \tag{10.3b}$$

and the corresponding approximations have the form

$$\hat{f}_n = (\alpha_{fn}, \phi_{gn}, \beta_{gn}) , \tag{10.4a}$$

$$\hat{g}_n = (\alpha_{gn}, \phi_{gn}, \beta_{gn}) . \tag{10.4b}$$

Remark 10.1 We cannot guarantee as much about convergence for the approximating estimators as we could for the approximating control problems in Section 6. Since the damping in this example does not couple the natural modes and the rigid-body mode is observable, we would have Part i) of Theorem 8.5 if we were using the natural mode shapes as basis vectors. Therefore, we know at

least that a solution to the infinite dimensional Riccati equation (7.16) exists and that the infinite dimensional estimator that we want to approximate exists. The numerical results indicate that the solutions to the finite dimensional Riccati equations are bounded in n and that the functional estimator gains converge in norm. The rigid-body mode prevents our guaranteeing a priori all the convergence that we want. If a torsional spring and damper were attached to the hub in the current example, we would have coercive stiffness and damping and Theorem 8.5 ii) would guarantee that the solutions to the finite dimensional Riccati equations converge strongly and that the functional estimator gains converge in norm for the basis vectors used here. Also, see Remark 6.2 and Remark 8.6. □

For damping coefficient $c_0 = 10^{-4}$, Figures 10.1 and 10.2 show $\phi''_{fn}$ and $\phi_{gn}$, and Tables 10.1 and 10.2 list the the scalars $\alpha_{fn}, \alpha_{gn}$ and $\beta_{gn}$. Since $\phi_{fn}(0) = \phi'_{fn}(0) = 0$, the convergence of $\phi''_{fn}$ implies the convergence of $\beta_{fn} = \phi_{fn}(\ell)$; as in the control problem, $\beta_{fn}$ is not an independent piece of information about the estimator gains while, as far as our results go, $\beta_{gn}$ is. We maintain analogy with the control problem and list only $\beta_{gn}$ in the table.

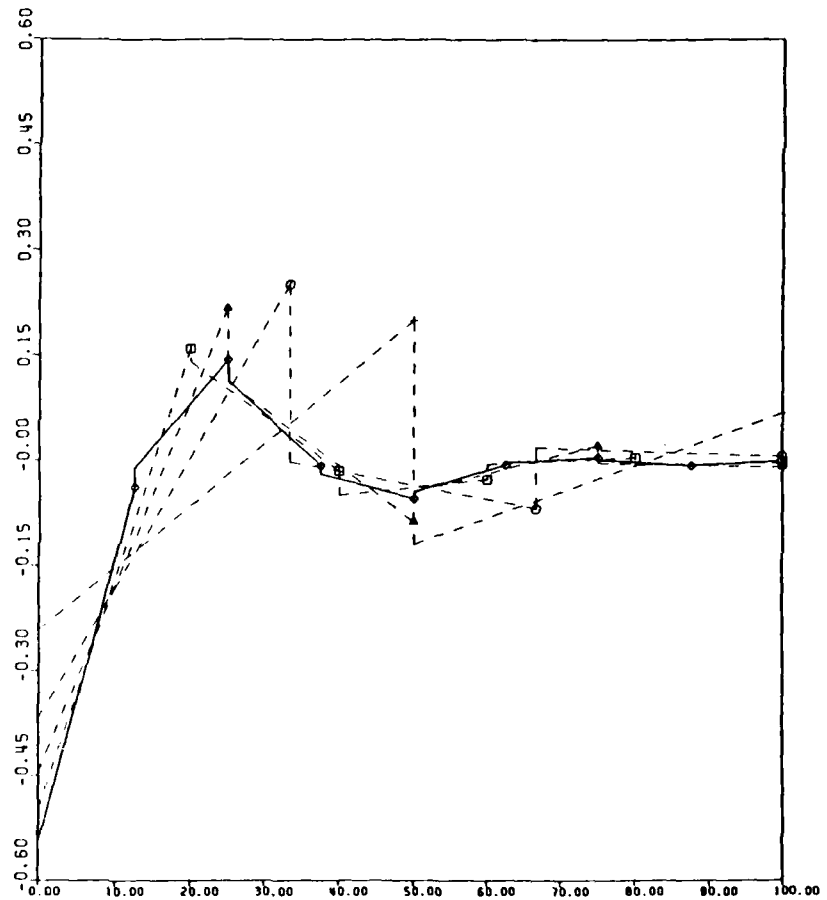



Figure 10.1a. Functional Estimator Gain Component $\phi''_{fn}$

Damping $c_0 = 10^{-4}$; estimator R = $10^{-4}$

number of elements $n_e$ = 2,3,4,5,8

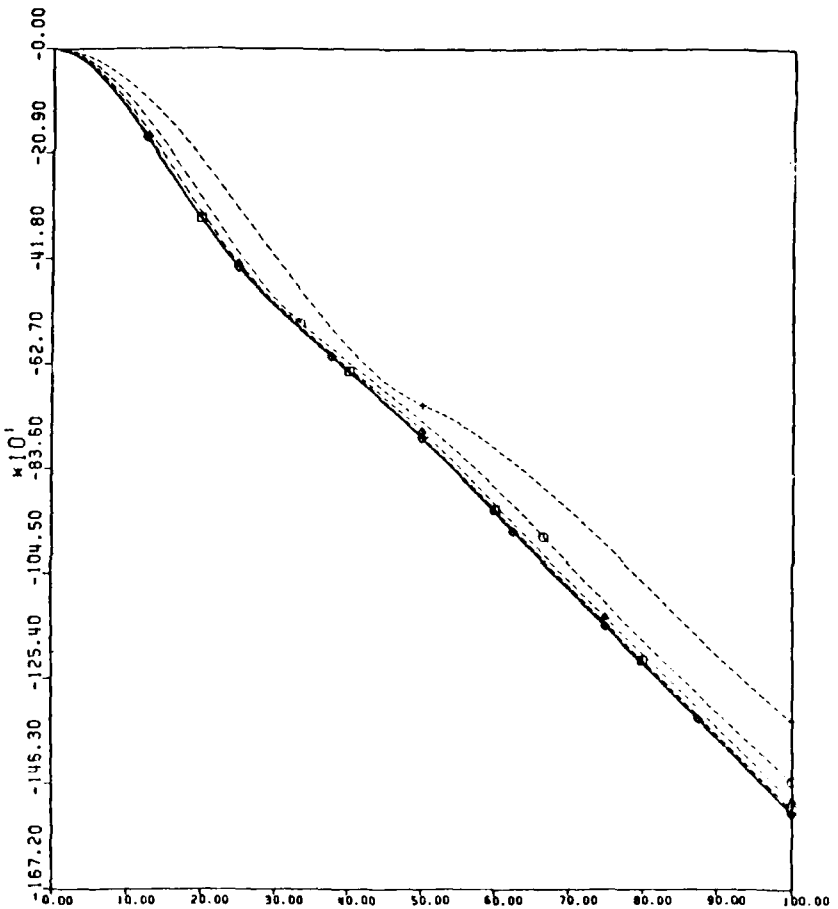



Figure 10.1b. Functional Estimator Gain Component $\phi_{gn}$

Damping $c_0 = 10^{-4}$; estimator R = $10^{-4}$

number of element $n_e$ = 2, 3, 4, 5, 8

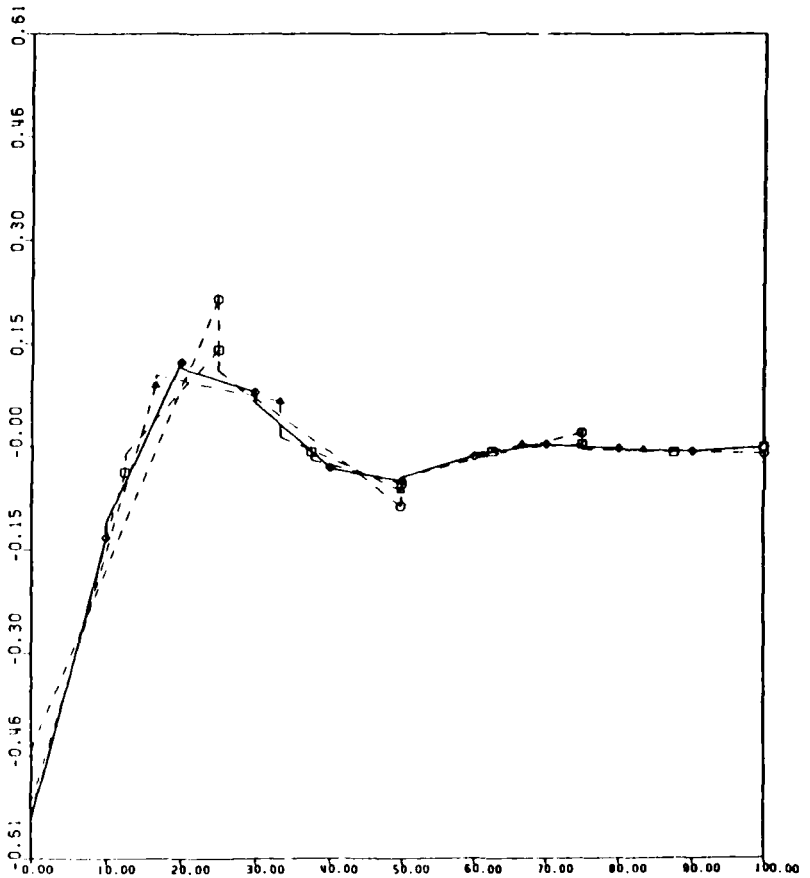



Figure 10.2a. Functional Estimator Gain Component $\phi''_{fn}$

Damping $c_0 = 10^{-4}$; estimator $R = 10^{-4}$

number of elements $n_e = 4, 6, 8, 10$

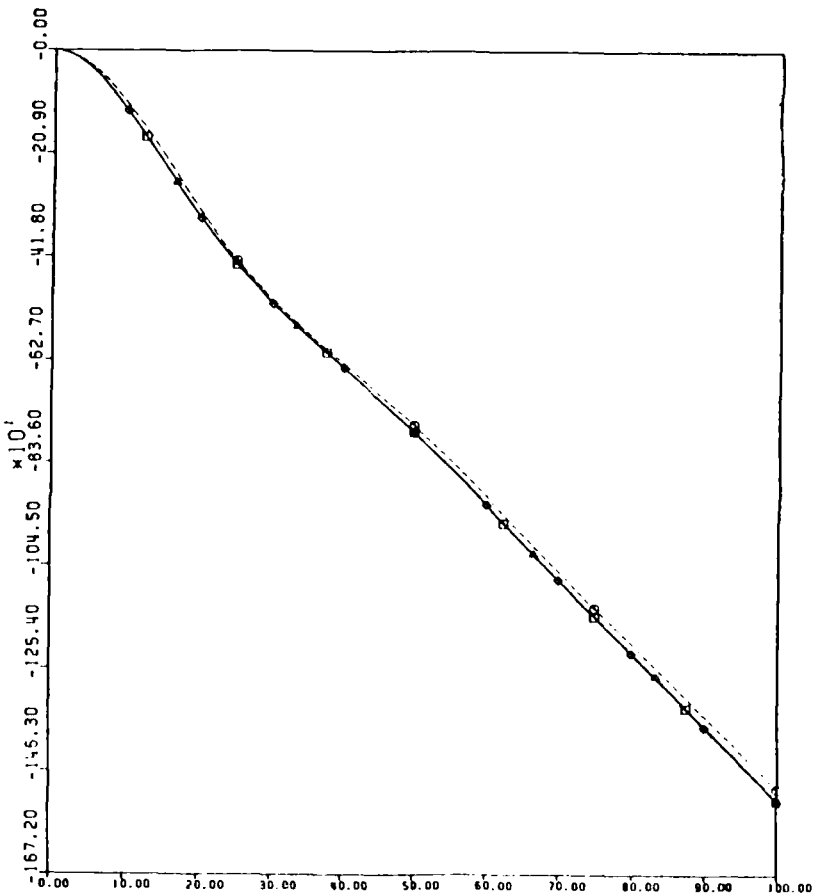



Figure 10.2b. Functional Estimator Gain Component $\phi_{gn}$

Damping $c_0 = 10^{-4}$; estimator $R = 10^{-4}$

number of elements $n_e = 4, 6, 8, 10$

| $n_e$ | $\alpha_{fn}$ | $\alpha_{gn}$ | $\beta_{gn}$ |
|---|---|---|---|
| 2 | 5.0358 | 12.680 | -1334.9 |
| 3 | 5.2514 | 13.789 | -1455.4 |
| 4 | 5.3195 | 14.149 | -1495.7 |
| 5 | 5.3478 | 14.300 | -1512.2 |
| 8 | 5.3611 | 14.371 | -1520.1 |

Table 10.1. Scalar Components of Functional Estimator Gains

Damping coefficient $c_0 = 10^{-4}$ ; estimator R = $10^{-4}$

number of elements $n_e$ = 2, 3, 4, 5, 8

| $n_e$ | $\alpha_{fn}$ | $\alpha_{gn}$ | $\beta_{gn}$ |
|---|---|---|---|
| 4 | 5.3195 | 14.149 | -1495.7 |
| 6 | 5.3567 | 14.347 | -1517.5 |
| 8 | 5.3611 | 14.371 | -1520.1 |
| 10 | 5.3623 | 14.377 | -1520.8 |

Table 10.2. Scalar Components of Functional Estimator Gains

Damping coefficient $c_0 = 10^{-4}$; estimator R = $10^{-4}$

number of elements $n_e$ = 4, 6, 8, 10

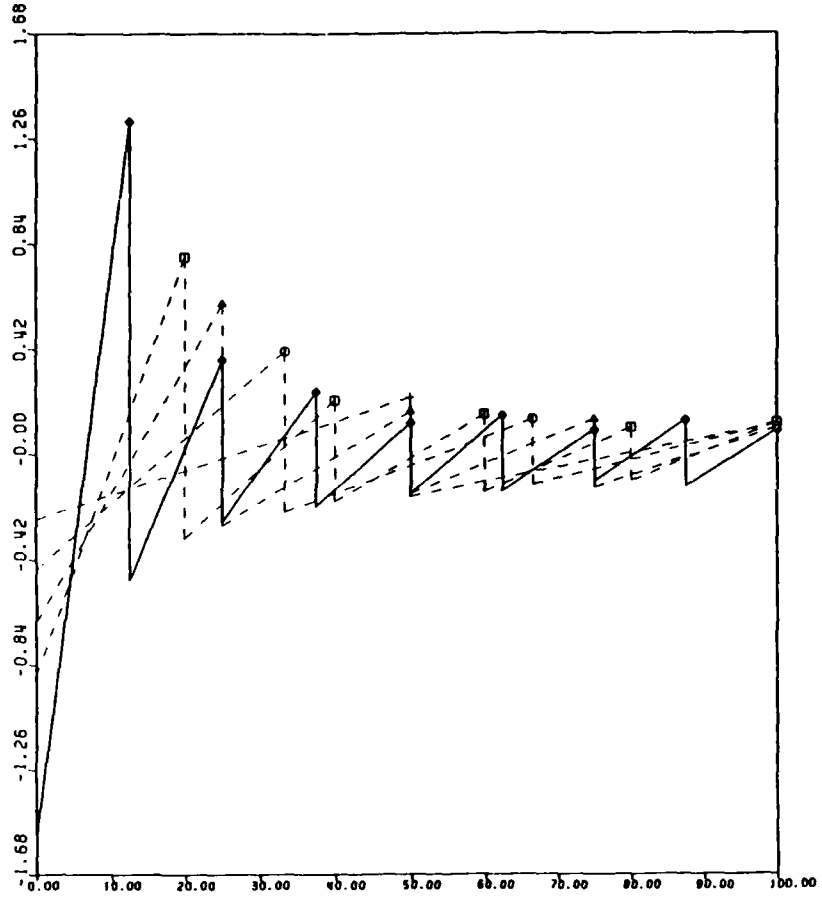
Figures 10.3a and 10.3b and Table 10.3 give the numerical results for the finite dimensional estimators when the structural damping is zero. While Theorem 8.4 says that the solutions to the finite dimensional Riccati equations for these estimators will not converge when the damping is zero and $\hat{Q}$ is coercive on E, we have no result to predict the convergence for zero damping when $\hat{Q}$ is not coercive (even though $\hat{Q}_2$ is coercive on H). From the numerical results though, $\hat{f}_n$ does not appear to converge.

| $n_e$ | $\alpha_{fn}$ | $\alpha_{gn}$ | $\beta_{gn}$ |
|---|---|---|---|
| 2 | 5.0730 | 12.868 | -1354.4 |
| 3 | 5.3390 | 14.253 | -1506.0 |
| 4 | 5.4417 | 14.806 | -1568.0 |
| 5 | 5.4894 | 15.067 | -1596.3 |
| 8 | 5.5398 | 15.345 | -1627.2 |

Table 10.3. Scalar Components of Functional Estimator Gains

Zero damping; estimator $R = 10^{-4}$

number $n_e = 2, 3, 4, 5, 8$




Figures 10.3a. Functional Estimator Gain Component $\phi''_{fn}$

Zero damping; estimator R = $10^{-4}$

number of elements $n_e$ = 2, 3, 4, 5, 8

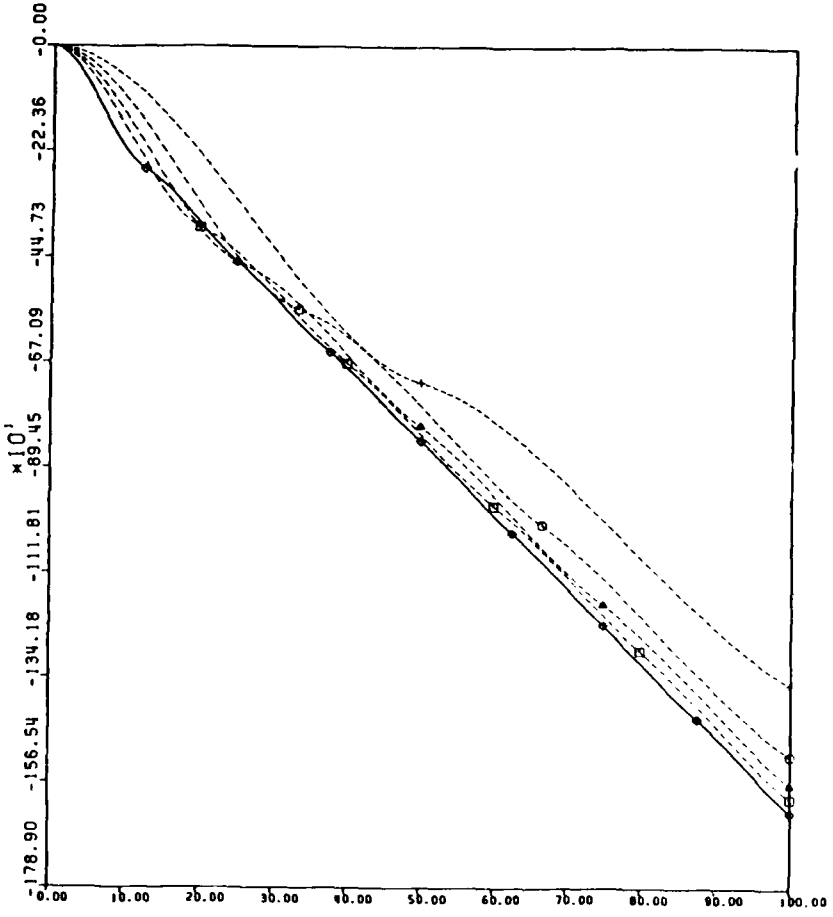



Figure 10.3b. Functional Estimator Gain Component $\phi_{gn}$

Zero damping; estimator R = $10^{-4}$

number of elements $n_e$ = 2, 3, 4, 5, 8

## 10.2 Approximation of the Optimal Compensator

Finally, for the damping $c_0 = 10^{-4}$, R = .05 in the control problem and R = $10^{-4}$ in the estimator problem, we construct the finite dimensional compensator in Figure 9.1; i.e, for each $n = 2n_e + 1$, we apply the $n_e^{th}$ control law represented by the functional gains in Figure 6.4 and Table 6.4 to the output of the $n_e^{th}$ estimator represented by the functional gains in Figure 10.2 and Table 10.2. As the number of elements increases, the transfer function in (9.16) of the finite dimensional compensator converges to the transfer function in (7.14) of the optimal infinite dimensional compensator, as described in Section 9.3. Theorem 9.5 and Remark 9.6 apply. Figure 10.4 shows the frequency response (bode plots) of the finite dimensional compensators for 4, 6, 8 and 10 elements. The phase plot is for 10 elements only. These plots indicate that the finite dimensional compensator for eight or more elements is virtually identical to the optimal infinite dimensional compensator, as predicted by the functional gain convergence in Figures 6.4 and 10.2.

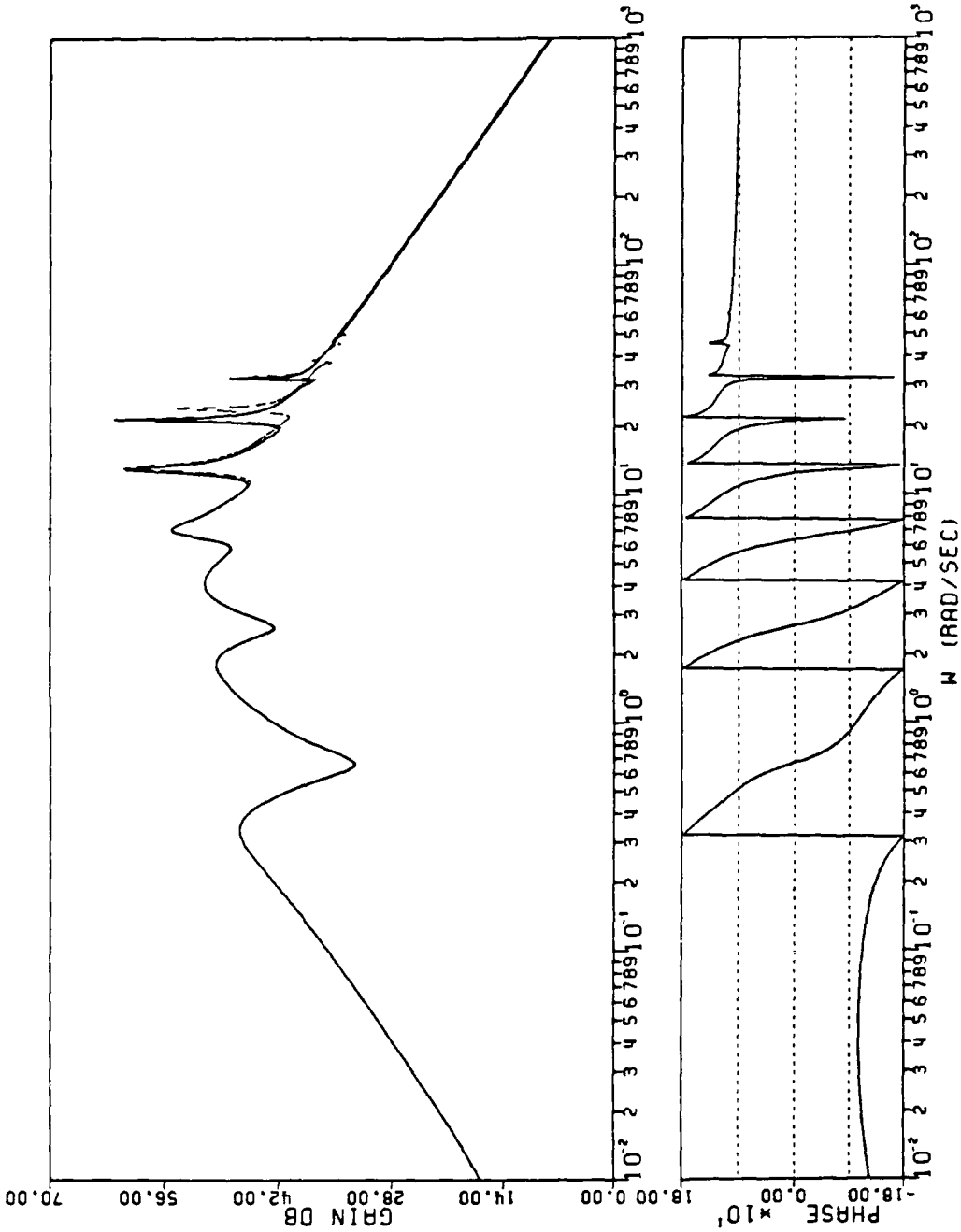



**Figure 10.4.** Frequency Response (Bode Plot) of Compensators

Damping $c_0 = 10^{-4}$; control R = .05, estimator R = $10^{-4}$

number of elements $n_e$ = 4, 6, 8, 10

## 10.3 Comments on the Structure and Dimension of the Implementable Compensators

Though this paper does not address the problem of obtaining the lowest-order compensator that closely approximates the infinite dimensional compensator, we should note that the compensators based on eight and ten elements here are unnecessarily large because the finite element scheme that we chose is not nearly the most efficient in terms of the dimension required for convergence. (The dimension of the first-order differential equation in the compensator is 2(2ne+1).) We used cubic Hermite splines here to demonstrate that the finite element scheme most often used to approximate beams in other engineering applications can be used in approximating the optimal compensator. In [G5], we compare the present scheme with one using cubic B-splines and one using the natural mode shapes as basis vectors. The natural mode shapes yield the fastest converging compensators, but the B-splines are almost as good. The only advantages of the Hermite splines result from the fact that the coding to build the basic matrices (mass, stiffness, etc.) is simpler than for B-splines and the fact that, before the Riccati equations based on, say, ten natural modes are solved, a much larger finite element approximation of the structure must be used to get the ten modes accurately.

To understand the redundancy in the large finite dimensional compensators here, it helps to consider the structure of the optimal compensator. It is based on an infinite dimensional state estimator that has a representation of each of the structure's modes. In the present example, the optimal compensator estimates and controls the the first six modes significantly, the next three modes slightly, and virtually ignores the rest. This observation is

based on the projections of the functional gains onto the natural modes and on comparison of the open-loop and closed-loop eigenvalues. (See [G5] for more detail, including the spectrum of closed-loop system -- which is stable -- obtained with the ten-element compensator here.) The infinite dimensional compensator then has an infinite number of modes that contribute nothing to the input-output map of the compensator. These inactive modes are just copies of all the open-loop modes past the first nine. They can be truncated from the compensator without affecting the closed-loop system response significantly. The number of active modes in the compensator -- i.e., the modes that contribute to the input-output map -- depends on the structural damping and the Q's and R's in the LQG problem statement. (See the discussion in Section 6.4 about the effect of damping and control weighting on performance.)

The compensator computed here based on ten elements has 21 modes (although we did not do the computations in modal coordinates). Nine of these compensator modes are virtually identical to the nine active modes in the infinite dimensional compensator, and the twelve inactive modes are approximations to the tenth through twenty-first open-loop modes of the structure. The inactive modes result from the large number of elements needed to approximate the active compensator modes accurately. Now that we essentially have the optimal compensator in the ten- element compensator, we could truncate the twelve inactive modes and implement a compensator with nine modes. And we probably could reduce the compensator even further using an order reduction method like balanced realizations.

## 11. Conclusions

For the deterministic linear-quadratic optimal regulator problem for a flexible structure with bounded input operator (the $B_0$ in (2.1)), the approximation theory in Sections 4 and 5 is reasonably complete. The most important extensions should be to the corresponding (very difficult) problem with unbounded input operator, for which there exists little approximation theory. Because of the different kinds of boundary input operators, stiffness operators and structural damping, all of which must be considered in detail when $B_0$ is unbounded, it seems unlikely that the approximation theory for the unbounded-input case can be made as complete as the theory here.

The convergence results in Section 8 for the estimation problem are less complete than those for the control problem because rigid-body modes present more technical difficulties for the proofs in the estimator case. However, our analysis and numerical experience suggest that the difficulties only make the proofs harder and that the convergence in the estimation problem is identical to the convergence in the control problem, and that controllable and observable rigid-body modes make no qualitative difference in either problem.

Where we would most like substantial improvement over the results of this paper is in Section 9.2, which considers how the approximating closed-loop systems obtained by controlling the distributed model of the structure with the finite dimensional compensators converge to the optimal closed-loop system, obtained with the infinite dimensional compensator. Theorem 9.4 gives us what we want for problems where the damping does not couple the natural modes of free vibration and the natural mode shapes are the basis vectors for the

approximation scheme. In particular, this theorem says that, if the optimal closed-loop system is uniformly exponentially stable, then so are the approximating closed-loop systems for sufficiently large order of approximation. We have verified numerically the stability of the approximating close-loop systems for the example in Sections 6 and 10, where the basis vectors are not the modes. This example and others have made us suspect that Theorem 9.4 is true when the basis vectors satisfy Hypothesis 4.1 only and when the damping couples the modes.

Another possible approach to analyzing the convergence of the approximating closed-loop systems to the optimal closed-loop system is to use the input-output description in frequency domain. Results like those in Section 9.3 are useful for this, although for the closed-loop stability we want, we probably need the transfer functions of the finite dimensional compensators to converge more uniformly on the compensator resolvent set than we have proved here. In our example, Figure 10.4 indicates that these transfer functions converge uniformly on the imaginary axis, but we have no theorem that guarantees this.

## References


B1 A.V. Balakrishnan, Applied Functional Analysis, Second Edition, Springer-Verlag, New York, 1981.

B2 A.V. Balakrishnan, "Strong Stability and the Steady-State Riccati Equation," Appl. Math. Opt., 7, 1981, pp. 335-345.

B3 M.J. Balas, "Feedback Control of Flexible Systems," IEEE Trans. AC, AC-23, pp 673-679, 1978.

B4 M.J. Balas, "Modal Control of Certain Flexible Systems", SIAM J. Contr. Opt., 16(1978) pp.450-462.

B5 H.T. Banks and J.A. Burns, "Hereditary Control Problems: Numerical Methods Based on Averaging Approximations," SIAM J. Contr. Opt., 16(1978) pp. 169-208.

B6 H.T. Banks and K. Kunisch, "The Linear Regulator Problem for Parabolic Systems," SIAM J. Contr. Opt., Vol. 22, No. 5, September 1984, pp. 684-698.

B7 D.S. Berstein and D.C. Hyland, "The Optimal Projection Approach to Designing Finite-dimensional Controllers for Distributed Parameter Systems," 23rd IEEE Conference on Decision and Control, Las Vegas, NV, December 1984.

B8 D.S. Berstein and D.C. Hyland, "The Optimal Projection Equations for Finite-dimensional Fixed-order Dynamic Compensation of Infinite-Dimensional Systems," SIAM J. Contr. Opt., to appear.

C1 G. Chen and D.L. Russell, "A Mathematical Model for Linear Elastic Systems with Structural Damping," Quarterly of Applied Mathematics, January 1982, pp. 433-454.

C2 R.W. Clough and J. Penzien, Dynamics of Structures, McGraw-Hill, New York, 1975.

C3 R.F. Curtain, "On Stabilizability of Linear Spectral Systems via State Boundary Feedback," SIAM J. Contr. Opt., Vol. 23, No. 1, 1985, pp. 144-152.

C4 R.F. Curtain and A.J. Pritchard, Infinite Dimensional Linear Systems Theory, Springer-Verlag, New York, 1978.

C5 R.F. Curtain and D. Salamon, "Finite Dimensional Compensators for Infinite Dimensional Systems with Unbounded Control Action," SIAM J. Contr. Optim., to appear.

D1 N. Dunford and J.T. Schwartz, Linear Operators, Part II, Wiley Interscience, New York, 1963.

G1 J.S. Gibson, "An Analysis of Optimal Modal Regulation: Convergence and Stability," SIAM J. Contr. Opt., Vol. 19, No. 5, September 1981, pp. 686-707.

G2 J.S. Gibson, "A Note on Stabilization of Infinite Dimensional Linear Oscillators by Compact Linear Feedback," SIAM J. Contr. Opt., 18(1980) pp. 311-316.

G3 J.S. Gibson, "Linear-quadratic Optimal Control of Hereditary Differential Systems: Infinite Dimensional Riccati Equations and Numerical Approximations," SIAM J. Contr. Opt., Vol. 21, No. 1, January 1983, pp. 95-139.

G4 J.S. Gibson, "The Riccati Integral Equations for Optimal Control Problems on Hilbert Spaces," SIAM J. Contr. Opt., Vol. 17, No. 4, July 1979, pp. 537-565.

G5 J.S. Gibson and A. Adamian, "A Comparison of Approximation Schemes for Optimal Control of Flexible Structures," to appear in SIAM Frontiers Edition on Control of Distributed Systems, 1988.

G6 J.S. Gibson and D.L. Mingori, "Approximation and Compensator Order Determination in Optimal Control of Flexible Structures," Proc. IFAC Workshop on Model Error Concepts and Compensation, Boston, MA, June 1985.

G7 J.S. Gibson, D.L. Mingori, A. Adamian and F. Jabbari, "Approximation of Optimal Infinite Dimensional Compensators for Flexible Structures," JPL Workshop on Identification and Control of Space Structures, San Diego, CA, June 1984.

G8 J.S. Gibson, D.L. Mingori, et al, Integrated Control/Structure Research for Large Space Structures, HR Textron Report #956541-Final, to JPL, September 1984.

G9 J.S. Gibson and M. Navid, "Approximate Solution of Riccati Algebraic Equations in Optimal Control and Estimation of Hyperbolic Systems," International Symposium on Mathematical Theory of Networks and Systems, Santa Monica, CA, August 1981.

K1 T. Kato, Perturbation Theory for Linear Operators, Second Edition, Springer-Verlag, New York, 1984.

L1 I. Lasiecka and R. Triggiani, "Riccati Equations for Hyperbolic Partial Differential Equations with $L2(0,T; L2(\Gamma))$ - Dirichlet Boundary Terms," Proceedings of 24th Conference on Decision and Control, Ft. Lauderdale, FL, December 1985, pp. 114-116; to appear in SIAM J. Contr. Opt.

L2 I. Lasiecka and R. Triggiani, "Dirichlet Boundary Control

Problem for Parabolic Equations with Quadratic Cost: Analyticity and Riccati's Feedback Synthesis," SIAM J. Contr. Opt., Vol. 21, No. 1, January 1983, pp. 41-67.

M1 D.L. Mingori, J.S. Gibson, P. Blelloch and A. Adamian, "Control of a Flexible Space Antenna: A Finite Dimensional Perspective Based on Distributed Parameter Theory," JPL Workshop on Identification and Control of Space Structures, San Diego, CA, June 1984.

M2 B.C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," IEEE Trans. Automat. Contr., vol. AC-26, pp. 17-32, Feb., 1981.

S1 M.H. Schultz, Spline Analysis, Prentice-Hall, Englewood Cliffs, N.J., 1973.

S2 J.M. Schumacher, "A Direct Approach to Compensator Design for Distributed Parameter Systems," SIAM J. Contr. Opt., Vol. 21, No. 6, 1983, pp. 823-836.

S3 R.E. Showalter, Hilbert Space Methods for Partial Diffenential Equations, Pitman, London, 1977.

S4 G. Strang and G.J. Fix, An Analysis of the Finite Element Method, Prentice-Hall, Englewood Cliffs, N.J., 1973.

## APPENDIX

### Errata for [G1]

In the first paragraph of the proof of Theorem 2.1 on page 689 of [G1], the first sentence should be:

> If a dissipative operator is invertible, its inverse is dissipative.

At the beginning of the fifth line of the same paragraph, the expression $(\alpha x+y)$ should be deleted the first time it occurs.

The next-to-last sentence of the paragraph should be:

> Hence, if a densely defined maximal dissipative operator has dense range, its inverse is maximal dissipative.

The theorem is correct as stated.

---

In the current paper, we use Theorem 2.1 of [G1] to conclude that the operator $\tilde{A}$ defined in Section 2 is maximal dissipative (see (2.10)-(2.12)) and that the operator $\mathring{A}$ in (2.16) has a unique maximal dissipative extension.

NASA

# Report Documentation Page

| 1. Report No | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| NASA CR-181705<br>ICASE Report No. 88-48 | | |

| 4. Title and Subtitle | 5. Report Date |
|---|---|
| APPROXIMATION THEORY FOR LQG OPTIMAL CONTROL OF FLEXIBLE STRUCTURES | August 1988 |
| | 6. Performing Organization Code |

| 7. Author(s) | 8. Performing Organization Report No |
|---|---|
| J. S. Gibson and A. Adamian | 88-48 |
| | 10. Work Unit No. |
| | 505-90-21-01 |

| 9. Performing Organization Name and Address | 11. Contract or Grant No |
|---|---|
| Institute for Computer Applications in Science and Engineering<br>Mail Stop 132C, NASA Langley Research Center<br>Hampton, VA 23665-5225 | NAS1-17070, NAS1-18107 |
| | 13. Type of Report and Period Covered |
| | Contractor Report |

| 12. Sponsoring Agency Name and Address | 14. Sponsoring Agency Code |
|---|---|
| National Aeronautics and Space Administration<br>Langley Research Center<br>Hampton, VA 23665-5225 | |

15. Supplementary Notes

Langley Technical Monitor:
Richard W. Barnwell

Submitted to SIAM Journal of Control and Optimization

Final Report

16. Abstract

This paper presents approximation theory for the linear-quadratic-Gaussian optimal control problem for flexible structures whose distributed models have bounded input and output operators. The main purpose of the theory is to guide the design of finite dimensional compensators that approximate closely the optimal compensator separates into an optimal linear-quadratic control problem lies in the solution to an infinite dimensional Riccati operator equation. The approximation scheme in the paper approximates the infinite dimensional LQG problem with a sequence of finite dimensional LQG problems defined for a sequence of finite dimensional, usually finite element or modal, approximations of the distributed model of the structure. Two Riccati matrix equations determine the solution to each approximating problem.

The finite dimensional equations for numerical approximation are developed, including formulas for converting matrix control and estimator gains to their functional representation to allow comparison of gains based on different orders of approximation. Convergence of the approximating control and estimator gains and of the corresponding finite dimensional compensators is studied. Also, convergence and stability of the closed-loop systems produced with the finite dimensional compensators are discussed. The convergence theory is based on the convergence of the solutions of the finite dimensional Riccati equations to the solutions of the infinite dimensional Riccati equations. A numerical example with a flexible beam, a rotating rigid body, and a lumped mass is given.

| 17. Key Words (Suggested by Author(s)) | 18. Distribution Statement |
|---|---|
| control, flexible structures, optimal control, approximation | 08 - Aircraft Stability and Control<br>63 - Cybernetics<br><br>Unclassified - unlimited |

| 19. Security Classif. (of this report) | 20. Security Classif. (of this page) | 21. No. of pages | 22. Price |
|---|---|---|---|
| Unclassified | Unclassified | 121 | A06 |

NASA FORM 1626 OCT 86