Finite Dimensional Approximations of Distributed Parameter Control Systems

by

David Dean Hill

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

APPROVED:

J. A. Burns, Chairman

T. L. Herdman

E.M. Cliff

K. B. Hannsgen

W. A. Cebuhar

August, 1989

Blacksburg, Virginia

Finite Dimensional Approximations of

Distributed Parameter Control Systems

by

David Dean Hill

Committee Chairman: John A. Burns
Mathematics

(ABSTRACT)

In this paper we consider two separate approaches to the development of finite dimensional control systems for approximating distributed parameter models. One method uses the "standard finite element" approximations to construct the basic system matrices. The resulting system can then be balanced by any of several balancing algorithms. The second method is based on truncating infinite dimensional balanced realizations of the the input-output map. Both approaches are applied to a control problem goverened by the heat equation. We present a comparison of the resulting finite dimensional models.

DEDICATION

This work is dedicated to my parents,

ACKNOWLEDGEMENTS

First and foremost, I wish to thank Profesor John A. Burns, my advisor, for his guidance and patience.

I also wish to acknowledge the support given to me by all other members of the mathematics faculty.

I would also like to thank Dr. Robert Miller for his help with some of the numerical calculations.

I would like to thank my wife, for all of the help and support she has given me.

Finally, I would like to acknowledge the financial support given to me under AFOSR Grant No. AFOSR-89-001, DARPA Grant No F49620-87-C-0116, and SDIO under contract F49620-87-C-0088.

TABLE OF CONTENTS

Chapter I

1.1 Introduction

The development of finite dimensional models for control of distributed parameter systems often requires that some type of numerical approximation scheme be introduced at some point of the modeling process. It is of great concern whether or not the matrices generated by this approximation scheme preserve important properties of the infinite dimensional system (e.g. stabilizability and controllability, [Ba-1], [Bu-1], [Bu-4], and [Bu-5]). Another important consideration is whether control designs for an approximating system are converging to an appropriate control design for the original distributed parameter system [Ba-1]. Even if the scheme converges and important properties of the system are preserved, how well the properties are preserved is important since we are using numerical approximations (for example the finite dimensional system should not be "near" an unstabilizable system).

In this paper we consider two separate approaches to the development of finite dimensional control systems for approximating distributed parameter models. One method uses the "standard finite element" approximations to construct the basic system matrices. The resulting system can then be balanced by any of several balancing algorithms. The second method is based on truncating infinite dimensional balanced realizations of the the input-output map. Both approaches are applied to a control problem governed by the heat equation. We present a comparison of the resulting finite dimensional models.

In Chapter II we present a definition for the measure of stabilizability for finite dimensional systems [E-1] as well as a definition for the measure of stabilizability for infinite dimensional systems. It will be proven that if an approximation scheme satisfies the property POES [Ba-1], then the finite dimensional matrices generated by this scheme have measures of stabilizabiliy uniformly bounded away from zero. In Chapter III we present results on the balancing of finite dimensional systems as well as results on infinite dimensional balanced systems. We develop a characterization of a large class of finite dimensional systems. A characterization of the balanced realization for special infinite dimensional systems is also discussed. In Chapter IV numerical methods are developed for the various schemes. The finite element method is developed for the heat equation. The truncated balanced realization matrices are described for the finite dimensional approximations and a method for calculating them is given for the heat equation. In chapter V the truncated balanced realization scheme will be compared to the standard finite element scheme for the one dimensional heat equation. The measures of controllability, observability, stabilizability, and detectability are computed for the finite element, balanced finite element and truncated balanced realization scheme.

1.2 Notation

Throughout this paper we will use the following notation. If H is a Hilbert space, then $\| \cdot \|$ and $\langle \cdot, \cdot \rangle$ will denote the norm and inner product on H. If T is a linear operator from the space H to a space U, we denote the domain of T by D(T). We denote the spectrum of A by $\sigma(A)$. For a Hilbert space Z, the set of all square integrable functions defined on [a,b] with values in Z will be denoted by $L_2(a,b;Z)$. The triple (A,B,C) will represent a system with dynamic operator A, control operator B, and observation operator C. The pair (A,B) represents a system with control operator B. The measure of controllability of a pair (A,B) will be denoted c(A,B) and the measure of stabilizability will be denoted s(A,B). The triple (A_M,B_M,C_M) will denote a M dimensional approximation of the system (A,B,C). If $T:X \rightarrow Y$ is a bounded linear operator from a Hilbert space X to a Hilbert space Y, then $\|T\|_2$ denotes the spectral norm $\|T\|_2 = \sup_{\|x\|=1} \|Tx\|_Y$.

Chapter II

2.1 Introduction

The convergence of numerical approximations to the solution of an infinite dimensional control problem is highly dependent on the schemes which are used to approximate the control problem. The approximating systems must preserve the appropriate control system properties, and in some cases these properties must be preserved uniformly. For example, numerical approximations of LQR-problems must preserve stabilizability and detectability uniformly under approximation. Moreover, most approximation systems are constructed using numerical schemes and, therefore, one is faced with several practical issues that need to be considered when using these systems in control design. In this chapter we discuss some of these issues and present practical methods which can be used to compute some of the aforementioned quantities.

2.2 Preservation of Stabilizability Under Bounded Perturbations

In this section a general theorem will be given which implies that stabilizability of a system is preserved under small bounded perturbations. The framework in which the theorem is presented applies to the finite dimensional case as well as many infinite dimensional cases.

We shall need the following result. A proof of this result can be found in [P-1] on page 76.

Lemma(2.2.1) Suppose A is the infinitesimal generator of a C_0-semigroup T(t) which satisfies $\|T(t)\| \leq Me^{\gamma t}$ where γ is a real number. If δA is any bounded linear operator on H, then $A + \delta A$ is the infinitesimal generator of a C_0- semigroup S(t) on H satisfying

$$\|S(t)\| \leq Me^{(\gamma + M\|\delta A\|)t}.$$

Theorem(2.2.2) Assume that A, B, and D are linear operators, where U and H are Hilbert spaces with $A:H \to H$, $B:U \to H$, and $D:H \to U$ is bounded. If A + BD generates an exponentially stable C_0-semigroup T(t), then there exists $\epsilon > 0$ such that if $\delta A:H \to H$ and $\delta B:U \to H$ are bounded linear operators satisfying $\|\delta A\| + \|\delta B\| < \epsilon$, then $(A + \delta A) + (B + \delta B)D$ generates an exponentially stable C_0-semigroup S(t).

Proof:

Since δA, δB, and D are bounded operators, $\delta A + \delta BD$ is bounded and

(2.2.1) $$\|\delta A + \delta BD\| \leq \|\delta A\| + \|\delta BD\| \leq \|\delta A\| + \|\delta B\|\|D\|.$$

Since A+BD generates an exponentially stable C_0-semigroup T(t), (see [C-1]), there exists an $\omega > 0$ such that T(t) satisfies

(2.2.2) $$\|T(t)\| \leq Me^{-\omega t}.$$

Thus, from Lemma 2.2.1 it follows that the semigroup S(t) satisfies

(2.2.3) $$\|S(t)\| \leq Me^{(-\omega + M\|\delta A + \delta BD\|)t}$$

which implies

$$(2.2.4) \qquad \|S(t)\| \leq Me^{(-\omega+M\|\delta A\|+M\|\delta B\|\|D\|)t}.$$

Clearly, if $\|\delta A\|$ and $\|\delta B\|$ are small, the quantity $(-\omega+M\|\delta A\|+M\|\delta B\|\|D\|)$ is negative and thus $S(t)$ is an exponentially stable C_0-semigroup.

Corollary(2.2.3) If the pair (A,B) is stabilizable, see [C-1], then it is stabilizable under small bounded perturbations (δA, δB).

Proof:

By the definition of stabilizability, see [C-1], there exists a bounded linear operator D such that the operator A+BD generates an exponentially stable C_0-semigroup on H. The result follows from Theorem(2.2.2).

Corollary(2.2.3) indicates that it may be possible to define a "radius of stabilizability" for a stabilizable pair (A,B). The following definition concerns bounded perturbations.

Definition(2.2.4) The radius of stabilizability r for a pair (A,B) is given by

$$(2.2.5) \qquad r = r(A,B) = \inf\{\|[\delta A, \delta B]\|_2 : (A + \delta A,\ B+\delta B) \text{ is not stabilizable}\}.$$

Note that $0 < r$, and if

(2.2.6) $$\|[\delta A,\delta B]\|_2 < r,$$

then $(A + \delta A, B + \delta B)$ is stabilizable.

We shall need the following result. A proof of this result can be found in [P-1] on page 11.

Theorem(2.2.5) If A generates a C_0-semigroup S(t), then $A+\epsilon I$ generates a C_0-semigroup T(t) and $S(t)=e^{-\epsilon t}T(t)$.

Theorem(2.2.6) If the radius of stabilizability for (A,B) is r, then for every real number q with $r < q$ there exists bounded linear operators δA and δB with $\|[\delta A,\delta B]\|_2 = q$ such that $(A + \delta A, B + \delta B)$ is not stabilizable.

Proof:

The definition of r implies that there exist bounded linear operators $\delta\tilde{A}$ and $\delta\tilde{B}$ such that $(A + \delta\tilde{A}, B + \delta\tilde{B})$ is not exponentially stabilizable by any feedback operator and

(2.2.7) $$r \leq \|[\delta\tilde{A},\delta\tilde{B}]\|_2 < q.$$

Define the continuous function $f(\epsilon)$ by

(2.2.8) $$f(\epsilon) = \|[\delta\tilde{A} + \epsilon I,\ \delta\tilde{B}]\|_2.$$

Clearly, $f(0) < q$ and $\lim_{\epsilon \to \infty} f(\epsilon) = \infty$. Since $f(\epsilon)$ is continuous, there exists $\epsilon_0 > 0$ such that $f(\epsilon_0) = q$. If $(A + \delta\tilde{A} + \epsilon_0 I,\ B + \delta\tilde{B})$ is stabilizable, then there exists a bounded operator K such that the C_0-semigroup T(t) generated by $(A + \delta\tilde{A} + \epsilon_0 I) + (B + \delta B)K$ is stable. Let S(t) be the C_0-semigroup generated by $(A + \delta\tilde{A}) + (B + \delta\tilde{B})K$. By Theorem(2.2.5) it follows that

(2.2.9) $$S(t) = e^{-\epsilon_0 t} T(t).$$

Since T(t) is stable, there exist M and $\omega > 0$ such that,

(2.2.10) $$\|T(t)\| \le Me^{-\omega t}.$$

By equations (2.2.9) and (2.2.10) we have

(2.2.11) $$\|S(t)\| = \|e^{-\epsilon_0 t} T(t)\| = e^{-\epsilon_0 t}\|T(t)\| \le Me^{-\epsilon_0 t} e^{-\omega t}.$$

Therefore,

(2.2.12) $$\|S(t)\| \le Me^{-(\omega+\epsilon_0)t},$$

which contradicts the assumption that $(A + \delta\tilde{A},\ B + \delta\tilde{B})$ is unstabilizable.

2.3 Measures of Controllability and Stabilizability

In the paper [E-1] the measure of controllability c(A,B) for a finite dimensional system (A,B) is defined by

(2.3.1) $$c(A,B) = \min\{\|[\delta A,\delta B]\|_2 : (A+\delta A,B+\delta B) \text{ is uncontrollable}\}.$$

A method for computing c(A,B) is presented in [E-2]. This method involves minimizing the functional

(2.3.2) $$\phi(x) = \frac{x^*A\left[I-\frac{xx^*}{x^*x}\right]A^*x}{x^*x} + \frac{BB^*}{x^*x},$$

over all $x\epsilon C^n - \{0\}$.

The functional $\phi(x)$ depends on 2n real variables (where A is nxn), and is "constrained" by the deletion of 0. In Section 2.4 a method will be presented which can be used to compute this measure of controllability for finite dimensional systems which does not involve a constrained minimization. A similar definition for the measure of stabilizability s(A,B) is given in [E-1]. The measure of controllability c(A,B) and the measure (radius) of stabilizability s(A,B) (r(A,B)) may be found by

(2.3.3) $$c(A,B) = \min_{\lambda\epsilon C} \sigma_{\min}[A-\lambda I,B],$$

(2.3.4) $$r(A,B)=s(A,B)=\min_{\lambda\epsilon C, re\lambda\geq 0}\sigma_{min}[A-\lambda I,B],$$

where σ_{min} denotes the smallest singular value of the respective matrices (see [E-1]). Using equations (2.3.3) and (2.3.4) to compute c(A,B) and s(A,B) can take a large amount of computer time since a singular value decomposition would have to be computed for each new lambda found in the minimization scheme. For the case where A is a real symmetric matrix with negative eigenvalues, the measure of stabilizability is given by (see [Bu-3])

(2.3.5) $$s(A,B)=\sigma_{min}[A,B].$$

2.4 Methods for Computing the Measure of Controllability; Finite Dimensional Systems

In this section we present a method which can be used to compute the measure of controllability and the measure of stabilizability for the finite dimensional case. The method for computing c(A,B) is based on an unconstrained minimization problem for a functional $f(x,\lambda)$ which has continuous second derivatives near its minimum values. Thus, a quasi-Newton method can be used to compute the minimum. The functional used to compute the measure of stabilizability s(A,B) is the same as the functional $f(x,\lambda)$ which is used to compute the measure of controllability. However, a constraint must be imposed on one of the variables.

Lemma(2.4.1) Let R be a Hermitian matrix with minimum eigenvalue $e_{min}\geq 0$. Suppose $K > e_{min}$. If $\tilde{x}$ is the minimizer of $f_R(x)$ over all $\mathbf{C}^n$ where

(2.4.1) $$f_R(x) = x^*Rx+K(1-|x|)^2,$$

then

(2.4.2) $$e_{min} = \frac{\tilde{x}^*R\tilde{x}}{|\tilde{x}|^2} = \frac{f_R(\tilde{x}) - K(1-|\tilde{x}|)^2}{|\tilde{x}|^2}.$$

Proof:

First it will be shown that $\tilde{x}$ actually exists. Since R is Hermitian and has nonnegative eigenvalues, for any $x_0 \epsilon C^n$ such that $|x_0| > 1$ we have

(2.4.3) $$0 \leq \frac{x_0^*Rx_0}{|x_0|^2} \leq x_0^*Rx_0.$$

If $x_b = \frac{x_0}{|x_0|}$, then

(2.4.4) $$f_R(x_b) = x_b^*Rx_b + (1 - |x_b|)^2$$

(2.4.5) $$= x_b^*Rx_b$$

(2.4.6) $$= \frac{x_0^*Rx_0}{|x_0|^2} \leq x_0^*Rx_0 \leq f_R(x_0).$$

Thus for every x_0 with $|x_0| > 1$, there exists x_b with $|x_b| = 1$ and

(2.4.7) $$f_R(x_b) \le f_R(x_0).$$

Therefore,

(2.4.8) $$\inf_{x \epsilon C^n} f_R(x) = \inf_{x \epsilon C^n, \|x\| \le 1} f_R(x).$$

The compactness of the unit ball in C^n implies the existence of a minimizer $\tilde{x}$, with $\|\tilde{x}\| \le 1$.

Now we show $\tilde{x} \neq 0$. If y is a normalized eigenvector of R for the eigenvalue e_{min}, then

(2.4.9) $$f_R(y) = y^*Ry + K(1 - \|y\|)^2 = \frac{y^*Ry}{\|y\|^2} = e_{min} < K = f_R(0).$$

Hence $x=0$ does not minimize f_R. Let $S = \{x \epsilon C^n : \|x\| = \|\tilde{x}\|\}$. For $x_s \epsilon S$

(2.4.10) $$f_R(x_s) = x_s^* R x_s + K(1 - \|\tilde{x}\|)^2.$$

Since $\tilde{x}$ minimizes $f_R(x)$, for each $x_s \in S$

(2.4.11) $$f_R(\tilde{x}) \le f_R(x_s).$$

Equations (2.4.10) and (2.4.11) imply that

(2.4.12) $$\tilde{x}^*R\tilde{x} \le x_s^*Rx_s, \text{ for all } x_s \epsilon S,$$

and Theorem 1 on page 142 of [F-1] implies

(2.4.13) $$e_{min} = \min_{x_s \epsilon S} \frac{x_s^* R x_s}{\|x_s\|^2} = \frac{\tilde{x}^* R \tilde{x}}{\|\tilde{x}\|^2}.$$

The previous lemma will be used to construct an algorithm for the calculation of the measures of controllability c(A,B) and stabilizability s(A,B). Note that the functional $f_R(x)$ has continuous derivatives except at $x=0$; hence f_R is differentiable at $x=\tilde{x}$.

Let $H(\lambda)$ denote the matrix,

(2.4.14) $$H(\lambda) = [A - \lambda I, B].$$

Recall that (2.3.3) implies that $c(A,B) = \min_{\lambda \epsilon C} [\sigma_{min} H(\lambda)]$.

Theorem(2.4.2) Let $c(A,B) = \min_{\lambda \epsilon C} [\sigma_{min} H(\lambda)]$ and assume that $K > c(A,B)$. If $\tilde{x}$ and $\tilde{\lambda}$ minimize

(2.4.15) $$F(x,\lambda) = x^* H(\lambda) H(\lambda)^* x + K(1 - \|x\|)^2,$$

over $C^n x C$, then

(2.4.16) $$c(A,B)^2 = \frac{F(\tilde{x},\tilde{\lambda}) - K(1 - \|\tilde{x}\|)^2}{\|\tilde{x}\|^2}.$$

Proof:

Assume that $c(A,B) = \min_{\lambda \epsilon C}[\sigma_{min} H(\lambda)] = \sigma_{min} H(\lambda_0)$ and assume $(\tilde{x},\tilde{\lambda})$ minimizes $F(x,\lambda)$. Observe that $[\sigma_{min} H(\lambda)]^2$ is the minimum eigenvalue of $R(\lambda) = H(\lambda)(H\lambda)^*$,see [Go-1], and hence $[c(A,B)]^2$ is the minimum eigenvalue of $R(\lambda_0)$.

Lemma(2.4.1) implies that there exists $x_0 \neq 0$ such that

(2.4.17) $$c(A,B)^2 = \frac{F(x_0,\lambda_0) - K(1 - \|x_0\|)^2}{\|x_0\|^2} = [\sigma_{min} H(\lambda_0)]^2.$$

Moreover, by the definition of λ_0 it follows that

(2.4.18) $$[\sigma_{min} H(\lambda_0)]^2 \leq [\sigma_{min} H(\tilde{\lambda})]^2.$$

By Lemma (2.4.1)

(2.4.19) $$[\sigma_{min} H(\lambda_0)]^2 \leq [\sigma_{min} H(\tilde{\lambda})]^2 = \frac{F(\tilde{x},\tilde{\lambda}) - K(1 - \|\tilde{x}\|)^2}{\|\tilde{x}\|^2}.$$

Define $S = \{x \epsilon C : \|x\| = \|\tilde{x}\|\}$. Clearly,

(2.4.20) $$F(\tilde{x},\tilde{\lambda}) \leq \min_{x \epsilon S} F(x,\lambda_0).$$

Thus,

$$(2.4.21) \qquad \frac{F(\tilde{x},\tilde{\lambda}) - K(1 - \|\tilde{x}\|)^2}{\|\tilde{x}\|^2} \le \frac{F(x,\lambda_0) - K(1 - \|\tilde{x}\|)^2}{\|\tilde{x}\|^2},$$

which in view of (2.3.13) implies that

$$(2.4.22) \qquad \frac{\tilde{x}^*H(\tilde{\lambda})H(\tilde{\lambda})\tilde{x}}{\tilde{x}^*\tilde{x}} \le \min_{x \epsilon S} \frac{x^*H(\lambda_0)H(\lambda_0)^*x}{x^*x} = [\sigma_{min}H(\lambda_0)]^2.$$

Thus we have

$$(2.4.23) \qquad \frac{F(\tilde{x},\tilde{\lambda}) - K(1 - \|\tilde{x}\|)^2}{\|\tilde{x}\|^2} = \frac{\tilde{x}^*H(\tilde{\lambda})H(\tilde{\lambda})\tilde{x}}{\tilde{x}^*\tilde{x}} \le [c(A,B)]^2.$$

Therefore, $[c(A,B)]^2 = \frac{F(\tilde{x},\tilde{\lambda}) - K(1 - \|\tilde{x}\|)^2}{\|\tilde{x}\|^2}$ and the proof is complete.

Theorem(2.4.3) Let $s(A,B) = \min_{\lambda \epsilon C, re(\lambda) \ge 0} [\sigma_{min}H(\lambda)]$ and assume that $K > s(A,B)$. If $\tilde{x}$ and $\tilde{\lambda}$ minimize

$$(2.4.24) \qquad F(x,\lambda) = x^*H(\lambda)H(\lambda)^*x + k(1 - \|x\|)^2,$$

over $C^n x C^+$, then

$$(2.4.25) \qquad s(A,B)^2 = \frac{F(\tilde{x},\tilde{\lambda}) - K(1 - \|\tilde{x}\|)^2}{\|\tilde{x}\|^2}.$$

Proof:

The proof is completely analogous to the above proof.

2.5 Measure of Stabilizability for Infinite Dimensional Systems

In the last section a method was given which can be used to compute the measure of stabilizability for finite dimensional systems. It is possible to compute the measure of stabilizability for a certain class of infinite dimensional systems. A theorem will be given which will allow the construction of an upper bound on the measure of stabilizability.

Theorem (2.5.1) Suppose A generates a C_0 semigroup on a Hilbert space H and B is a linear operator form the Hilbert Space U to the Hilbert space H. If the pair (A,B) is stabilizable, then γ defined by

$$\gamma^2 = \inf_{\substack{x \in D(A) \\ \mathrm{Re}\,\lambda \geq 0 \\ \|x\|=1}} <(A-\lambda I)x,(A-\lambda I)x\rangle + <B^*x,B^*x>, \tag{2.5.1}$$

is an upper bound on s(A,B).

Proof:

Let $\epsilon>0$ be given. Since $0 \leq \gamma^2$ is an infimun, there exists $x_0 \epsilon D(A)$, $\|x_0\|=1$, and and λ_0 with $\mathrm{Re}(\lambda_0) \geq 0$, such that

$$\gamma^2 \leq <(A-\lambda_0 I)x_0,(A-\lambda_0 I)x_0> + <B^*x_0,B^*x_0> < \gamma^2+\epsilon. \tag{2.5.2}$$

Let $S_0 = \text{span}\{x_0\}$. Decompose H by $H = S_0 \oplus S_0^{\perp}$. Also suppose that $\delta A: H \rightarrow H$ and $\delta B: U \rightarrow H$ are bounded linear operators defined by:

(2.5.3) $$\delta A = (-A + \lambda_0 I)|_{S_0} \oplus 0|_{S_0^{\perp}}$$

(2.5.4) $$\delta B^* = -B^*|_{S_0} \oplus 0|_{S_0^{\perp}}.$$

Note that δA and δB satisfy

(2.5.5) $$\|\delta A, \delta B\|_2^2 \leq \gamma^2 + \epsilon.$$

If the perturbed system $(\hat{A}, \hat{B})$ is defined by

(2.5.6) $$\hat{A} = A + \delta A, \text{ and } \hat{B} = B + \delta B,$$

then for any bounded linear operator $K: H \rightarrow U$, $\hat{A} + \hat{B}K$ forms a C_0 semigroup $\hat{T}(t)$. Fix $K: H \rightarrow U$. Since H is a Hilbert space, Corollary 10.6 on page 41 of [P-1] implies that $\hat{A}^* + K^*\hat{B}^*$ generates the adjoint semigroup $\hat{T}^*(t)$. By Theorem 10.4 on page 39 of [P-1], $\|\hat{T}(t)\| \leq Me^{\omega t}$ implies that $\|\hat{T}^*(t)\| \leq Me^{\omega t}$. Thus we only need to establish that $\hat{A}^* + K^*\hat{B}^*$ does not generate an exponentially stable semigroup $\hat{T}^*(t)$. By Theorem 8.1 on page 33 of [P-1]

(2.5.7) $$\hat{T}^*(t)x_0 = \lim_{n \rightarrow \infty} (I - \tfrac{t}{n}(\hat{A}^* + K^*\hat{B}^*))^{-n} x_0.$$

However, note that $x_0 \in D(A^*)$ and for all $n \geq 1$,

(2.5.8) $$[I - \tfrac{t}{n}(\hat{A}^* + K^*\hat{B}^*)]x_0 = (I - \tfrac{t}{n}\lambda_0^*)x_0.$$

Thus we have

(2.5.9) $$\hat{T}^*(t)x_0 = \lim_{n\to\infty}(I - \tfrac{t}{n}\lambda_0^*)^{-n}x_0 = e^{\lambda_0^* t}x_0.$$

Thus we see that

(2.5.10) $$\|\hat{T}^*(t)\| \geq e^{re(\lambda_0^*)t} \geq 1,$$

which implies that $\hat{T}(t)$ is not exponentially stable for this choice of K. However K was arbitary, so (A,B) is unstabilizable and hence $s^2(A,B) \leq \gamma^2 + \epsilon$. Since $\epsilon > 0$ is arbitary, $s^2(A,B) \leq \gamma^2$.

The following theorem can be found on page 293 of [K-1].

Theorem(2.5.2) Suppose that A is a self-adjoint operator on a Hilbert space H with compact resolvent. Also assume A has simple eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots$. Futher assume that $\lim_{h\to\infty} \lambda_{h-1} - \lambda_h = \infty$. For any bounded perturbation δA to the operator A, the operator $A + \delta A$ is closed with compact resolvent and the spectrum of $A + \delta A$ consists of eigenvalues, which can be indexed as $\{\mu_{0k}, \mu_h\}$, where $k = 1, \ldots, m < \infty$ and $h = n+1, \ldots$ with $n \geq 0$, and $|\mu_h - \lambda_h|$ is uniformly bounded.

The following definitions are a summary of definitions given in [C-1].

Definition(2.5.3) Suppose A is an operator on a Hilbert space H and $\sigma(A)$ denotes the spectrum of A. For a fixed $\delta > 0$

$$\sigma_u(A) = \sigma(A) \cap \{\lambda \epsilon C : re(\lambda) \geq -\delta\}, \tag{2.5.11}$$

$$\sigma_s(A) = \sigma(A) \cap \{\lambda \epsilon C : re(\lambda) < -\delta\}. \tag{2.5.12}$$

Definition(2.5.4) A C_0-semigroup T(t) with generator A is said to satisfy the spectrum determined growth assumption if

$$\sup \operatorname{Re} \sigma(A) = \lim_{t \to \infty} \frac{\log \| T(t) \|}{t} = \omega_0. \tag{2.5.13}$$

Definition(2.5.5) An operator A is said to satisfy the spectrum decomposition assumption if the set $\sigma_u(A)$ is bounded and is separated from the set $\sigma_s(A)$ in such a way that a rectifiable, simple, closed curve can be drawn so as to enclose an open set containing $\sigma_u(a)$ in its interior and $\sigma_s(A)$ in its exterior.

The following result may be found on page 178 in [K-1].

Lemma(2.5.6) Suppose A is a linear operator on a Hilbert space H satisfies the spectrum decomposition assumption. Then the following three conditions hold:

(2.5.14) $$H = H_u \oplus H_s\ ,$$

where H_u is finite dimensional. If P denotes the projection onto H_u, then

(2.5.15) $$PD(A) \subset D(A),$$

(2.5.16) $$AH_s \subset H_s,\ AH_u \subset H_u,\ A_s = AP,\ A_u = A(I\text{-}P),$$

and A is a bounded operator on Z_u. Furthermore $\sigma(A_s) = \sigma_s(A)$ and $\sigma(A_u) = \sigma_u(A)$.

The following theorem and its proof can be found in [C-1].

Theorem(2.5.7) The pair (A,B), where B is bounded, is exponentially stabilizable on a Hilbert space H if the generator A satisfies the spectrum decomposition assumption, the semigroup $T_s(t)$ (generated by A_s) satisfies the spectrum determined growth assumption, and the projection onto Z_u is exponentially stabilizable by feedback control $U = D_u Z_u$, where D_u is a bounded operator with domain in Z_u.

Theorem(2.5.8) Suppose that A is a self-adjoint operator on a Hilbert space H with compact resolvent and A generates an analytic semigroup. Further assume A has simple eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots$ where $\lim_{n \to \infty} \lambda_{n-1} - \lambda_n = \infty$. If B is a bounded operator then, the measure of stabilizability s(A,B) of the system (A,B) is given by

(2.5.17) $$s(A,B)^2=\gamma^2=\inf_{\substack{x\epsilon D(A)\\ \text{Re }\lambda\geq 0\\ \|x\|=1}} \langle (A-\lambda I)x,(A-\lambda I)x\rangle+\langle B^*x,B^*x\rangle$$

Proof:

Suppose to the contrary that (δA, δB) is a bounded perturbation to (A,B) and makes (A,B) unstabilizable where $\|[\delta A,\delta B]\|_2^2<\gamma^2$. Since $A+\delta A$ generates an analytic semigroup [P-1], it satisfies the spectrum determined growth assumption [C-1]. By Theorem (2.5.2), the spectrum of $A+\delta A$ consists of separated eigenvalues. Thus $A+\delta A$ satifies the spectrum decompositon assumption. In light of Theorem(2.5.7), the projection of the system ($A+\delta A,B+\delta B$) onto the space H_u (see Equations (2.5.16) and (2.5.17)) is not stabilizable, because stabilizability on this space would imply that the system was stabilizable on H. This implies that ($P(A+\delta A)P,P(B+\delta B)$), which is defined on the finite dimensional space H_u, is not stabilizable. Since H_u is finite dimensional, the finite dimensional measure of stabilizability must be 0 for ($P(A+\delta A)P,P(B+\delta B)$). By equation (2.3.4) there must exist $x_0 \in H_u$ and λ_0 with $re(\lambda_0)\geq 0$ such that

(2.5.18) $$\langle P(A+\delta A - \lambda_0 I)Px_0,P(A+\delta A - \lambda_0 I)Px_0\rangle + \langle (P(B+\delta B))^*x_0,(P(B+\delta B))^*x_0\rangle=0.$$

Since P is the projection onto H_u, $x_0\in H_u$ and $(A+\delta A)H_u \subset H_u$, it follows that

(2.5.19) $$\langle (A+\delta A-\lambda_0 I)x_0,(A+\delta A-\lambda_0 I)x_0\rangle+ \langle P(B+\delta B)^*x_0,P(B+\delta B)^*x_0\rangle=0.$$

Therefore,

(2.5.20) $$\|(A-\lambda_0 I)x_0\| \leq \|(A+\delta A-\lambda_0 I)x_0\| + \|\delta A x_0\| = \|\delta A x_0\|$$

and

(2.5.21) $$\|PB^*x_0\| \leq \|P(B + \delta B)^*x_0\| + \|P(\delta B)^*x_0\| = \|P(\delta B)^*x_0\|.$$

Thus, we have that

(2.5.22) $$\langle (A - \lambda_0 I)x_0,(A - \lambda_0 I)x_0\rangle + \langle B^*x_0,B^*x_0\rangle \leq \|[\delta A,\delta B]\|_2^2 < \gamma^2,$$

which contradicts the definition of γ^2 in equation (2.5.8).

Corollary(2.5.9) Suppose that A is a self-adjoint operator on a Hilbert space H with compact resolvent and A generates an analytic semigroup. Further assume A has simple eigenvalues $0 \geq \lambda_n \geq \lambda_{n+1}$ where $\lim_{n\to\infty} \lambda_{n-1} - \lambda_n = \infty$. If B is a bounded operator, then the measure of stabilizability s(A,B) of the system (A,B) is given by

(2.5.23) $$s(A,B)^2 = \inf_{\substack{x\epsilon D(A) \\ \|x\|=1}} \langle Ax,Ax\rangle + \langle B^*x,B^*x\rangle.$$

Proof:

Since A is self-adjoint with a countable number of simple eigenvalues, by the spectral theorem [W-1] we have for $y \in D(A)$

$$\text{(2.5.24)} \qquad Ay = \sum_{n=0}^{\infty} \langle y, x_n \rangle x_n,$$

where $\{x_n\}$ is the orthonormal basis consisting of the eigenvectors corresponding to the eigenvalues $\{\lambda_n\}$. This implies for $y \in D(A)$

$$\text{(2.5.25)} \qquad \langle (A-\lambda I)y, (A-\lambda I)y \rangle = \sum_{n=0}^{\infty} |\lambda_n - \lambda|^2 \langle y, x_n \rangle^2 .$$

Since A is self-adjoint, the eigenvalues λ_n are all real. Suppose α has a nonzero imaginary part, and let $\beta = \text{re}(\alpha)$, then

$$\text{(2.5.26)} \qquad |\lambda_n - \alpha| > |\lambda_n - \beta|.$$

This implies

$$\text{(2.5.27)} \qquad \langle (A-\beta I)x, (A-\beta I)x \rangle \leq \langle (A-\alpha I)x, (A-\alpha I)x \rangle .$$

If $\nu > 0$ then, since $\lambda_n \leq 0$,

$$\text{(2.5.28)} \qquad |\lambda_n - \nu| > |\lambda_n - 0|.$$

Thus we have,

(2.5.29) $$\langle Ax,Ax\rangle \leq \langle (A-\nu I)x,(A-\nu I)x\rangle .$$

Therefore,

(2.5.30) $$\gamma^2 = \inf_{\substack{x\epsilon D(A) \\ \mathrm{Re}\ \lambda \geq 0 \\ \|x\|=1}} \langle (A-\lambda I)x,(A-\lambda I)x\rangle + \langle B^*x,B^*x\rangle$$

(2.5.31) $$= \inf_{\substack{x\epsilon D(A) \\ \|x\|=1}} \langle Ax,Ax\rangle + \langle B^*x,B^*x\rangle .$$

2.6 The Relationship to POES

The following is a summary of the material related to POES in [Ba-1]. We suppose throughout this section that H and U are Hilbert spaces, that $A:H\rightarrow H$ is the infinitesimal generator of a C_0-semigroup T(t) on H and that $B:U\rightarrow H$ is a bounded linear operator. We consider a control system in H given by

(2.6.1) $$\dot{y}(t) = Ay(t) + Bu(t), \quad t>0$$

and an associated performance measure

(2.6.2) $$J(y_0,u) = \int_0^\infty \{\langle Dy(t), y(t)\rangle + \langle Qu(t),u(t)\rangle\}\, dt,$$

where $D:H\rightarrow H$ $Q:U\rightarrow U$ are bounded linear self-adjoint operators that satisfy $D\geq 0$, $Q>0$. Our fundamental abstract linear optimal regulator problem can then be stated as

($\mathcal{R}$) Minimize J (y_0,u) over $u \in L^2(0,\infty;U)$ subject to $y=y(\cdot;u)$ satisfying Equation (2.6.1).

Definition(2.6.1) A function $u \in L^2(0,\infty;U)$ is an admissible control for the initial state $y_0 \in H$ if $J(y_0,u)$ is finite.

Definition(2.6.2) A bounded linear operator $\Pi:H \rightarrow H$ is called a solution of the algebraic Riccati equation (A.R.E.) if Π maps D(A) into D(A^*) and satisfies on H the equation

$$(2.6.3) \qquad A^*\Pi + \Pi A - \Pi B Q^{-1} B^* \Pi + D = 0.$$

By Corollary 10.6 on page 37 of [P-1] A^* generates C_0-semigroup $T(t)^*$ which is adjoint to T(t). Note that if Π satisfies (2.6.3) on D(A), which is dense in H, then Equation (2.4.3) can be taken as an equation on H since $\Pi BQ^{-1}B^*\Pi - D$ is bounded so that $A^*\Pi + \Pi A$ has a bounded extension to all of H.

The following result is taken from [G-1].

Theorem(2.6.3) Let A, B, Q, D be as given above. Then there exists a nonnegative self-adjoint solution Π of the algebraic Riccati equation (2.6.3) if and only if for each $y_0 \in H$, there exists an admissible control. The unique optimal control and corresponding trajectory for ($\mathcal{R}$) are given by

$$\bar{u}(t) = -Q^{-1}B^*\Pi_\infty \bar{y}(t), \tag{2.6.4}$$

$$\bar{y}(t) = S(t)y_0, \tag{2.6.5}$$

where Π_∞ is the minimal nonnegative self-adjoint solution of A.R.E. equation (2.6.3) and $S(t)$ is the C_0-semigroup generated by $A - BQ^{-1}B^*\Pi_\infty$. If $|y(t;u)| \to 0$ as $t \to \infty$ for any admissible control (guaranteed for example by the condition $D>0$), then Π_∞ is the unique nonnegative self-adjoint solution of the A.R.E. If $D > 0$, then we also have that $S(t)$ is uniformly exponentially stable.

The term minimal for a self-adjoint operator is in reference to the usual ordering of self-adjoint nonnegative operators on a Hilbert space. Note that the minimal solution Π_∞ of Equation (2.4.3) can be obtained as the limit of a sequence of Riccati operators for associated finite interval regulator problems (see [G-1]) in a manner analogous to the usual procedure for finite dimensional state space regulator problems.

We next formulate a sequence of approximate regulator problems and present a convergence result for the corresponding Riccati operators. Let H^N, $N = 1,2,\cdots$, be a sequence of finite dimensional linear subspaces of H and $P^N:H \to H^N$ be the canonical orthogonal projections. Assume that $T^N(t)$ is a sequence of C_0-semigroups on H^N with infinitesimal generators $A^N:H^N \to H^N$. Given operators $B^N:U \to H^N$ and $D^N:H^N \to H^N$, we consider the family of regulator problems:

$$(\mathcal{R}^N) \text{ Minimize } J^N(y^N(0),u) \text{ over } u \in L^2(0,\infty;U), \tag{2.6.6}$$

where

$$\dot{y}^N(t) = A^N y^N(t) + B^N u(t), \quad t > 0 \tag{2.6.7}$$

$$y^N(0) = y_0^N \equiv P^N y_0, \tag{2.6.8}$$

and

$$J^N(y^N(0),u) = \int_0^\infty \{\langle D^N y^N(t), y^N(t)\rangle + \langle Qu(t), u(t)\rangle\} \, dt. \tag{2.6.9}$$

We note that since $B^N:U \rightarrow H^N$, the trajectories of (2.6.7) evolve in H^N and consequently $(\mathcal{R}^N)$ is a linear regulator problem in the finite dimensional state space H^N so that finite dimensional control theory is applicable here. We shall need several assumptions in a convergence statement regarding solutions of $(\mathcal{R}^N)$ and $(\mathcal{R})$.

(H1). For each $y_0^N \in H^N$ there exists and admissible control $u^N \in L^2(0,\infty;U)$ for $(\mathcal{R}^N)$ and any admissible control for (2.6.7), (2.6.8) drives the state of (2.6.7) to zero asymptotically.

(H2). (i) For each $z \in H$, we have $T^N(t)P^N z \rightarrow T(t)z$ with the convergence uniform in t on bounded subset of $[0,\infty)$.

(ii) For each $Z \in H$, we have $T^N(t)^* P^N z \rightarrow T(t)^* z$ with the convergence uniform in t on bounded subsets of $[0,\infty]$.

(iii) For each $v \in U$, $B^N v \to Bv$ and for each $z \in H$, $B^{N^*} z \to B^* z$.

(iv) For each $z \in H$, $D^N P^N z \to Dz$.

Note that (H2)(i) implies in particular (take t=0) that $P^N z \to z$ for each $z \in H$ and in this sense we have that the subspaces H^N approximate H.

If assumption (H1) holds, then the optimal control $\bar{u}^N$ for $(\mathcal{R}^N)$ is given in feedback form by

$$\bar{u}^N(t) = -Q^{-1} B^{N^*} \Pi^N y^N(t), \tag{2.6.10}$$

where $\Pi^N : H^N \to H^N$ is the unique nonnegative self-adjoint solution of the algebraic Riccati equation on H^N. Thus,

$$A^{N*}\Pi + \Pi^N A^N - \Pi^N B^N Q^{-1} B^{N*} \Pi^N + D^N = 0, \tag{2.6.11}$$

and $\bar{y}^N$ is the corresponding solution of (2.6.6) with $u = \bar{u}^N$. Moreover,

$$J^N(y_0^N, \bar{u}^N) = \langle \Pi^N y_0^N, y_0^N \rangle. \tag{2.6.12}$$

A proof of the following theorem is given in [Ba-1].

Theorem(2.6.4) Suppose (H1), (H2) hold, $Q > 0$, $D \geq 0$ and $D^N \geq 0$ and let Π denote the unique nonnegative self-adjoint Riccati operators on H^N for the problem $(\mathcal{R}^N)$. Further assume that a unique nonnegative self-adjoint Riccati operator on H for the problem $(\mathcal{R})$ exists. Let

$S(t)$ and $S^N(t)$ be the semigroups generated by $A - BQ^{-1}B^*\Pi$ and $A^N - B^NQ^{-1}B^{N*}\Pi^N$ on H and H^N, respectively, and suppose $|S(t)z| \to 0$, $t \to \infty$, for all $z \in H$. If there are positive constants M_1, M_2 and ω independent of N and t such that

(2.6.13) $$|S^N(t)|_{H^N} \le M_1 e^{-wt} \text{ for } t \ge 0, \quad N = 1,2,\cdots,$$

and

(2.6.14) $$|\Pi^N|_{H^N} \le M_2,$$

then

(2.6.15) $$\Pi^N P^N z \to \Pi z \quad \text{for every } z \in H,$$

(2.6.16) $$S^N(t)P^N z \to S(t)z \quad \text{for every } z \in H,$$

where the convergence is uniform in t on bounded subsets of $(0,\infty)$, and

(2.6.17) $$|S(t)| \le M_1 e^{-\omega t} \quad \text{for } t \ge 0.$$

The above theorem places conditions on the stabilizability of the approximating semigroups in order to get convergence. The following condition is sufficent to ensure these stability conditions are satisfied. The following is the definiton of POES (preservation of expontential stabilizability) which is given in [Ba-1].

Definition(2.6.5) (POES) The condition POES is defined as follows: Suppose that (A,B) is stabilizable by the feedback operator K. Then there exists an integer N_0 such that for all $N \geq N_0$ the pairs $(A^N, P^N B)$ are uniformly exponentially stabilizable by the feedback operator K, i.e., there exist positive constants (independent of N) M_S and $\omega_S > 0$ such that the C_0-semigroups generated by $A^N + P^N BK$ satisfy $\|T_S^N(t)\| \leq M_S e^{-\omega_S t}$ for all $N \geq N_0$ and $t \geq 0$.

The following theorem shows that POES implies that the finite dimensional measure of stabilizability for the approximating systems does not go to zero.

Theorem(2.6.6) Suppose that (A^N, B^N) is a matrix representation of an approximation scheme as defined by (H1) and (H2). If POES is satisfied, then there exists $\alpha > 0$ such that $s(A^N, B^N) \geq \alpha$, where $s(A^N, B^N)$ is the measure of stabilizability of the finite dimensional system (A^N, B^N).

Proof:

Since POES is satisfied, there exists an bounded linear operator $K: U \rightarrow H$ integer N_0 such that for all $N \geq N_0$ the pairs (A^N, B^N) are uniformly exponentially stabilizable by the operator K. Let M_S and ω_S be as above and define $q = \max\{1, \|K\|\}$. Let $(\delta A^N, \delta B^N)$ be a perturbation to (A^N, B^N) and $\alpha = \frac{\omega_S}{\sqrt{8 q M_S}}$. If

(2.6.18) $$\|[A^N, \delta B^N]\|_2 < \alpha,$$

then

$$(2.6.19) \qquad \left\|\delta A^N + \delta B^N K\right\|_2 \leq \left[\left\|\delta A^N\right\|_2 + \left\|\delta B^N K\right\|_2\right]$$

$$(2.6.20) \qquad \leq q\left[\left\|\delta A^N\right\|_2 + \left\|\delta B^N\right\|_2\right] \leq \sqrt{2}q\left\|[A^N,\delta B^N]\right\|_2 \leq \frac{\omega_s}{2M_s}.$$

If $\hat{T}^N(t)$ is the C_0-semigroup generated by $A^N + \delta A^N + (B^N + \delta B^N)K$, then by Lemma(2.2.1),

$$(2.6.21) \qquad \left|\hat{T}^N(t)\right| \leq M_s e^{\left[-\omega_s + M_s \frac{\omega_s}{2M_s}\right]t} = M_s e^{-\frac{\omega_s}{2}t}.$$

Thus $s(A^N,B^N) \geq \alpha$ and the proof is complete.

Note that this provides a means to check POES numerically provided that one can compute $s(A^N,B^N)$.

Chapter III

3.1 Introduction

In this chapter we present two methods for constructing balanced finite dimensional approximating systems. The first method was introduced by Curtain and Glover [C-2] and may be described as the process of truncating balanced infinite dimensional realizations. The second approach is based on balancing finite dimensional approximations of the dynamic equations, e.g. a finite element model. The second approach is fairly straight forward to implement, whereas the first method can become rather complex to apply.

3.2 Infinite Dimensional Realizations

A balanced realization of a finite dimensional system is defined as a realization of the system for which the controllability and the observability grammians are diagonal and equal [M-1]. Balanced realizations for infinite dimensional systems are presented in [C-2]. We shall review the basic definitions and results found in [C-2]. We limit our discussion to finitely many inputs and outputs.

Consider the class of linear infinite dimensional systems defined by the following input-output map:

$$y(t) = \int_0^t h(t-s)u(s)ds, \tag{3.2.1}$$

where the outputs $y \epsilon L_2(0,\infty;R^m)$, the inputs $u \epsilon L_2(0,\infty;R^m)$, and the impulse response function h satisfies

(3.2.2) $$h \in L_1 \cap L_2(0,\infty;R^{kxm}) \text{ and } t^{\frac{1}{2}} h(t) \in L_2(0,\infty,R^{kxm}).$$

This induces the Hankel operator Γ: $L_2(0,\infty;R^m) \rightarrow L_2(0,\infty;R^k)$ defined by

(3.2.3) $$(\Gamma u)(t) = \int_0^\infty h(t+s)u(s)ds.$$

Assumption(3.2.2) implies that Γ is compact and Hilbert Schmidt on $L_2(0,\infty)$ and is a continuous operator on each of the spaces $L_p(0,\infty)$ $1 \leq p < \infty$ and on $\tilde{C}^1(0,\infty)$, which is the space of continuously differentiable functions with the norm

(3.2.4) $$\|f\|_1 = \|f\|_\infty + \int_0^\infty |\dot{f}(s)|ds.$$

The operator $\Gamma^*\Gamma$ is compact and positive on $L_2(0,\infty;\mathbf{R}^m)$ and has countably many positive eigenvalues $\sigma_1^2 \geq \sigma_2^2 \geq \ldots \sigma_i^2 \geq \ldots$, where the $\sigma_i \geq 0$ are the singular values of Γ. If v_i and w_i are the corresponding normalized eigenvectors of $\Gamma^*\Gamma$ and $\Gamma\Gamma^*$, respectively, then (v_i,w_i) are called the Schmidt pairs of Γ and

(3.2.5) $$\begin{cases} \Gamma v_i = \sigma_i w_i \\ \Gamma^* w_i = \sigma_i v_i \qquad ; i = 1,2,\ldots . \end{cases}$$

Since w_i is an eigenvector of $\Gamma\Gamma^*$, and $\Gamma\Gamma^*$ is compact in the spaces $L_2(0,\infty)$, $L_p(o,\infty)$; $p > 1$ and $C^1(0,\infty)$, it follows that, see [C-2], (3.2.6) $$w_i \epsilon L_1 \cap L_2 \cap \tilde{C}^1(0,\infty;R^k)$$

and

(3.2.7) $$v_i \epsilon L_1 \cap L_2 \cap C^1(0,\infty;R^m).$$

The Hankel kernels can be expressed in terms of their singular values and Schmidt pairs. A proof of the following Lemma can be found in [C-2].

Lemma(3.2.1) With the above notation the following relationships hold:

(3.2.8) $$(\Gamma u)(t) = \int_0^\infty h(t+s)u(s)ds = \sum_{i=1}^\infty w_i(t)\sigma_i \int_0^\infty v_i^*(s)u(s)ds;$$

(3.2.9) $$(\Gamma^* y)(t) = \int_0^\infty h^*(t+s)y(s)\,ds = \sum_{i=1}^\infty v_i(t)\sigma_i \int_0^\infty w_i^*(s)y(s)ds;$$

(3.2.10) $$h^*(t+s) = \sum_{i=0}^\infty v_i(t)\sigma_i w_i^*(s) \quad \text{a.e. in s for all } t \geq 0;$$

(3.2.11) $$(\Gamma^* y)(t+s) = \int_0^\infty h^*(t+s+r)y(r)dr = \sum_{i=1}^\infty v_i(t)\sigma_i \int_0^\infty w_i^*(s+r)y(r)dr.$$

For simplicity of exposition we shall assume henceforth that the singular values are distinct, but the proofs can be extended to the case of multiple singular values in the usual way.

Definition(3.2.2) A realization for the impulse response function h on the state space H is the triple (A,B,C), where $C{:}H \rightarrow \mathbf{R}^k$, $B{:}\mathbf{R}^m \rightarrow H$ and A is the infinitesimal generator of a C_0-

semigroup, T(t), on H such that $h(t) = CT(t)B$ a.e. in t.

If $B:\mathbf{R}^m \rightarrow H$ and $C:H \rightarrow \mathbf{R}^k$, are bounded operators and T(t) is exponentially stable, then h satisfies the assumption (3.2.2) and the reachability and observability operators $\underline{B}$ and $\underline{C}$ defined below are bounded:

$$\text{(3.2.12)} \qquad \underline{B}:L_2(0,\infty) \rightarrow H, \quad \underline{B}u = \int_0^\infty T(t)Bu(t)dt,$$

$$\text{(3.2.13)} \qquad \underline{C}:H \rightarrow L_2(0,\infty), \quad [\underline{C}x](t) = CT(t)x.$$

With the above definitions, the Hankel operator Γ is defined by

$$\text{(3.2.14)} \qquad \Gamma:L_2(0,\infty) \rightarrow L_2(0,\infty), \quad \Gamma = \underline{C}\underline{B} = C\int_0^\infty T(t+s)Bu(s)ds.$$

Since $\underline{B}$ and $\underline{C}$ are bounded, the infinite dimensional controllability grammian P and observability grammian Q can be defined by:

$$\text{(3.2.15)} \qquad P = \underline{B}\underline{B}^* = \int_0^\infty T(t)BB^*T^*(t)dt$$

$$\text{(3.2.16)} \qquad Q = \underline{C}^*\underline{C} = \int_0^\infty T^*(t)C^*CT(t)dt,$$

respectively.

The following definition for the balanced realization for Γ may be found in [C-2].

Definition(3.2.3) (T(t),B,C) is a balanced realization for the Hankel operator Γ if it has bounded reachability and observability operators defined by (3.2.12) and (3.2.13), (3.2.14) holds and the controllability and the observability grammians are equal to the same positive diagonal operator.

The following two results may be found in [C-2].

Theorem(3.2.4) If a T(t) is a stable C_0-semigroup, B and C are bounded control and observation operators, then $(\tilde{T}(t),B,\tilde{C})$ defines a balanced realization for (T(t),B,C) on ℓ^2 where

(3.2.17) $$\tilde{T}_{ij}(t) = \sqrt{\frac{\sigma_j}{\sigma_i}} \int_0^\infty w_i^*(s)w_j(s+T)ds$$

(3.2.18) $$\tilde{C}=\left[\sqrt{\sigma_1}w_1(0),\sqrt{\sigma_2}w_2(0), \ldots, \sqrt{\sigma_i}w_i(0),\ldots\right],$$

and

(3.2.19) $$\tilde{B}=\left[\sqrt{\sigma_1}v_1(0),\sqrt{\sigma_2}v_2(0), \ldots, \sqrt{\sigma_i}v_i(0),\ldots\right].$$

Theorem(3.2.5) If $(\tilde{T}(t),\tilde{B},\tilde{C})$ is the balanced realization defined by (3.2.17)-(3.2.19) and $\tilde{A}$ is the infinitesimal generator of $\tilde{T}(t)$, then

(3.2.20) $$\tilde{A}_{ij}=\left[\frac{\sigma_j}{\sigma_i}\right]^{\frac{1}{2}}\int_0^\infty w_i^*(s)\dot{w}_j(s)ds,$$

(3.2.21) $$\tilde{C}=\left[\sqrt{\sigma_1}w_1(0),\sqrt{\sigma_2}w_2(0),\ \ldots,\ \sqrt{\sigma_i}w_i(0),\ldots\right],$$

and

(3.2.22) $$\tilde{B}=\left[\sqrt{\sigma_1}v_1(0),\sqrt{\sigma_2}v_2(0),\ \ldots,\ \sqrt{\sigma_i}v_i(0),\ldots\right].$$

3.3 Single-Input Single-Output Finite Dimensional Case; Symmetric A

In this section the finite dimensional case will be examined where there is only one input and one output and the matrix A is symmetric. We are restricting ourselves to this case since the infinite dimensional systems we consider below are generated by self-adjoint operators. Consequently, most finite dimensional approximations of A are self-adjoint, and this greatly simplifies the computational problems. In [M-1] an introduction is given to the balancing transformation and some of the uses of it. For the special case of single-input single-output systems, with symmetric A, a characterization of the balanced system can be obtained rather easily.

Definition(3.3.1) If T is an invertible nxn matrix and (A,B,C) is a finite dimensional system, where A is nxn, B is nx1 and C is nx1, then the system $(\tilde{A},\tilde{B},\tilde{C})$ is called the system transformed by T if

(3.3.1) $$\tilde{A}=T^{-1}AT,$$

(3.3.2) $$\tilde{B}=T^{-1}B,$$

and

(3.3.3) $$\tilde{C}=CT.$$

The following algorithm is given in [L-1] for the computation of the balanced realization of a controllable, observable finite dimensional system (A,B,C) where A has eigenvalues with negative real parts.

Let L_r and L_o denote the Cholesky factors of the control and observation grammians P and Q, i.e.

(3.3.4) $$P=L_rL_r^*$$

and

(3.3.5) $$Q=L_oL_o^*,$$

respectively. Let $U\Lambda V^*$ be the singular value decomposition of $L_o^*L_r$, i.e.

(3.3.6) $$U\Lambda V^*=L_o^*L_r.$$

Then, the balanced realization $(\tilde{A},\tilde{B},\tilde{C})$ of the system (A,B,C) is given by

(3.3.7) $$\tilde{A}=\Lambda^{-\frac{1}{2}}U^*L_o^*AL_rV\Lambda^{-\frac{1}{2}},$$

(3.3.8) $$\tilde{B}=\Lambda^{-\frac{1}{2}}U^*L_o^*B,$$

and

(3.3.9) $$\tilde{C}=CL_rV\Lambda^{-\frac{1}{2}}.$$

The following sequence of lemmas is needed to establish the properties of a finite dimensional balanced realization of finite dimensional system with a single input and single output where the matrix A is symmetric.

Lemma(3.3.2) If (A,B,C) is controllable and observable, A is a stable diagonal matrix and $B^*=CI_c$ where I_c is a diagonal matrix then $Q=I_cPI_c$.

Proof:

Let

$$P = \int_0^\infty e^{As} BB^* e^{As} ds, \tag{3.3.10}$$

denote the controllability grammian. Multiply both sides of (3.3.10) by I_c on the left and I_c on the right to obtain

$$I_c P I_c = \int_0^\infty I_c C^{As} BB^* e^{As} I_c ds. \tag{3.3.11}$$

Since A is a diagonal matrix, e^{As} is a diagonal matrix and hence commutes with I_c. Thus, it follows that

$$\begin{aligned} I_c P I_c &= \int_0^\infty e^{As} [I_c B][B^*] I_c e^{As} ds \\ &= \int_0^\infty e^{As} C^* C e^{As} ds = Q. \end{aligned} \tag{3.3.12}$$

Lemma(3.3.3) If (A,B,C) is a controllable and observable system with A, a stable diagonal matrix, and $B^* = CI_c$ where I_c is a diagonal matrix, then for the balanced system $(\tilde{A},\tilde{B},\tilde{C})$, $\tilde{A}$ is a stable symmetric matrix and $B^* = C$.

Proof:

In this proof we refer to the algorithm defined by (3.3.4)-(3.3.9), for the computation of the balanced realization of the system (A,B,C). By Lemma(3.3.2) the Cholesky decompositions in Equations (3.3.4) and (3.3.5) for the control and observation grammians P and Q are related as follows:

(3.3.13) $$P=LL^*$$

(3.3.14) $$Q=I_cLL^*I_c.$$

Let $U^*\Lambda U$ be the singular value decomposition of L^*I_cL, i.e.

(3.3.15) $$L^*I_cL = U^*\Lambda U.$$

Hence by Equation (3.3.7), $\tilde{A}$ becomes

(3.3.16) $$\tilde{A}=\Lambda^{-\frac{1}{2}}U^*L^*I_cALU\Lambda^{-\frac{1}{2}}$$

which is obviously symmetric, since I_cA is diagonal. By Equation (3.3.8) $\tilde{B}$ becomes

(3.3.17) $$\tilde{B} = \Lambda^{-\frac{1}{2}}U^*L^*I_cB = \Lambda^{-\frac{1}{2}}U^*L^*C^*.$$

By Equation (3.3.9) $\tilde{C}$ becomes

(3.3.18) $$\tilde{C}=CLU\Lambda^{-\frac{1}{2}}.$$

Clearly, $\tilde{B}^*=\tilde{C}$.

Lemma(3.3.4) If (A,B,C) is a controllable and observable system with A, a stable diagonal matrix, B is nx1, and C is 1xn, then the system can be transformed into a system $(\acute{A},\acute{B},\acute{C})$ where $\acute{B}=\acute{C}I_c$ where I_c is a diagonal matrix with the diagonal elements being either 1 or -1.

Proof:

Since A is diagonal and the system is both controllable and observable the matrices B and C cannot have any zero elements [Le-1]. Let H be the diagonal matrix

(3.3.19) $$[h_{ii}]=\left[\sqrt{\left|\frac{b_i}{c_i}\right|}\right].$$

The system which is obtained by using H as the system transformation has

(3.3.20) $$[\acute{b}_i] = \left[b_i \sqrt{\left|\frac{c_i}{b_i}\right|}\right],$$

(3.3.21) $$[\acute{c}_i] = \left[c_i \sqrt{\left|\frac{b_i}{c_i}\right|}\right].$$

Clearly $B^* = CI_c$ for some diagonal matrix I_c with the diagonal elements of I_c being either 1 or-1.

Theorem(3.3.5) If (A,B,C) is a controllable and observable system with A, a stable symmetric matrix, B is nx1, and C is 1xn and $(\tilde{A},\tilde{B},\tilde{C})$ is a balanced realization of (A,B,C), then $\tilde{A}$ is symmetric and $\tilde{B}^* = \tilde{C}$.

Proof:

Since A is symmetric the system can be transformed by the orthogonal matrix U which diagonalizes A producing a system (A',B',C'). By Lemma(3.3.4) there exits a system transformation H which transforms the system (A',B',C') into a system $(\dot{A},\dot{B},\dot{C})$ where $(\dot{A},\dot{B},\dot{C})$ satisfies the hypothesis of Lemma(3.3.3). By Lemma(3.3.3) there exits a transformation R which transforms $(\dot{A},\dot{B},\dot{C})$ into a balanced system $(\tilde{A},\tilde{B},\tilde{C})$ where $\tilde{A}$ is symmetric and $\tilde{B}^* = \tilde{C}$.

Chapter IV

4.1 Introduction

In this chapter two methods will be discussed which can be used to construct finite dimensional approximations to infinite dimesional systems. The first method addressed is the finite element method which is a standard method for approximations of partial differential equations. The second method discussed is the truncated balanced realization method proposed by Curtain and Glover [C-2]. In the final section of this chapter the truncated balanced realization method will be applied to the one dimensional heat equation and compared to the finite element method.

4.2 The Finite Element Method

We shall consider the one dimensional heat equation:

(4.2.1) $$y_t = y_{xx} + b(x)u(t),\ x\epsilon[0,1],\ t>0,$$

with ouput

(4.2.2) $$r(t) = \int_0^1 c(x)y(x,t)dt,$$

and boundary conditions

(4.2.3) $$y(0,t)=y(1,t)=0,\ t>0.$$

Here b(x) and c(x) belong to $L_2(0,1)$. For each $i=1,2,\ldots,N$ let x_i be defined by

(4.2.4) $$x_i=\frac{i}{N+1}, \text{ and}$$

let $h_i^N(x)$ denote the hat function

(4.2.5) $$h_i^N(x)=\begin{cases} (N+1)(x-x_i) & x_{i-1}\le x\le x_i \\ -(N+1)(x-x_{i+1}) & x_i\le x\le x_{i+1} \\ 0 & \text{elsewhere.} \end{cases}$$

If y(x,t) is approximated by

(4.2.6) $$y^N(x,t)=\sum_{i=1}^{N} z_i^N(t)h_i^N(x),$$

then the standard Galerkin procedure ([R-1],[Bu-2]) leads to the finite element approximation

(4.2.7) $$E_E^N\dot{z}^N(t) = F_E^N z^N(t) + G_E^N u(t)$$

(4.2.8) $$y^N(t) = C_E^N z^N(t).$$

Moreover,

(4.2.9) $$E_E^N = [\langle h_i^N, h_j^N \rangle] = \frac{1}{6(N+1)} \begin{bmatrix} 4 & 1 & 0 & & & \\ 1 & 4 & 1 & & 0 & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots \\ & 0 & & 1 & 4 & 1 \\ & & & & 1 & 4 \end{bmatrix},$$

(4.2.10) $$F_E^N = -[\langle \dot{h}_i^N, h_j^N \rangle] = (N+1) \begin{bmatrix} -2 & 1 & 0 & & & \\ 1 & -2 & 1 & & 0 & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots \\ & 0 & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix},$$

(4.2.11) $$G_E^N = \text{col}[\langle b, h_1^N \rangle, \langle b, h_2^N \rangle, \ldots, \langle b, h_N^N \rangle],$$

and

(4.2.12) $$C_E^N = [\langle c, h_1^N \rangle, \langle c, h_2^N \rangle, \ldots, \langle c, h_N^N \rangle].$$

Let

(4.2.13) $$A_N^E = [E_E^N]^{-1} F_E^N, \quad B_E^N = [E_E^N]^{-1} G_E^N$$

and we define the N dimensional finite element model $\Sigma_E^N = (A_E^N, B_E^N, C_E^N)$ by

(4.2.14) $$\dot{z}^N(t) = A_E^N z^N(t) + B_E^N u(t),$$

(4.2.15) $$y^N(t) = C_E^N z^N(t).$$

4.3 The Curtain-Glover Technique of Approximation

In section (4.3) we review the construction of the truncated balanced realization for a large class of infinite dimensional systems. In [C-2] the authors define a scheme for the approximation of the infinite dimensional by simply taking truncations of the infinite dimensional matrices. Referring to Section (3.2) we find that the approximating systems of order M is defined by

(4.3.1) $$A_M = [a_{ij}] = \left[\sqrt{\frac{\sigma_j}{\sigma_i}} \int_0^\infty w_i^*(s)\dot{w}_j(s)ds\right] \quad i,j=1,2,\dots,M,$$

(4.3.2) $$= \left[\sqrt{\frac{\sigma_i}{\sigma_j}} \int_0^\infty v_i^*(s)\dot{v}_j(s)ds\right] \quad i,j=1,2,\dots,M$$

(4.3.3) $$B_M = [b_i] = [\sqrt{\sigma_1}v_1(0),\dots,\sqrt{\sigma_M}v_M(0)]^*,$$

(4.3.4) $$C_M = [c_i] = [\sqrt{\sigma_1}w_1(0),\dots,\sqrt{\sigma_M}w_M(0)],$$

The following result can be found in [C-2].

Theorem(4.3.1) The system defined in Equations (3.2.1)-(3.2.3) is balanced. The controllability and observability grammians are diagonal and both equal to Σ_M where

(4.3.5) $$\Sigma_M = \text{diagonal}(\sigma_1, \ldots, \sigma_M).$$

From this theorem it is seen that the diagonal values of Σ_M are the same as the first M diagonal values of the grammians for the infinite dimensional the balanced realization. This fact will be used later in Section 4.5 when discussing the controllability and stabilizability properties of the approximating systems.

4.4 Stability Properties of the Finite Element Method

In Section (2.5) we saw that a system and the approximation scheme possessing the property POES guaranteed that the Riccati operators for the approximating systems converged to the Riccati operator for the orgional system. The proof of the following theorem can be found in [Ba-1].

Theorem(4.4.1) The finite element scheme for the heat equation defined in Equations (4.2.1)-(4.2.3) satisfies the POES property.

Therefore, we see that the finite element method is a convergent method in the sense that solutions to the approximating L-Q-R problems (i.e. the gain operators) converge to the solution of the infinite dimensional L-Q-R problem. In general, it is not known if the the finite element scheme preserves controllability. However, for special cases it can be shown that the finite element scheme preserves controllability (see [Bu-3]).

4.5 Relationship of the Truncated Balanced Realization Method to Stabilizability

In the last section it was noted that the finite element scheme preserved stabilizability uniformly under approximation. An open question was whether the finite element method preserved controllability. The situation for the balanced realization method of approximation is almost the opposite. For this method it is known that the approximating systems are controllable if the original system was approximately controllable. However, it is not known whether the approximating systems possesses the POES property. The following result may be found in [C-2].

Theorem(4.5.1) If the original system is approximately controllable (c.f. [C-1]), then the finite dimensional approximating systems obtained by truncating the balanced realization method are controllable.

Although it is not know in general if the truncated balanced realization method satisfies POES, it is true that the system which results for the truncated balanced realization method is stabilizable if the original system is approximately controllable.

Corollary(4.5.2) If the original system is approximately controllable, then the system which results from the truncated balanced realization scheme is stabilizable.

Proof:

By the above theorem, the system which results from the truncated balanced realization is controllable. In finite dimensions controllability implies stabilizability [Le-1].

4.6 Truncated Balanced Realization for the Heat Equation

Consider the heat equation on the interval (0,1)

(4.6.1) $$y_t = y_{xx} + b(x)u(t) \qquad z(0)=z(1)=0$$

with observation

(4.6.2) $$r(t)=\int_0^1 c(x)y(x,t)dx.$$

We will use the orthonormal basis $\{\sin(n\pi x)\}$ for $L_2(0,1)$ to convert this into a system defined on ℓ^2, since the balanced realization of the system is defined on ℓ^2. Also, it will make the construction of truncated balanced systems easier to follow. Clearly, we could simply think of the problem on $L_2(0,1)$ with the Fourier coefficents in ℓ^2 where

(4.6.3) $$y_n(t)=\int_0^1 y(x,t)\sin(n\pi x)dx,$$

(4.6.4) $$b_n = \int_0^1 b(x)\sin(n\pi x)dx,$$

(4.6.5) $$c_n = \int_0^1 c(x)\sin(n\pi x)dx.$$

If $y(t) = \{y_n(t)\}$, then Equations (4.6.1) - (4.6.2) become

(4.6.6) $$\dot{y}(t) = \begin{bmatrix} -\pi^2 & 0 & 0 & 0 \\ 0 & \ddots & & 0 \\ 0 & & -n^2\pi^2 & 0 \\ 0 & 0 & 0 & \ddots \end{bmatrix} \begin{bmatrix} y_1(t) \\ \vdots \\ y_n(t) \\ \vdots \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \\ \vdots \end{bmatrix} u(t)$$

and

(4.6.7) $$C(y(t)) = \begin{bmatrix} c_1 & c_2 & \cdots \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \end{bmatrix}.$$

From the results in [C-1] we see that the above matrix operator forms a C_0−semigroup $T(t)$ on ℓ^2 which has the following representation

(4.6.8)
$$T(t)y=\begin{bmatrix} e^{-\pi^2 t} & 0 & \cdots & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & & e^{-n^2\pi^2 t} & \\ 0 & \cdots & 0 & \ddots \end{bmatrix}\begin{bmatrix} y_1 \\ \vdots \\ y_n \\ \vdots \end{bmatrix}.$$

Thus, we have Γ defined by (3.2.14) is given by

(4.6.9)
$$\Gamma u(t) = \int_0^\infty \sum_{n=1}^\infty b_n c_n e^{-n^2\pi^2(t+s)} u(s)ds.$$

For the rest of this section assume that the coefficients b_n, c_n satisfy

(4.6.10)
$$b_n c_n > 0.$$

This assumption is not neccessary for the constuction of the truncated balanced realization matrices. However, it reduces the problem to finding a singular value decomposition of a large matrix to that of finding the eigenvalues of a large positive symmetric matrix. This hypothesis applies to the problem which we consider below.

Theorem(4.6.1) If $b_n c_n > 0$ for $n=1,2,\ldots$, then Γ is a nonnegative self–adjoint compact operator on $L_2(0,\infty;\mathbf{R}^1)$.

Proof:

The kernel is symmetric so self-adjointness follows from material on page 23 [H-1]. Since $\{b_n\}$ and $\{c_n\}$ are in ℓ^2, we can switch the order of integration and summation. When we do this and take the inner product on $L_2(0,\infty)$ we get

$$(4.6.11) \qquad <u,\Gamma u>_{L_2(0,\infty)} = \sum_{n=1}^{\infty} b_n c_n \int_0^\infty \int_0^\infty e^{-n^2\pi^2(t+s)} u(s)u(t)dsdt \geq 0.$$

Thus, Γ is nonnegative. The compactness of Γ follows from Theorem(3.2.4) Section 3.2. (It is also easy to see that we can write Γ as the limit of finite rank operators in the uniform operator topology.)

In order to compute the realization given in Section 3.2 we need to compute the eigenvalues and eigenvectors of the operator $\Gamma\Gamma^*$. Therefore, we could use numerical methods to approximate the eigenvalues and eigenvectors of $\Gamma\Gamma^*$. Since Γ defined on $L_2(0,\infty,\mathbb{R}^1)$ is self-adjoint when hypothesis (4.6.10) is satisfied, the eigenvectors of $\Gamma\Gamma^*$ are the same as the eigenvectors of Γ. Also, the eigenvalues of $\Gamma\Gamma^*$ are simply the square of the eigenvalues of Γ. Therefore, it is necessary only to compute the eigenvalues and eigenvectors of Γ. Since Γ is compact we will use the following theorem dealing with compact operators.

Theorem(4.6.2) Suppose $T_n:H \rightarrow H$ and $T:H \rightarrow H$ are compact self-adjoint operators on a Hilbert space H and T_n converges to T in the uniform operator topology. If λ is an eigenvalue of T, then there exists an eigenvector x for λ and a subsequence T_{n_k} of T_n with eigenvalues λ_{n_k} and corresponding eigenvectors x_{n_k} such that

(4.6.12) $$\lambda_{n_k} \to \lambda,$$

and

(4.6.13) $$\|x_{n_k} - x\| \to 0.$$

Proof:

By Theorem 4.10 on page 291 of [K-1] there is a sequence λ_n which is composed of eigenvalues of T_n such that $\lambda_n \to \lambda$. By Theorem 5.10 on page 236 of [Ch-1] we have the existence of the subsequence of eigenvectors and eigenvalues.

The proof of the following theorem follows from the previous theorem and induction on the algebraic order of the eigenvalue.

Theorem(4.6.3) Suppose $T_n:H \to H$ and $T:H \to H$ are compact self-adjoint operators on a Hilbert space H. If T_n converges to T in the uniform operator topology and $\lambda \neq 0$ is an eigenvalue of T with multiplicity k, $1 \leq k < \infty$, then there is a subsequence T_{n_l}, of the sequence of operatorsT_n, with k eigenvalues $\lambda_{n_{l_i}}$ $1 \leq i \leq k$ and corresponding eigenvectors $x_{n_{l_i}}$, $\lambda_{n_{l_i}}$ not necessarily distinct, which converge to k eigenvectors of T which span the eigenspace for the eigenvalue λ.

Theorem(4.6.4) If the operator Γ is self-adjoint, then $T_{i,j}(t) = T_{j,i}(t)$ for the balanced system.

Proof:

By Lemma 3.1 of [C-2]

(4.6.14) $$\sigma_j\int_0^\infty w_i(s)w_j(s+t)ds = \sigma_i\int_0^\infty w_i(s+t)w_j(s)ds .$$

Thus,

(4.6.15) $$\sqrt{\frac{\sigma_j}{\sigma_i}}\int_0^\infty w_i(s)w_j(t+s)ds = \sqrt{\frac{\sigma_i}{\sigma_j}}\int_0^\infty w_j(s)w_i(t+s)ds .$$

Theorem(4.6.5) If Γ is self-adjoint, then $A_{i,j}=A_{j,i}$ for the balanced system.

Proof:

By Equation 3.12 of [C-2] the result follows directly from the above theorem.

Corollary(4.6.6) If $\Gamma:L_2(0,\infty,\mathbf{R})\rightarrow L_2(0,\infty,\mathbf{R})$ is selfadjoint the value of $A_{i,j}$ is given by

(4.6.16) $$A_{i,j}=\frac{-s_{i,j}}{1+s_{i,j}^2}w_i^*(0)w_j(0),$$

where w_k is the k-th eigenvector of Γ with corresponding eigenvalue σ_k and

(4.6.17) $$s_{i,j}=\sqrt{\frac{\sigma_j}{\sigma_i}}.$$

Proof:

By equation (3.2.6) $w_i(s),w_j(s)\epsilon C^1(0,\infty;\mathbf{R}^k)$; therefore integration by parts makes sense.

By equation (3.2.20)

(4.6.18) $$A_{i,j}=\sqrt{\frac{\sigma_j}{\sigma_i}}\int_0^\infty w_i^*(s)\dot{w}_j(s)ds=-\sqrt{\frac{\sigma_j}{\sigma_i}}w_i^*(0)w_j(0)-\sqrt{\frac{\sigma_j}{\sigma_i}}\int_0^\infty \dot{w}_i^*(s)w_j(s)ds$$

(4.6.19) $$= -\sqrt{\frac{\sigma_j}{\sigma_i}}w_i^*(0)w_j(0) - \frac{\sigma_j}{\sigma_i}\sqrt{\frac{\sigma_i}{\sigma_j}}\int_0^\infty w_j^*(s)\dot{w}_i(s)ds$$

(4.6.20) $$= -\sqrt{\frac{\sigma_j}{\sigma_i}}w_i^*(0)w_j(0) - \frac{\sigma_j}{\sigma_i}A_{j,i}.$$

The result follows from the fact $A_{i,j} = A_{j,i}$.

Using the above results it is possible to calculate the value of the balanced realization by simply computing the eigenvalues and the value of the corresponding eigenvectors when t=0. However, this requires computation of the eigenvalues and eigenvectors for (3.5.11). However, the above theory can be used to get approximations of the eigenvalues and eigenvectors of (3.5.11) by computing the eigenvalues and eigenvectors of Γ_N where

(4.6.21) $$\Gamma_N u(t) = \int_0^\infty \sum_{n=1}^{N} b_n c_n e^{-n^2\pi^2(t+s)} u(s)ds.$$

This is equivalent to solving an integral equation with degenerate kernel. Thus, we can use the techniques outlined on page 37 of [H-1] to solve this problem. We need the following notation:

(4.6.22) $$g_n(t)=\sqrt{b_n c_n}\,e^{-n^2\pi^2 t}$$

(4.6.23) $$R^N=[r_{i,j}]^N=\left[\langle g_i(t),g_j(t)\rangle\right]^N=\left[\int_0^\infty \sqrt{b_ib_jc_ic_j}\,e^{-\pi^2(i^2+j^2)t}\right]^N,\ 1\leq i,j\leq N.$$

The eigenvalues of Γ_N are the eigenvalues of the matrix R^N, (see [H-1] page 38), and the eigenvectors $w_k^N(t)$ of Γ_N are:

(4.6.24) $$w_k^N(t)=\sum_{n=1}^N v_{k,n}^N\sqrt{b_nc_n}\,e^{-n^2\pi^2t},$$

where

(4.6.25) $$V_k^N=\begin{bmatrix} v_{k,1}^N \\ v_{k,2}^N \\ \vdots \\ v_{k,N}^N \end{bmatrix}$$

is the k-th eigenvector of the matrix R^N.

The construction of the matrices in Section 3.2 required normalized eigenvectors. From (3.5.31) the norm of $w_k^N(t)$ is given by:

(4.6.26) $$\left|w_k^N(t)\right|^2=\sum_{i=1}^N\sum_{j=1}^N v_{k,i}^N v_{k,j}^N\int_0^\infty g_i(t),g_j(t)dt.$$

Using Equation (4.6.33) the right hand side of Equation (3.5.32) can be represented as

(4.6.27) $$V_k^{N*}R^NV_k^N=\sigma_k^N.$$

Therefore, if we assume the V_k^N's are already normalized, the normalized $w_k^N(t)$ denoted $\tilde{w}_k^N(t)$ is given by:

$$\tilde{w}_k^N(t) = \frac{1}{\sqrt{\sigma_k^N}} \sum_{n=1}^{N} V_{k,n}^N \sqrt{b_n c_n} e^{-n^2 \pi^2 t} \tag{4.6.28}$$

Thus, M dimensional approximations of the matrices defined in Section 4.3 for the balanced system can be characterized as follows:

$$A_M^N = \left[\frac{-1}{\sigma_i^N + \sigma_j^N} \left(\sum_{n=1}^{N} V_{i,n}^N \sqrt{b_n c_n}\right)\left(\sum_{n=1}^{N} V_{j,n}^N \sqrt{b_n c_n}\right) \right], \ 1 \leq i,j \leq M, \tag{4.6.29}$$

and

$$B_M^N = [C_M^N]^* = \left[\sum_{n=1}^{N} V_{i,n} \sqrt{b_n c_n} \right], \ 1 \leq i \leq M. \tag{4.6.30}$$

Since Γ is a compact self-adjoint operator and the Γ_N operators are self-adjoint and converge to Γ, Theorem(4.6.3) implies that we should be able to get approximations of any finite number of eigenvalues and corresponding eigenvectors of Γ by simply taking N large enough and finding eigenvalues and eigenvectors of Γ_N. Since the the values of the the truncated balanced matrices depend on these values, we have that the truncation of the balanced realization of the system defined on ℓ^2 can be approximated to any degree of accuracy by simply finding the eigenvalues and eigenvectors of a symmetric matrix (equation (4.6.33)) and then applying equations (4.6.38) and (4.6.39).

Chapter V

5.1 Introduction

In this chapter numerical results will be presented for the items discussed in the preceding chapters. The numerical schemes were implemented on a VAX 8800 and a Cray-II. In Section 5.2 there is a discussion on the finite dimensional balancing of the finite element matrices. Section 5.3 will deal with the implementation of the algorithm detailed in Section 4.6 for construction of the matrices for the truncated balanced realization scheme. Included in Section 5.3 is a listing of the various checks which were made against theory to ensure the computer program was working properly and that the matrices which were generated were, in fact, very close to the actual matrices. Also, there is a discussion on the speed of convergence of the numerical approximations to the truncated balanced realization for the M=10 case. In Section 5.4 a comparison of the measures of controllability, observability, stabilizability, and detectability the matrices mentioned above will be made.

5.2 Finite Element

The FORTRAN code used to generate the M dimensional approximations by the finite element method was implemented on the VAX 8800. The FORTRAN code used to balance the resulting systems was also implemented on the VAX 8800. The results we compared against the results in [Bu-2]. Also, for some of the smaller matrices the balancing algorithm was compared to the balancing algorithm in Matlab. Also, the matrices generated by the balancing routine satisfy the theoretical properties of Section 3.3. The resulting system for the M = 5 finite element approximation of the system

(5.2.1) $$z_t = z_{xx} + xu(t), \quad 0 \le x \le 1,$$

(5.2.2) $$c(t) = \int_0^1 x^2 z(x,t)dx,$$

(5.2.3) $$z(0,t) = z(1,t) = 0.$$

is given in TABLE 1 and the balanced finite element system is given in TABLE 2.

TABLE 1

Finite Element Method

$$A_5 = \begin{bmatrix} -1.31 & 93.05 & -24.92 & 6.646 & -1.662 \\ 93.05 & -156.2 & 99.69 & -26.58 & 6.646 \\ -24.92 & 99.69 & -157.8 & 99.69 & -24.92 \\ 6.646 & -26.58 & 99.69 & -156.2 & 93.05 \\ -1.662 & 6.646 & -24.92 & 93.05 & -131.3 \end{bmatrix}$$

$$B_5 = \begin{bmatrix} 0.1679 \\ 0.3282 \\ 0.5192 \\ 0.5949 \\ 1.101 \end{bmatrix}$$

$$C_5 = \begin{bmatrix} 0.1944 & 0.6944 & 1.528 & 2.694 & 4.194 \end{bmatrix}$$

TABLE 2

Balanced Finite Element

$$A_5 = \begin{bmatrix} -12.78 & -11.22 & -4.424 & -1.185 & 0.216 \\ -11.22 & -62.57 & -44.76 & -12.64 & 2.31 \\ -4.424 & -44.76 & -150.9 & -76.80 & 14.68 \\ -1.185 & -12.64 & -76.80 & -214.4 & 76.27 \\ 0.216 & 2.31 & 14.68 & 76.27 & -292.1 \end{bmatrix}$$

$$B_5 = \begin{bmatrix} -2.420 \\ -1.108 \\ -0.4201 \\ -0.1123 \\ 0.02046 \end{bmatrix}$$

$$C_5 = \begin{bmatrix} -2.420 & -1.108 & -0.4201 & -0.1123 & 0.02046 \end{bmatrix}$$

5.3 Truncated Balanced Realization Method

The FORTRAN code used to generate the M dimensional approximations of the balanced realization was implemented on the Cray-2. The subroutine RS from the Eispack library of FORTRAN subroutines was called to find the eigenvalues and eigenvectors of a real symmetric matrix. In order to ensure the matrices generated by this code were close to the actual matrices, the following checks were implemented: (1) The generated A_M matrices were checked for symmetry which was mandated by Corollary(4.6.6); (2) The B_M and C_M matrices were checked to see if $B^*_M = C_M$ as mandated by equation (4.6.36); (3) The eigenvalues of the A_M matrix were computed to see how they compared with the infinite dimensional eigenvalues; (4) The A_M, B_M, and C_M matrices were run through a finite dimensional balancing algorithm to see if this caused many changes (i.e. to see if the system was in fact balanced); (5) The singular values of the controllability and observability grammians of the system obtained by balancing A_M, B_M, and C_M were compared against the singular values used to construct the system (equation (4.3.5) requires that they be the same). In TABLE 3 the matrices A^N_M, B^N_M, and C^N_M are given for the case M=5 and N = 1500 for the control problem defined by equations (5.2.1)-(5.2.3). The time required by the Eispack subroutine RS, see [S-1], to find the eigenvalues and eigenvectors for R^N (equation (4.6.33)) was 280 seconds on the Cray-2. To get convergence for the M = 10 system the size of R^N that had to be used was 1500x1500. However, it may have been possible to decrease the time if another method had been used to find the eigenvalues and eigenvectors of Γ. The method used in Section 4.6 was chosen because it had only the one possibility for the introduction of numerical error.

In TABLE 4 the singular values of the grammians for a 10x10 truncated balanced realizations $(A_{10}^N, B_{10}^N, C_{10}^N)$ which were generated using R^N N=100,...,1000 are given. The singular values of the gramians for M=5 are given in TABLE 5. The infinity norm of the difference between the truncated balanced realizations (A_M^N, B_M^N, C_M^N) generated using R^N for N=100,...,1000 and R^N for N=1500, with M=10, are given in TABLE 6. The M=5 results are given in TABLE 7. Note that A_5^N converges to A_5 monotonically and this is not true for the larger system A_{10}^N.

TABLE 3

Truncated Balanced Realization Method

$$A_5 = \begin{bmatrix} -12.10 & -11.08 & -6.743 & 4.474 & 3.171 \\ -11.08 & -75.49 & -81.57 & 58.09 & 41.59 \\ -6.743 & -81.57 & -289.7 & 335.1 & 296.3 \\ 4.474 & 58.09 & 335.1 & -879.4 & -1046 \\ 3.171 & 41.59 & 264.3 & -1046 & -2307 \end{bmatrix}$$

$$B_5 = \begin{bmatrix} -0.3235 \\ -0.1535 \\ -0.9044 \\ 0.05984 \\ 0.04240 \end{bmatrix}$$

$$C_5 = \begin{bmatrix} -0.3235 & -0.1535 & -0.9044 & 0.05984 & 0.04240 \end{bmatrix}$$

TABLE 4

Singular Values, σ_i^N, of Grammians for the System

$(A_{10}^N, B_{10}^N, C_{10}^N)$

i	N=100	N=500	N=1000	N=1500	σ_i
1	$4.325x10^{-3}$	$4.324x10^{-3}$	$4.325x10^{-3}$	$4.325x10^{-3}$	$4.325x10^{-3}$
2	$1.561x10^{-4}$	$3.884x10^{-6}$	$3.778x10^{-6}$	$1.561x10^{-4}$	$1.561x10^{-4}$
3	$1.142x10^{-5}$	$9.258x10^{-8}$	$9.004x10^{-8}$	$1.412x10^{-5}$	$1.412x10^{-5}$
4	$2.036x10^{-6}$	$2.977x10^{-8}$	$2.908x10^{-8}$	$2.036x10^{-6}$	$2.036x10^{-6}$
5	$3.895x10^{-7}$	$1.269x10^{-8}$	$1.244x10^{-8}$	$3.896x10^{-7}$	$3.896x10^{-7}$
6	$9.022x10^{-8}$	$4.026x10^{-9}$	$3.956x10^{-9}$	$9.039x10^{-8}$	$9.039x10^{-8}$
7	$2.388x10^{-8}$	$1.810x10^{-9}$	$1.780x10^{-9}$	$2.415x10^{-8}$	$2.415x10^{-8}$
8	$6.683x10^{-9}$	$8.914x10^{-10}$	$8.837x10^{-10}$	$7.191x10^{-9}$	$7.191x10^{-9}$
9	$2.032x10^{-9}$	$3.404x10^{-10}$	$3.365x10^{-10}$	$2.336x10^{-9}$	$2.336x10^{-9}$
10	$5.970x10^{-10}$	$5.311x10^{-11}$	$5.257x10^{-11}$	$8.148x10^{-10}$	$8.148x10^{-10}$

TABLE 5

Singular Values, σ_i^N, of Grammians for the system

(A_5^N, B_5^N, C_5^N)

i	N=100	N=500	N=1000	N=1500	σ_i
1	4.325×10^{-3}	4.324×10^{-3}	4.325×10^{-3}	4.325×10^{-3}	4.325×10^{-3}
2	1.561×10^{-4}	2.971×10^{-5}	2.971×10^{-5}	1.561×10^{-4}	1.561×10^{-4}
3	1.142×10^{-5}	2.693×10^{-8}	2.693×10^{-6}	1.412×10^{-5}	1.412×10^{-5}
4	2.036×10^{-6}	5.720×10^{-7}	5.720×10^{-7}	2.036×10^{-6}	2.036×10^{-6}
5	3.895×10^{-7}	2.075×10^{-7}	2.075×10^{-7}	3.896×10^{-7}	3.896×10^{-7}

TABLE 6

Infinity Norm Difference 10x10

N	$\left\|A_{10}^{N}-A_{10}^{1500}\right\|_{\infty}$	$\left\|B_{10}^{N}-B_{10}^{1500}\right\|_{\infty}$	$\left\|C_{10}^{N}-C_{10}^{1500}\right\|_{\infty}$
100	29901	0.12	0.12
200	37956	0.12	0.12
300	18136	0.12	0.12
400	9216	0.12	0.12
500	5089	0.12	0.12
600	116985	0.024	0.024
700	117487	0.024	0.024
800	117780	0.025	0.025
900	795	$5.3x10^{-5}$	$5.3x10^{-5}$
1000	525	$3.5x10^{-5}$	$3.5x10^{-5}$

TABLE 7

Infinity Norm Difference 5x5

N	$\|A_5^N - A_5^{1500}\|_\infty$	$\|B_5^N - B_5^{1500}\|_\infty$	$\|C_5^N - C_5^{1500}\|_\infty$
100	37	0.12	0.12
200	4.9	0.12	0.12
300	1.47	0.12	0.12
400	0.62	0.12	0.12
500	0.31	0.12	0.12
600	0.174	1.6×10^{-6}	1.6×10^{-6}
700	0.11	1.2×10^{-6}	1.2×10^{-6}
800	0.067	6.1×10^{-7}	6.1×10^{-7}
900	0.044	4.0×10^{-7}	4.0×10^{-7}
1000	0.029	2.6×10^{-7}	2.6×10^{-7}

5.4 A Comparison of the System Measures

The method used to compute the measures of controllability and observability was outlined in Section 2.3. The FORTRAN programs which were used to compute these measures were run on the VAX 8800. The results of these programs for the finite element method were checked against the results in [Bu-2]. The results for the problem given in equations (5.2.1)-(5.2.3) are given in TABLE 8. It is interesting to note how well the finite element method compares to the truncated balanced realization method. In fact, the measure of controllability for the finite element method is almost an order of magnitude higher than the measure of controllability for the truncated balanced realization method. In addition, the balanced finite element model is an order of magnitude more "robust" than the truncated balanced realization model.

The measures of stabilizability and detectability were computed using the algorithm outlined in Section 4.2. The results presented here are consistent with the results given in [Bu-2]. The results are given in TABLE 9 are for the problem defined by equations (5.2.1)-(5.2.3). Note that $\gamma^2 = \pi^2$ is the measure of stabilizability (and dectability) for the infinite dimensional problem and the truncated balanced realization reflects this property.

TABLE 8

Measures of Controllability and Observability

	Controllability			Observability		
N	F.E	B.F.E.	T.B.R.	F.E.	B.F.E.	T.B.R.
2	0.2221	0.6659	0.04223	1.1993	0.6659	0.04223
3	0.1155	0.2828	0.02112	0.6923	0.2828	0.02212
4	0.06698	0.1595	0.01166	0.3798	0.1595	0.01166
5	0.04187	0.09302	0.006921	0.2066	0.09302	0.006921
6	0.02773	0.06094	0.006859	0.1339	0.06094	0.006859
7	0.01922	0.04091	0.004459	0.08706	0.04091	0.004459
8	0.01383	0.02932	0.002998	0.06218	0.02932	0.002998
9	0.01026	0.02137	0.003159	0.04453	0.02137	0.003159
10	0.007811	0.01625	0.001466	0.03383	0.01625	0.001466

TABLE 9

Measures of Stabilizability and Detectability

	Stabilizability			Detectability		
N	F.E	B.F.E.	T.B.R.	F.E.	B.F.E.	T.B.R.
2	10.83	10.90	10.22	11.09	10.90	10.22
3	10.43	10.52	9.949	10.82	10.52	9.949
4	10.25	10.37	9.893	10.77	10.37	9.893
5	10.16	10.31	9.879	10.80	10.31	9.879
6	10.11	10.28	9.875	10.87	10.28	9.875
7	10.08	10.28	9.874	10.95	10.28	9.874
8	10.06	10.29	9.874	11.04	10.29	9.874
9	10.05	10.31	9.874	11.14	10.31	9.874
10	10.05	10.33	9.874	11.24	10.33	9.874

5.5 Conclusions

The truncated balanced realization method was applied to control a system governed by the heat equation and compared to the finite element and balanced finite element methods. The truncated balanced realization method requires considerable effort be devoted to analytical preparation and large amounts of computational times simply to construct the system matrices (c.f. Section 5.3). Conversely, the finite element method requires little analytical preparation and takes almost no computer time to construct the system matrices. The balanced finite element method uses a moderate amount of computational resources in applying the balancing alogrithm outlined in [L-1]. If we compare the measures of controllability and observability given in TABLE 7, we see that the balanced finite element model is more robust under perturbation than the finite element model, and the finite element model is more robust under perturbation than truncated balanced realization model. The measures of stabilizability and detectability listed in TABLE 8 are about the same for each of the models. The magnitude of the matrix entries for the truncated balanced realization model (TABLE 3) varies more than the magnitude of the corresponding entries for either the finite element model (TABLE 1) or the balanced finite element model (TABLE 2). Thus, the truncated balanced realization gives approximations which are numerically "stiffer" than the approximations from the finite element model or the balanced finite element model. Therefore, we conclude that either the finite element method or the balanced finite element method provides a more robust finite dimensional model for controlling the heat equation.

This study raises many unanswered questions concerning the appropriate approach to modeling control systems described by partial differential equations. In particular, it is not known if the truncated balanced realization is always less robust than the finite element scheme or if this is only true for the problem (or class of problems) considered here.

References

Ba-1 Banks, H.T. and Kunisch, K., The Linear Regulator Problem for Parabolic Systems, SIAM Journal of Control and Optimization, Vol. 22, N0. 5, September 1984

Bu-1 Burns, J.A., Peichl, G.H., A note on the asympotic behavior of controllability radii for a scalar herediatry system, submitted

Bu-2 Burns, J.A., Peichl, G.H., On Robustness and Controllability for Finite Dimensional Approximations of Distributed Parameter Systems, Distributed Parameter Systems: Modelling and Simulation, IMACS Transactions on Scientific Computing, to appear. Also in IMACS/IFAC Symposium on Distributed Parameter Systems, Hiroshima, Japan, October 1987, 491-496.

Bu-3 Burns, J.A., Peichl, G.H., Open questions concerning appxoximation and control of infinite dimensional systems, Preprint

Bu-4 Burns, J.A., Peichl, G.H., Preservation of Controllability under Approximation and Controllability Radii for Hereditary Systems, Journal of Differential and Integral Equations, to appear.

Bu-5 Burns, J.A., Peichl, G.H., System Radii of Semi-Discretized Hereditary Control Systems, 27th IEEE Conference on Decision and Control, 1988, 2003-2005.

Ch-1 Chatelin, F., Spectral Approximation of Linear Operators, Academic Press, New York, (1983).

C-1 Curtain, R.F. and Pritchard, A.J., Infinite Dimensional Linear Systems Theory, Springer, Berlin (1978).

C-2 Curtain, RF. and Glover, K., Realization and Approximation of Linear Infinite-Dimenisional Systems With Error Bounds, SIAM Journal of Control and Optimization, Vol. 26, N0. 4, July 1988.

E-1 Eising, R., Between controllable and uncontrollable, Systems & Control Letters, Vol 4, No 5, July 1984, 263-264.

E-2 Eising, R., The distance between a system and the set of uncontrollable systems, proc. Mtns, Beer Sheva, Lecture Notes in Control and InformationSciences 8, 1978, Springer, New York

F-1 Franklin, J.N., Matrix Theory, Prentice-Hall, Englewood Cliffs, NJ (1968).

G-1 Gibson, J.S., The Riccati Integral Equations for Optimal Control Problems on Hilbert spaces, Siam J Control and Optimization, 17 (1979), pp. 537-565.

Go-1 Golub, G. and Van Loan, C., Matrix Computations, The Johns Hopkins University Press, Baltimore, (1983).

H-1 Hochstadt, H., Integral Equations, John Wiley & Sons, New York (1973)

K-1 Kato, T., Perturbation Theory for Linear Operators, 2nd ed., Springer-Verlag, Berlin (1976).

L-1 Laub, A.J., Heath, M.T., Pagie, C.C., and Ward, R.C., Computation of System Balancing Transformations, Proodeedings of the 25th Conference on Decision and Control, Athens, Greece December 1986

Le-1 Lee, E.B., Markus, L., Foundations of Optimal Control Theory, John Wiley & Sons, New York (1967).

M-1 Moore, B.C., Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction, Transactions on Automatic Control Vol. AC-26 No. 1, February 1981

P-1 Pazy, A., Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer-Verlag, New York (1983).

R-1 Reddy, J.N., An Introduction to the Finite Element Method, McGraw-Hill, New York (1984).

S-1 Smith, B.T., Boyle, J.M., Dongarra, J.J., Garbow, B.S., Ikebe, Y., Klema, V.C. and Moler, C.B., Matrix Eigensystem Routines - EISPACK Guide, 2nd ed., Springer-Verlag, Berlin (1976).

W-1 Weidmann, J., Linear Operators in Hilbert Spaces, Springer-Verlag, New York, (1980).

The vita has been removed from the scanned document